Alt proof of VI.

$V^* = \underline{T} V^*$.

$\forall f \in \mathbb{R}^S$. $(\underline{T} f)(s) = \max_a \left(R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)} [f(s')] \right)$!

Init $f_0 \equiv \vec{0}$. $f_{k+1} \leftarrow \underline{T} f_k$.

$f_1(s) = \max_a R(s,a)$.

$f_2(s) = \max_a \left(R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)} \left[\max_{a'} R(s',a') \right] \right)$.

define $V_H^\pi(s) = \mathbb{E}_\pi \left[\sum_{t=1}^{H} \gamma^{t-1} r_t \,\middle|\, s_1 = s \right]$.

Claim: $f_k(s) = V_H^*(s) \equiv \max_\pi V_H^\pi(s)$.

(π: depend t.)

Want to bound $\|V^* - f_K\|_\infty \le \dots$

$\Leftarrow V^*(s) - \dots \le f_K(s) \le V^*(s) + \dots$

$\le V_H^*(s)$. $(K = H)$.

① $V_H^*(s) \le V^*(s)$.

(π_H^*, $*$)

(dep. on non-neg reward)

$$V_H^*(s) = V_H^{\pi_H^*}(s) \le V^{\pi_H^* \circ \text{anything}}(s)$$

$$\le V^*(s).$$

② $V_H^*(s) \ge V^*(s) - \boxed{\;\;}$

$$\ge V_H^{\pi^*}(s) = \mathbb{E}_{\pi^*}\left[\sum_{t=1}^{H} \gamma^{t-1} r_t \mid s_1 = s\right]$$

$$= \mathbb{E}_{\pi^*}\left[\sum_{t=1}^{\infty} \gamma^{t-1} r_t \mid s_1 = s\right]$$

$$- \mathbb{E}_{\pi^*}\left[\sum_{t=H+1}^{\infty} \gamma^{t-1} r_t \mid s_1 = s\right].$$

$$\ge V^*(s) - \sum_{t=H+1}^{\infty} \gamma^{t-1} R_{max}.$$

$$= V^*(s) - \gamma^H \cdot \frac{R_{max}}{\boxed{1-\gamma}}.$$

$$\|Q^* - f_K\|_\infty \leq \frac{\gamma^K R_{max}}{1-\gamma}.$$

π_{f_K}.