PS 4 5.1 * !?

$$V'(s) = R(s) + \gamma \max_a \sum_{s'} P_{sa}(s') \underline{V(s')}$$

Show: $\|B(V_1) - B(V_2)\|_\infty \leq \gamma \|V_1 - V_2\|_\infty$

$\|V\|_\infty = \max_s |V(s)|$ = biggest vector value

$$V'(s) = R(s) + \gamma \max_a \underbrace{E_{s' \sim P_{sa}}[V(s')]}_{\text{weighted avg of all entries}}$$

Example

	V_1	V_2
S_1	500	1000
S_2	600	1200

$B(V_1) = \cancel{R}\ R(s)$

$B(V_2) = R($

$\Rightarrow V_1 - V_2 = \begin{vmatrix} -500 \\ -600 \end{vmatrix}$

$\Rightarrow \|V_1 - V_2\|_\infty$ 600

For all s:

$B(V_1) - B(V_2) = \cancel{R(}$ $R(s) + \gamma \max_a E_{s' \sim P_{sa}} V_1(s')$

$- R(s) \cancel{+} - \gamma \max_a E_{s' \sim P_{sa}} V_2(s')$

PSG 5.1 2/2

$$= \gamma \left(\max_a \underset{s' \sim P_{sa}}{E} V_1(s') - \max_a \underset{s' \sim P_{sa}}{E} V_2(s') \right)$$

vs.

$$\underbrace{\qquad\qquad\qquad}_{\leq}$$

$$V_1(s) - V_2(s)$$

$$\max_a \underset{s' \sim P_{sa}}{E} V(s')$$

is ~~largest if~~ $\leq \max_{(\text{any } s)} V(s')$

(upper bound: $P_{sa}(s') = 1$ for highest $V(s')$)

$$\max_s R(s) - R(s) + \gamma \left(\max_a \underset{s' \sim P_{sa}}{E} [V_1(s')] \right) - \gamma \left(\max_a E [V_2(s')] \right)$$

$$= \max_s \left\| \gamma \left(\max_a \underset{s' \sim P_{sa}}{E} [V_1(s')] - \max_a \underset{s' \sim P_{sa}}{E} [V_2(s')] \right) \right\|$$

$$\leq \max_s \gamma \left(\max_a \left| \underset{s' \sim P_{sa}}{E} [V_1(s')] - \underset{s' \sim P_{sa}}{E} [V_2(s')] \right| \right)$$

second term can only be smaller or equal

$$\leq \max_s \gamma \left(\max_a \underset{s' \sim P_{sa}}{E} \left| [V_1(s') - V_2(s')] \right| \right) \leq \max_s \gamma \left(|V_1(s) - V_2(s)| \right)$$

PS 9. 5b)

Assume V_1 and V_2 are two fixed points

$$\Rightarrow \| B(V_1) - B(V_2) \|_\infty \leq \gamma \| V_1 - V_2 \|_\infty$$

$$\Rightarrow \| V_1 - V_2 \|_\infty \leq \gamma \| V_1 - V_2 \|_\infty$$

Left side can never be smaller than right side because $\gamma < 1$

Can only be equal when $\| V_1 - V_2 \|_\infty = 0$

But $\| V_1 - V_2 \|_\infty$ is only 0 when $V_1 = V_2$ = one fixed point