RL - Assignment 3

Nathani Jawahar Sai
735000830

1) Given $\|x+y\| \leq \|x\| + \|y\|$

let $x = x-y$ and $y = y$

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

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

$$\Rightarrow \|x-y\| \geq \|x\| - \|y\|.$$

Hence Proved

2) In MDP $V_\pi(s) = E\left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)\right]$ For any policy π

$$r(s_t, a_t) \leq |r(s_t, a_t)| \leq \max_{s,a}\left(|r(s_t, a_t)|\right)$$

$$\Rightarrow \quad E\left[\sum_{t=0}^{\infty} \gamma^t \cdot r(s_t, a_t)\right] \leq E\left[\sum_{t=0}^{\infty} \gamma^t \cdot \max_{s,a}\left(|r(s_t, a_t)|\right)\right]$$

$$= E\left[\max_{s,a}\left(|r(s_t, a_t)|\right) \sum_{t=0}^{\infty} \gamma^t\right]$$

$$= E\left[\max_{s,a}\left(|r(s_t, a_t)|\right) \cdot \frac{1}{1-\gamma}\right]$$

$$\left[\text{As } \sum^{\infty} \gamma^t = \frac{1}{1-\gamma} \quad \gamma < 1\right]$$

$$= \frac{\max_{s,a} |r(s,a)|}{1-\gamma}$$

* estimation of a constant is also a constant.

Therefore, for any π and s

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

* superior of a constant is also a constant.

3) We have $v_1 \geq v_2$ element-wise

$$\Rightarrow v_1(s) \geq v_2(s)$$

$$\Rightarrow \sum_{s'} P(s'|s,a)\, v_1(s') \geqslant \sum_{s'} P(s'|s,a)\, v_2(s')$$

$$\Rightarrow \gamma \sum_{s'} P(s'|s,a)\, v_1(s') \geqslant \gamma \sum_{s'} P(s'|s,a)\, v_2(s') \quad [\text{As } \gamma > 0]$$

$$\Rightarrow \underbrace{\max\left(r(s,a) + \gamma \sum_{s'} P(s'|s,a)\, v_1(s')\right)}_{(Tv_1)(s)} \geqslant \underbrace{\max\left(r(s,a) + \gamma \sum_{s'} P(s'|s,a) \cdot v_2(s')\right)}_{(Tv_2)(s)}$$

$$\Rightarrow (Tv_1)(s) \geqslant (Tv_2)(s)$$

* Therefore, if $v_1 \geqslant v_2$ element-wise $(Tv_1)(s) \geqslant (Tv_2)(s)$.

* Hence Bellman operator T is monotone.

4) Let $u, v \in \mathbb{R}^n$

$$\|Au - Av\|_\infty = \|A(u-v)\|_\infty = \max_{i,\, 1 \le i \le n} \left| \sum_{j=1}^{n} a_{ij}(u_j - v_j) \right|$$

[we know $|\Sigma xy| \le \Sigma |xy| = \Sigma |x||y|$]

$$\Rightarrow \max_{i,\, 1 \le i \le n} \left| \sum_{j=1}^{n} a_{ij}(u_j - v_j) \right| \le \max_{i,\, 1 \le i \le n} \sum_{j=1}^{n} |a_{ij}||u_j - v_j|$$

$$\le \max_{i,\, 1 \le i \le n} \sum_{j=1}^{n} |a_{ij}| \left(\|u-v\|_\infty \right)$$

$$= \max_{i,\, 1 \le i \le n} \|u-v\|_\infty \cdot \underbrace{\sum_{j=1}^{n} |a_{ij}|}_{\le \alpha}$$

$$\le \alpha \|u-v\|_\infty$$

$$\Rightarrow \|Au - Av\| \le \alpha \|u-v\|_\infty$$

Hence $f(u) = Au$ $R \to R$ is a contraction mapping with respect to $\|\cdot\|_\infty$

5)

Given $g_1 : U \to R$, $g_2 : U \to R$. Need to prove

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

$$\max_x (g_1(x)) = \max_x \left(g_1(x) + g_2(x) - g_2(x) \right)$$

$$= \max_x \left(g_2(x) + (g_1(x) - g_2(x)) \right)$$

$$a+b \leq a + |b|$$

$$\Rightarrow \quad \max_x (g_1(x)) \leq \max_x \left(g_2(x) + |g_1(x) - g_2(x)| \right)$$

$$= \max_x (g_2(x)) + \max_x \left(| g_1(x) - g_2(x) | \right)$$

$$\Rightarrow \quad \max_x (g_1(x)) - \max_x (g_2(x)) \leq \max_x \left(| g_1(x) - g_2(x) | \right) \quad \text{--- } ①$$

Similarly,

$$\max_x (g_2(x)) = \max_x (g_2(x) + g_1(x) - g_1(x))$$

$$\leq \max_x (g_1(x) + |g_2(x) - g_1(x)|)$$

$$|a-b| = |b-a|$$

$$\Rightarrow \max_x (g_2(x)) \leq \max_x (g_1(x) + |g_1(x) - g_2(x)|)$$

$$= \max_x (g_1(x)) + \max_x (|g_1(x) - g_2(x)|)$$

$$\Rightarrow \max_x (g_2(x)) - \max_x (g_1(x)) \leq \max_x (|g_1(x) - g_2(x)|) \quad (2)$$

From equation 1 and 2 we can write

$$\boxed{\left|\max_x (g_1(x)) - \max_x (g_2(x))\right| \leq \max_x (|g_1(x) - g_2(x)|)}$$

Hence Proved.

6)

a) Given $V_{k+1} = TV_k$, $n>m$ need to prove

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

$$\|V_k - V_{k+1}\|_\infty = \|TV_{k-1} - TV_k\|_\infty$$

As Bellman operator T is a contraction mapping we can write

$$\|TV_{k-1} - TV_k\|_\infty \leq \gamma \|V_{k-1} - V_k\|_\infty \quad [\text{for } \gamma < 1]$$

$$\Rightarrow \|V_k - V_{k+1}\| \leq \gamma \|V_{k-1} - V_k\|_\infty$$

$$= \gamma \|TV_{k-2} - TV_{k-1}\|_\infty$$

$$\leq \gamma.\gamma \|V_{k-2} - V_{k-1}\|_\infty$$

$$\vdots$$

$$\vdots$$

$$\leq \gamma^{k} \| v_0 - v_1 \|_\infty$$

$$\Rightarrow \quad \| v_k - v_{k+1} \| \leq \gamma^k \| v_0 - v_1 \|_\infty$$

And

$$\| v_m - v_n \|_\infty = \| v_m - v_{m+1} + v_{m+1} - v_{m+2} + v_{m+2} \cdots\cdots v_n \|_\infty$$

$$\left[\|a+b\| \leq \|a\| + \|b\| \right]$$

$$\leq \| v_m - v_{m+1} \| + \| v_{m+1} - v_{m+2} \| \cdots + \| v_{n-1} - v_n \|$$

$$= \sum_{k=m}^{n-1} \| v_k - v_{k+1} \|_\infty$$

$$\leq \sum_{k=m}^{n-1} \gamma^k \| v_0 - v_1 \|_\infty$$

$$\leq \| v_0 - v_1 \|_\infty \cdot \sum_{k=m}^{n-1} \gamma^k$$

$$= \| v_0 - v_1 \|_\infty \cdot \frac{\gamma^m (1 - \gamma^{n-m})}{(1-\gamma)}$$

$$\left[\begin{array}{l} 1 - \gamma^{n-m} < 1 \\ \Rightarrow \dfrac{\gamma^m (1 - \gamma^{n-m})}{1-\gamma} < \dfrac{\gamma^m}{1-\gamma} \end{array} \right]$$

$$\leq \|v_0 - v_1\|_\infty \cdot \frac{\gamma^m}{(1-\gamma)}$$

$$\Rightarrow \quad \|v_m - v_n\|_\infty \leq \frac{\gamma^m}{(1-\gamma)} \|v_0 - v_1\|_\infty$$

Hence proved.

b) According to value iteration v_0 converges to v^*.

$$\Rightarrow \lim_{n \to \infty} v_n = v^*$$

As $n = \infty > m$

$$\|v_m - v^*\|_\infty \leq \frac{\gamma^m}{1-\gamma} \|v_0 - v_1\|_\infty$$

c) $V_{k+1} = TV_k$

$$\Rightarrow \quad \|V_m - V^*\|_\infty = \|TV_{m-1} - TV^*\|_\infty$$

[$TV^* = V^*$]

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

[convergence property]

$$= \gamma \|V_{m-1} - V_m + V_m - V^*\|_\infty$$

$$\leq \gamma \left(\|V_{m-1} - V_m\|_\infty + \|V_m - V^*\|_\infty \right)$$

[Triangle Property]

$$\Rightarrow \quad \|V_m - V^*\|_\infty \leq \gamma \|V_{m-1} - V_m\|_\infty + \gamma \|V_m - V^*\|_\infty$$

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

Hence proved

7) Given $\|\bar{Q} - Q^*\|_\infty \leq e$

and $\bar{\pi} = \underset{a}{\operatorname{argmax}}(\bar{Q}(s,a))$

$\Rightarrow V_{\bar{\pi}} = Q^{-}(s, \bar{\pi}(s))$ $\quad [a = \bar{\pi}(s)]$

$\Rightarrow V^* = Q^*(s, \pi^*(s))$

$$\Rightarrow V^* - V_{\bar{\pi}} = Q^*(s, \pi^*(s)) - Q^{-}(s, \bar{\pi}(s))$$

$$= Q^*(s, \pi^*(s)) - Q^*(s, \bar{\pi}(s)) + Q^*(s, \bar{\pi}(s))$$
$$- Q^-(s, \bar{\pi}(s))$$

$$Q^*(s, \bar{\pi}(s)) - Q^-(s, \bar{\pi}(s)) = r(s,a) + \gamma E_{s' \sim P(s,a)}[V^*(s)]$$
$$- r(s,a) - \gamma E_{s' \sim P(s,a)}[V_{\bar{\pi}}(s)]$$
$$= \gamma E_{s' \sim P(s,a)}[V^*(s) - V_{\bar{\pi}}(s)]$$
$$\leq \gamma \left| E_{s' \sim P(s,a)}[V^*(s) - V_{\bar{\pi}}(s)] \right| \qquad [x \leq |x|]$$
$$\leq \gamma E_{s' \sim P(s,a)}[|V^*(s) - V_{\bar{\pi}}(s)|] \qquad \left[\begin{array}{l} \|x\|_\infty = \max(|x|) \\ \Rightarrow |x| \leq \|x\|_\infty \end{array} \right]$$
$$\leq \gamma E_{s' \sim P(s,a)}[\|V^*(s) - V_{\bar{\pi}}(s)\|_\infty]$$
$$= \gamma \|V^*(s) - V_{\bar{\pi}}(s)\|_\infty \qquad [E[\text{constant}] = \text{constant}]$$

$$\Rightarrow Q^*(s, \bar{\pi}(s)) - Q^-(s, \bar{\pi}(s)) \leq \gamma \|V^*(s) - V_{\bar{\pi}}(s)\|_\infty$$

$$\Rightarrow \quad V^* - V_{\bar{\pi}} \leq Q^*(s, \pi^*(s)) - Q^*(s, \bar{\pi}(s)) + \gamma \| V^*(s) - V_{\bar{\pi}}(s) \|_\infty$$

Apply superior on both sides.

$$\|V^* - V_{\bar{\pi}}\|_\infty \leq \| Q^*(s, \pi^*(s)) - Q^*(s, \bar{\pi}(s)) \|_\infty + \gamma \| V^*(s) - V_{\bar{\pi}}(s) \|_\infty$$

we know $\bar{Q}(s, \bar{\pi}(s)) \geq \bar{Q}(s, \pi^*(s))$

$$\Rightarrow \quad \|V^* - V_{\bar{\pi}}\|_\infty \leq \| Q^*(s, \pi^*(s)) - \bar{Q}(s, \pi^*(s)) + \bar{Q}(s, \bar{\pi}(s)) - Q^*(s, \pi^*(s)) \|_\infty + \gamma \| V^* - V_{\bar{\pi}} \|_\infty$$

[triangle inequality]

$$\leq \| Q^*(s, \pi^*(s)) - \bar{Q}(s, \pi^*(s)) \|_\infty + \| \bar{Q}(s, \bar{\pi}(s)) - Q^*(s, \pi^*(s)) \|_\infty + \gamma \| V^* - V_{\bar{\pi}} \|_\infty$$

$$\|V^* - V_{\bar{\pi}}\|_\infty \leq e + e + \gamma \| V^* - V_{\bar{\pi}} \|_\infty \qquad \left[\text{As } \|\bar{Q} - Q^*\|_\infty \leq e\right]$$

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

<u>Hence Proved</u>