Lec-11, IT567, 24-25

→ BE (Bellman equality) → does not involve max(·) & is defined for any policy π - whether good or bad.

→ BOE (Bellman optimality equation)

↳ defined for the optimal policies - involves a max(·) operation.

BE → set of $|S|$ simultaneous equations in $|S|$ unknowns given π:

$S = \{s_1, s_2, s_3, s_4\}$

$|S| = 4$

$V_\pi(s_1), V_\pi(s_2), V_\pi(s_3)$
$V_\pi(s_4)$

↳ considering the computational complexity, we go for iterative solution methods.

Choose $\underline{V_0}$ arbitrarily → except that the terminal state, if any, must be given a zero value.

$$\underline{V_0} = \begin{bmatrix} V_\pi(s_1) \\ V_\pi(s_2) \\ \vdots \\ V_\pi(s_n) \end{bmatrix}$$

iteration no.

$s \rightarrow s^+$

If the model has terminal states then the set of " is denoted as s^+

$$\| \underline{V}_{i+1} - \underline{V}_i \|_2 < \epsilon \; ; \; \max_n | \underline{V}_{i+1} - \underline{V}_i | < \epsilon$$

$$\|a\|_2 = \sqrt{1^2 + 2^2 + (-1)^2}$$

$$a = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} ; \; \|a\|_p = \left(a_1^p + a_2^p + a_3^p\right)^{1/p}$$

fixed point of a function $f(x)$? → suppose it is x_1 then $x_1 = f(x_1)$

FP:- also called invariant point is a value that does not change under a given transformation. Specifically, for functions, a fixed point is an element that is mapped to itself by the function.

in-line :- suppose $S = \{s_1, s_2, s_3, s_4\}$

$$\underline{V_0} = \begin{bmatrix} -1 \\ 0.5 \\ 0.3 \\ -2 \end{bmatrix}$$

$V_1(s_1)$ = use $V_0(s_1), V_0(s_2), V_0(s_3)$ & $V_0(s_4)$ to evaluate

$V_1(s_2)$ = → "

→ use $V_1(s_1), V_0(s_2), V_0(s_3), V_0(s_4)$