$Q^* = \mathcal{T} Q^*$. $\forall f \in \mathbb{R}^{S \times A}$, $(\mathcal{T} f)(s,a) := R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)}[\max_{a'} f(s',a')]$

VI: $f_k \leftarrow \mathcal{T} f_{k-1}$. $\forall f, f' \in \mathbb{R}^{S \times A}$, $\|\mathcal{T} f - \mathcal{T} f'\|_\infty \le \gamma \|f - f'\|_\infty$

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

Similarly, define $\mathcal{T}^\pi \Rightarrow$ $V^\pi = \mathcal{T}^\pi V^\pi$

$Q^\pi = \mathcal{T}^\pi Q^\pi$.

$f_k \leftarrow \mathcal{T}^\pi f_{k-1}$.

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

$$V^\pi(s) = R(s,\pi) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,\pi)}[V^\pi(s')].$$

$f$

$$=: (\mathcal{T}^\pi V^\pi)(s).$$

| | |
|---|---|
| $V^*(s)$, | $Q^*(s,a)$ |
| $V^\pi(s)$. | $Q^\pi(s,a)$ |

PE : $Q^{\pi_k}$.

$\forall f \in \mathbb{R}^{S \times A}$. $\pi_f(s) = \arg\max_{a \in A} f(s,a)$.

Policy Improv : $\pi_{k+1} \leftarrow \pi_{Q^{\pi_k}}$.

$\forall s$, $V^{\pi_{k+1}}(s) \geq V^{\pi_k}(s)$.

---

$\forall \pi, \pi'. s,$

$A^{\pi_k}(s', \pi_{k+1})$

$$V^{\pi'}(s) - V^{\pi}(s) = \frac{1}{1-\gamma} \mathbb{E}_{s' \sim d^{\pi', s}}[A^{\pi}(s', \pi')].$$

$$A^{\pi}(s,a) = Q^{\pi}(s,a) - V^{\pi}(s).$$

$$= Q^{\pi}(s,a) - Q^{\pi}(s,\pi).$$

$V^{\pi_{k+1}}(s) - V^{\pi_k}(s)$

$\|f\|_\infty = \max_{s,a} |f(s,a)|$.

step 1. abs value.

step 2. max.

Thm 3. $\|Q^* - Q^{\pi_{k+1}}\|_\infty \le \gamma \|Q^* - Q^{\pi_k}\|_\infty$.

Fact 1. $\mathcal{T}^{\pi_{k+1}} Q^{\pi_k} \ge \mathcal{T}^{\pi} Q^{\pi_k} \quad \forall \pi$.

$$(\mathcal{T}^{\boxed{\pi}} Q^{\pi_k})(s,a) = R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)}[Q^{\pi_k}(s', \boxed{\pi})].$$

Fact 2. $\mathcal{T}^{\pi_{k+1}} Q^{\pi_k} \le Q^{\pi_{k+1}}$.

Proof: $\mathcal{T}^{\pi_{k+1}} Q^{\pi_k} \le \mathcal{T}^{\pi_{k+1}} Q^{\pi_{k+1}}$

$$(\mathcal{T}^{\boxed{\pi_{k+1}}} f)(s,a) := R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)}[f(s', \boxed{\pi_{k+1}})].$$

suffice to show: $Q^{\pi_k}(s', \pi_{k+1})$ vs. $Q^{\pi_{k+1}}(s', \pi_{k+1})$.

$$Q^{\pi_k}(s,a) = R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)}[V^{\pi_k}(s')].$$

$V^{\pi_{k+1}}$

$\wedge$

---

$$Q^* - Q^{\pi_{k+1}} = Q^* - \mathcal{T}^{\pi_{k+1}} Q^{\pi_k} + \mathcal{T}^{\pi_{k+1}} Q^{\pi_k} - Q^{\pi_{k+1}}$$

$$\le Q^* - \mathcal{T}^{\pi_{k+1}} Q^{\pi_k}.$$

$Q^* = \mathcal{T} Q^*$

$$\le Q^* - \mathcal{T}^{\pi^*} Q^{\pi_k}$$

$Q^* = Q^{\pi^*}$

$$Q^* = Q^*$$

$$= \mathcal{T}^{\pi^*} Q^* \quad = \mathcal{T}^{\pi^*} Q^* - \mathcal{T}^{\pi^*} Q^{\pi_k}$$

$$\leq \gamma \cdot \| Q^* - Q^{\pi_k} \|_\infty .$$

$$= \mathcal{T} Q^* - \mathcal{T} Q^{\pi_k}$$

$$\forall f : \mathcal{T}^{\pi_f} f = \mathcal{T} f . \qquad \pi_f$$

$$R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)} \Big[ \underbrace{\max_{a'} f(s',a')}_{f(s', \pi_f)} \Big].$$

$$\mathcal{T}^{\pi_f} \equiv \mathcal{T} ,$$

$$\forall f \in \mathbb{R}^{S \times A} .$$

$$\|V^* - V^{\pi_f}\|_\infty \le \frac{2\cdot\|f - Q^*\|_\infty}{1-\gamma}.$$

Prove.

$\forall \pi, \pi'. s,$

$$V^{\pi'}(s) - V^{\pi}(s) = \frac{1}{1-\gamma}\mathbb{E}_{s'\sim d^{\pi',s}}\left[A^{\pi}(s',\pi')\right]$$

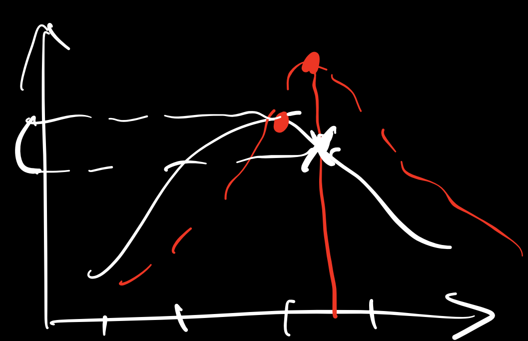
Proof: $V^*(s) - V^{\pi_f}(s) = V^{\pi^*}(s) - V^{\pi_f}(s)$

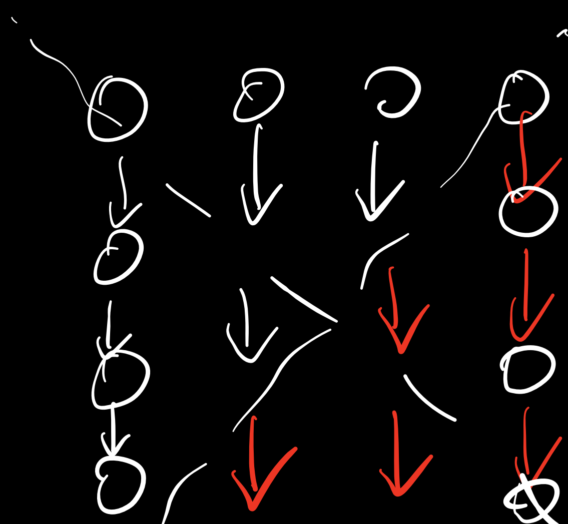
$$= \frac{1}{1-\gamma}\mathbb{E}_{s'\sim d^{\pi_f,s}}\left[Q^*(s',\pi^*) - Q^*(s',\pi_f)\right].$$

$$Q^*(s',\pi^*) - f(s',\pi^*) + f(s',\pi^*) - Q^*(s',\pi_f)$$

$$\le \|f - Q^*\|_\infty + f(s',\pi_f) - Q^*(s',\pi_f).$$

$$\le 2\|f - Q^*\|_\infty.$$





$<\varepsilon$

$<\varepsilon.$ $\quad 2\varepsilon.$

$\pi^*$ $\quad \pi_f.$