Assignment - 2 (Theory) RL : AI 3000

Donal Loitam

AI21BTECH11009

1(a) Bellman Optimality operator is given by :-

$$L(v) = \max_{a \in A} \left(R^a + \gamma P^a v\right)$$

Let's take

$$\|L(u) - L(v)\|_\infty = \left\|\max_{a \in A}\left(R^a + \gamma P^a u\right) - \max_{a \in A}\left(R^a + \gamma P^a v\right)\right\|_\infty$$

R.T.P $\|L(v) - L(u)\|_\infty \le \gamma \|v - u\|_\infty$ for some $\gamma < 1$

Consider, w.l.o.g

$$\left|\max_a f(a) - \max_a g(a)\right| = \max_a f(a) - \max_a g(a)$$

$$= f(a^*) - \max_a g(a)$$

$$\boxed{\text{where, } a^* = \arg\max_a f(a)}$$

$$\le f(a^*) - g(a^*) \quad \left[\because g(a^*) \le \max_a g(a)\right]$$

$$\le \max_a \left(f(a) - g(a)\right)$$

$$\le \max_a \left|f(a) - g(a)\right|$$

$$\Rightarrow \|L(v) - L(u)\|_\infty \le \max_a \left\|\left(R^a + \gamma P^a v\right) - \left(R^a + \gamma P^a u\right)\right\|_\infty$$

$$\le \max_a \|\gamma P^a (v - u)\|_\infty$$

$$\le \|\gamma P^{a'} (v - u)\|_\infty$$ where, a' is the argmax of R.H.S

P.T.O →

Contd

Ans 1(a)

$$\leq \gamma \|P^{a}\| \, \|(v-u)\|_\infty$$

$$\leq \gamma \|v-u\|_\infty$$ Proved

$$[\because \|Ax\|_\infty \leq \|A\|_\infty \|x\|_\infty]$$

$\therefore$ The Bellman optimality operator is also a contraction map

Proved

---

1 b) The Bellman evaluation operator is :-

$L^\pi(v) = R^\pi + \gamma P^\pi v$, and it follows

$$\boxed{\|L^\pi(u) - L^\pi(v)\|_\infty \leq \gamma \|u-v\|_\infty} \quad \text{— (i)}$$

and $L^\pi(v_k) = v_{k+1}$, $L^\pi(v^\pi) = v^\pi$ — (ii)

Putting $u = v_k$ & $v = v^\pi$ in (i)

$$\|L^\pi(v_k) - L^\pi(v^\pi)\|_\infty \leq \gamma \|v_k - v^\pi\|_\infty$$

$$\Rightarrow \|v_{k+1} - v^\pi\| \leq \gamma \|L^\pi(v_{k-1}) - L(v^\pi)\|_\infty \quad [\text{Use (i)}]$$

$$\leq \gamma^2 \|v_{k-1} - v^\pi\|_\infty$$

$$\vdots$$

$$\leq \gamma^k \|v_1 - v^\pi\|_\infty$$

Proved

Note that, $\boxed{\gamma < 1}$

1(c)

Ans

Given, V^* be the optimal value f^n

Value iteration algorithm terminated after $(K+1)$ iterations

$$\boxed{\|V_{K+1} - V_K\|_\infty < \varepsilon}$$ [ε = tolerance]

$\varepsilon > 0$

Before we proved in 1(a) that

$$\boxed{\|L(u) - L(v)\|_\infty \le \gamma \|u - v\|_\infty} \quad \text{-----} \quad (iii)$$

We know, $V_{n+1} = L(V_n) \quad \forall n$

Put, $V = V_K$ & $U = V_{K+1}$ in eq^n (iii),

$$\|V_{K+2} - V_{K+1}\|_\infty \le \gamma \|V_{K+1} - V_K\|_\infty$$

$$\le \gamma \epsilon \qquad \text{[From given]}$$

Put $U = V_{K+2}$, $V = V_{K+1}$ again in eq^n (iii),

$$\|V_{K+3} - V_{K+2}\|_\infty \le \gamma \|V_{K+2} - V_{K+1}\|_\infty$$

$$\le \gamma^2 \varepsilon$$

Similarly, we can get

$$\boxed{\|V_{K+n+1} - V_{K+n}\|_\infty \le \gamma^n \varepsilon} \quad \text{———} \quad (iv)$$

Now, $$\sum_{n=1}^{l} \|V_{K+n+1} - V_{K+n}\|_\infty \le \sum_{n=1}^{l} \gamma^n \varepsilon \quad \text{— [Use (iv)]}$$

$$\Rightarrow \left\|\sum_{n=1}^{l} (V_{K+n+1} - V_{K+n})\right\|_\infty \le \sum_{n=1}^{l} \gamma^n \varepsilon \quad \text{[Norm property]}$$

P.T.O →

Contd.

1 (c)

$$\Rightarrow \|V_{k+l+1} - V_{k+1}\|_\infty \le \sum_{n=1}^{l} \gamma^n \epsilon$$

$$\Rightarrow \lim_{k+l+1\to\infty} \|V_{k+l+1} - V_{k+1}\|_\infty \le \lim_{k+l+1\to\infty} \sum_{n=1}^{l} \gamma^n \epsilon$$

$$\Rightarrow \|V^* - V_{k+1}\|_\infty \le \frac{\epsilon\gamma}{1-\gamma}, \quad |\gamma| < 1$$

(Assume infinite horizon $\gamma \in [0,1)$)

~~So V_{k+1}~~ is

$\therefore$ After $k+1$ iterations, the value iteration algorithm result V_{k+1} is far from the optimal V^* by at most $\left(\frac{\epsilon\gamma}{1-\gamma}\right)$