ECEN 743 | Homework 3

Muhammed U. Ersoy
M.S. ECEN Student
Texas A&M University
mue@tamu.edu

Question 1

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

Let $u = x - y$ and $v = y$ then using $\triangle$·Inequality;

$$||u + v|| \leq ||u|| + ||v||$$

$$||(x - y) + y|| \leq ||(x - y)|| + ||y||$$

$$||x|| \leq ||x - y|| + ||y||$$

$$||x|| - ||y|| \leq ||x - y||$$

Question 2

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

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

$$\begin{aligned} V_\pi(s) &\doteq \mathbb{E}_\pi[\sum_{t=0}^{\infty} \gamma^t r(s, a) | S_t = s, A_t = a] \\ &= \sum_a \pi(s, a) \sum_{s'} P(s'|s, a) \sum_{t=0}^{\infty} \gamma^t r(s, a) \end{aligned} \tag{2}$$

We can take supremum of both sides which just equals to $\sup\limits_{\pi} V_\pi = V^*$, after this we can see that the optimal value function is bounded by the case where maximum reward was gained at every step. To show this we can take $\max_{s,a}$ of both sides.

$$\sup_{\pi} \max_{s,a} V_\pi = ||V^*||_\infty \leq \sum_{t=0}^{\infty} \gamma^t \max_{s,a} |r(s, a)| \tag{3}$$

Since $\gamma \leq 1$ we can simply the sum operator using geometric series;

$$\sup_{\pi} ||V_\pi||_\infty \leq \frac{\max_{s,a} |r(s, a)|}{(1 - \gamma)} \tag{4}$$

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Question 3

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

Let's consider the Bellman optimality equation;

$$TV = R_\pi + \gamma \sum_{s' \in S} P_\pi(s|s')V(s') \tag{5}$$

if $TV_1 \geq TV_2$ then $TV_1 - TV_2 \geq 0$;

$$\begin{aligned} R_\pi + \gamma \sum_{s' \in S} P_\pi(s|s')V_1(s') - R_\pi - \gamma \sum_{s' \in S} P_\pi(s|s')V_2(s') \geq 0 \\ \gamma \sum_{s' \in S} P_\pi(s|s')(V_1(s') - V_2(s')) \geq 0 \end{aligned} \tag{6}$$

and since $V_1 \geq V_2$ and P_π is bounded by $[0, 1]$ (only positive values) this expression must hold true. Hence bellman operator is a monotone operator.

Question 4

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 $||\cdot||_\infty$.

Let's remember the definition of a contraction;

$$||Au - Av||_\infty \leq \alpha ||u - v||_\infty \tag{7}$$

where $0 \leq \alpha < 1$. We can refactor the left side as $Au - Av = A(u - v)$.Then recall that the supremum norm of a vector is simplify the maximum absolute element of the vector. Then the left side can be expressed as;

$$\lfloor A(u - v) \rfloor_i = \sum_j a_{ij}(u_i - v_i) \tag{8}$$

Taking the absolute value of both sides we get the following inequality;

$$|\lfloor A(u - v) \rfloor_i| \leq \sum_j |a_{ij}||(u_i - v_i)| \tag{9}$$

and since the the supremum norm is the maximum absolute element of a vector we can add the following;

$$|\lfloor A(u - v) \rfloor_i| \leq ||A(u - v)||_\infty \leq \sum_j |a_{ij}|||(u - v)||_\infty \tag{10}$$

From the problem statement we know $\sum_j |a_{ij}| \leq \alpha < 1$;

$$||A(u - v)||_\infty \leq \alpha ||(u - v)||_\infty \tag{11}$$

Thusfort $f(\cdot)$ is a contraction mapping with respect to the supremum norm.

Question 5

Let U be a given set, and $g_1 : U \rightarrow \mathbb{R}$ and $g_2 : U \rightarrow \mathbb{R}$ be two real-valued functions on U. Also assume that both functions are bounded. Show that

$$|\max_u g_1(u) - \max_u g_2(u)| \leq \max_u |g_1(u) - g_2(u)| \tag{12}$$

Assume that $\max_u g_1(u) \geq \max_u g_2(u)$. Let $u^* = arg \max_u g_1(u)$;

$$\begin{aligned} |\max_u g_1(u) - \max_u g_2(u)| &= g_1(u^*) - \max_u g_2(u) \\ &\leq g_1(u^*) - g_2(u^*) \end{aligned} \tag{13}$$

Based on our initial assumption this value must be necessarily positive thus we rewrite this inequality without loss of generality;

$$\begin{aligned} |\max_u g_1(u) - \max_u g_2(u)| &= g_1(u^*) - \max_u g_2(u) \\ &\leq |g_1(u^*) - g_2(u^*)| \\ &\leq \max_a |g_1(u) - g_2(u)| \end{aligned} \tag{14}$$

The same logic follows if we change our initial assumption to $\max_u g_2(u) \geq \max_u g_1(u)$. Thus the initial statement must be true;

$$|\max_u g_1(u) - \max_u g_2(u)| \leq \max_a |g_1(u) - g_2(u)| \tag{15}$$

Question 6

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 $n > m$

$$||V_m - V_n||_\infty \leq \frac{\gamma^m}{(1-\gamma)}||V_0 - V_1||_\infty \tag{16}$$

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

$$||V_m - V^*||_\infty \leq \frac{\gamma^m}{(1-\gamma)}||V_0 - V_1||_\infty \tag{17}$$

(c) Show that

$$||V_m - V^*||_\infty \leq \frac{\gamma}{(1-\gamma)}||V_{m-1} - V_m||_\infty \tag{18}$$

(a)

$$\begin{aligned} ||V_m - V_n||_\infty &= ||T^m V_0 - T^n V_0||_\infty \\ &\leq \gamma^n ||T^{m-n} x_0 - x_0||_\infty \end{aligned} \tag{19}$$

We can expand the right side as a sum;

$$\begin{aligned} &\leq \gamma^m [||T^{n-m}V_0 - T^{n-m-1}V_0||_\infty + \cdots + ||TV_0 - V_0||_\infty] \\ &\leq \gamma^m [\sum_{k=0}^{m-n-1} \gamma^k] ||V_1 - V_0||_\infty \\ &\leq \gamma^m [\sum_{k=0}^{\infty} \gamma^k] ||V_1 - V_0||_\infty \end{aligned} \tag{20}$$

Using geometric series we get;

$$||V_m - V_n||_\infty \leq \frac{\gamma^m}{1-\gamma}||V_0 - V_1||_\infty \tag{21}$$

(b) Optimal Value function is the unique fixed-point of the bellman equation according to the banach fixed point theorem.

$$V^* = TV^* = T^n V_0,\ n \geq 0 \tag{22}$$

Then based on (a), for $n > m$ the following holds;

$$||V_m - V^*||_\infty = ||V_m - V_n||_\infty \leq \frac{\gamma^m}{1-\gamma}||V_0 - V_1||_\infty \tag{23}$$

(c)

$$||V_m - V_n||_\infty \leq \frac{\gamma^m}{1-\gamma}||V_0 - V_1||_\infty = \frac{\gamma^m}{1-\gamma}||V_0 - TV_0||_\infty \tag{24}$$

Let $n \to \infty$.

$$||V_m - V^*||_\infty = ||TV_{m-1} - TV^*||_\infty \leq \gamma ||V_{m-1} - V^*||_\infty \tag{25}$$

We can apply $\triangle$·Inequality;

$$||V_m - V^*||_\infty \leq \gamma [||V_{m-1} - V_m||_\infty + ||V_m - V^*||_\infty \tag{26}$$

This can be re-arranged;

$$||V_m - V^*||_\infty \leq \frac{\gamma}{1-\gamma}||V_{m-1} - V_m||_\infty \tag{27}$$

Question 7

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

$$||V^* - V_{\overline{\pi}}||_\infty \leq \frac{2\epsilon}{1-\gamma} \tag{28}$$

We can fix state s and let $a = \pi_Q(s)$. Then;

$$\begin{aligned}
V^*(s) - V_{\overline{\pi}}(s) &= Q^*(s, \pi^*(s)) - Q_{\overline{\pi}}(s, a) \\
&= Q^*(s, \pi^*(s)) - Q^*(s, a) + Q^*(s, a) - Q_{\overline{\pi}}(s, a) \\
&= Q^*(s, \pi^*(s)) - Q^*(s, a) + \gamma \mathbb{E}_{s' P(\cdot|s,a)}[V^*(s') - V_{\overline{\pi}}(s')] \\
&\leq Q^*(s, \pi^*(s)) - \overline{Q}(s, \pi^*(s)) + \overline{Q}(s, a) - Q^*(s, a) \\
&+ \gamma \mathbb{E}_{s' P(\cdot|s,a)}[V^*(s') - V_{\overline{\pi}}(s')] \\
&\leq 2||\overline{Q} - Q^*||_\infty + \gamma ||V^* - V_{\overline{\pi}}||_\infty
\end{aligned} \tag{29}$$

We can take the supremum norm of both sides; and norm of a supremum norm is itself since the norm of a scalar is itself;

$$\begin{aligned}
||V^*(s) - V_{\overline{\pi}}(s)||_\infty &\leq ||2||\overline{Q} - Q^*||_\infty + \gamma ||V^* - V_{\overline{\pi}}||_\infty||_\infty \\
&= 2||\overline{Q} - Q^*||_\infty + \gamma ||V^* - V_{\overline{\pi}}||_\infty
\end{aligned} \tag{30}$$

We can use the fact that $||\overline{Q} - Q^*||_\infty \leq \epsilon$;

$$(1-\gamma)||V^*(s) - V_{\overline{\pi}}(s)||_\infty \leq 2||\overline{Q} - Q^*||_\infty \leq 2\epsilon \tag{31}$$

We can rearrange to get the inequality in the problem statement;

$$||V^* - V_{\overline{\pi}}||_\infty \leq \frac{2\epsilon}{1-\gamma} \tag{32}$$

Hello world