DATE ______
PAGE ______

CS747 : Weekly Quiz 5

VIBHAV AGGARWAL
190050128

a) In CV, we were calculating V_{t+1} using V_t, so the order in which we calculate $V_{t+1}(s)$ would not matter since V_t is fixed for that iteration.

However in QV, we are ~~using~~ immediately updating the value of V and using these updated values in the calculation of next state.

The ~~~~ operator B corresponding to the pseudocode is given by:-

$$(B(F))(s) \stackrel{\text{def}}{=} \max_{a \in A} \Big(\sum_{s' < s} T(s,a,s')\{R(s,a,s') + \gamma (B(F))(s')\} + \sum_{s' \geq s} T(s,a,s')\{R(s,a,s') + \gamma F(s')\} \Big)$$

$$= (B^*(G_s))(s)$$

where,

$$G_s(s') = \begin{cases} (B(F))(s') & ; s' < s \\ F(s') & ; s' \geq s \end{cases}$$

DATE
PAGE

(b) (1) B is a contraction mapping

We need to show $\|B(F) - B(G)\|_\infty \leq \gamma \|F-G\|_\infty$

Equivalently, we can show, $|(B(F))(s) - (B(G))(s)| \leq \gamma \|F-G\|_\infty \quad \forall s \in \{1 \cdots n\}$

We do this by induction.

Base case:

For $s=1$, $(B(F))(s) = (B^*(F))(s)$ by applying the definition of B.

Since B^* is a contraction mapping, we have

$$|(B(F))(1) - (B(G))(1)| \leq \gamma \|F-G\|_\infty$$

Induction hypothesis:

For some state $s_0 \geq 1$, let $|(B(F))(s) - (B(G))(s)| \leq \gamma \|F-G\|_\infty \quad \forall s \leq s_0$

Induction step :-

We need to show that $|(B(F))(s_0+1) - (B(G))(s_0+1)| \leq \gamma \|F-G\|_\infty$

$$|(B(F))(s_0+1) - (B(G))(s_0+1)| = \Big| \max_{a\in A} \Big(\sum_{s' < s_0+1} T(s_0+1, a, s') \{ R(s_0+1, a, s') + \gamma (B(F))(s') \}$$

$$+ \sum_{s' \geq s_0+1} T(s_0+1, a, s') \{ R(s_0+1, a, s') + \gamma F(s') \} \Big)$$

$$- \max_{a\in A} \big(\text{< same expression for G >} \big) \Big|$$

DATE
PAGE

$$\leq \gamma \max_{a \in A} \left(\left| \sum_{s' < s_0+1} T(s_0+1, a, s') \left((B(F))(s') - (B(G))(s') \right) + \sum_{s' \geq s_0+1} T(s_0+1, a, s') \left(F(s') - G(s') \right) \right| \right)$$

$$\leq \gamma \max_{a \in A} \left(\sum_{s' < s_0+1} T(s_0+1, a, s') \left| (B(F))(s') - (B(G))(s') \right| + \sum_{s' \geq s_0+1} T(s_0+1, a, s') \left| F(s') - G(s') \right| \right)$$

We know that $|(B(F))(s') - (B(G))(s')| \leq \gamma \|F-G\|_\infty$
for $s' < s_0+1$
[Induction hypothesis]

Also, $|F(s') - G(s')| \leq \|F-G\|_\infty$

$$\Rightarrow |(B(F))(s_0+1) - (B(G))(s_0+1)| \leq \gamma \max_{a \in A} \left(\sum_{s' \in S} T(s_0+1, a, s') \|F-G\|_\infty \right)$$

$$= \gamma \|F-G\|_\infty$$

Hence, by strong induction, we get $\|B(F) - B(G)\|_\infty \leq \gamma \|F-G\|_\infty$ as desired.

(2) V^* is the fixed point of B

Let some V' be the fixed point of B.

So, we have, $V' = B(V')$

$$\Rightarrow V'(s) = \max_{a \in A} \left(\sum_{s' < s} T(s, a, s') \{ R(s, a, s') + \gamma (B(V'))(s') \} + \sum_{s' \geq s} T(s, a, s') \{ R(s, a, s') + \gamma V'(s') \} \right)$$

Since $V' = B(V')$,

$$V'(s) = \max_{a \in A} \left(\sum_{s' < s} T(s,a,s') \{ R(s,a,s') + \gamma V'(s') \} + \sum_{s' \geq s} T(s,a,s') \{ R(s,a,s') + \gamma V'(s') \} \right)$$

$$= \max_{a \in A} \sum_{s' \in S} T(s,a,s') \{ R(s,a,s') + \gamma V'(s') \}$$

This is the same expression as V^* and therefore, $V' = V^*$.

Hence, QV converges to V^* as $T \to \infty$.