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

$$V^*(s) = \max_{a \in A} \underbrace{\left(R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)}[V^*(s')]\right)}_{Q^*(s,a).} \quad \forall s.$$

$$\boxed{Q^*(s,a) = R(s,a) + \gamma \mathbb{E}_{s' \sim P(\cdot|s,a)}\left[\max_{a'} Q^*(s',a')\right] \quad \forall s, a.}$$

Bellman Operator

Optimality.

$$T : \mathbb{R}^{S \times A} \to \mathbb{R}^{S \times A}.$$

$\mathbb{R}^{m \times n} = \mathbb{R}^m$

$\forall f : \mathbb{R}^{S \times A} \quad \forall s, a.$

$$(Tf)(s,a) := R(s,a) + \mathbb{E}_{s' \sim P(\cdot|s,a)}\left[\max_{a'} f(s',a')\right]$$

Bellman opt. eq.: $Q^* = TQ^*$.

Value Iteration

Q^*

init f_0 arbitrarily. ($f_0 \equiv 0$)

for $k = 1, 2, 3, \ldots$ $\quad f_k = T f_{k-1}$.

T is γ-Contraction under $\|\cdot\|_\infty$.

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

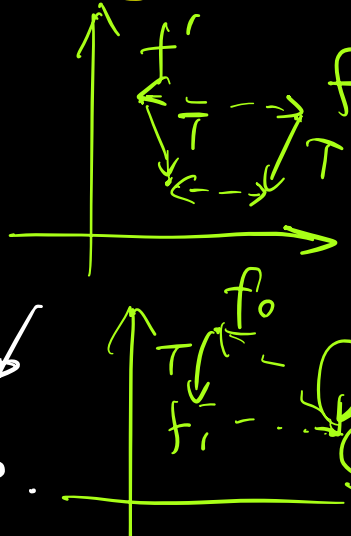
$$\|Tf - Tf'\|_\infty \le \gamma \|f - f'\|_\infty.$$

$$\|f_k - Q^*\|_\infty \le \square \gamma^k$$

$$= \|Tf_{k-1} - TQ^*\|_\infty$$

$$\le \gamma \|f_{k-1} - Q^*\|_\infty$$

$$\le \ldots \le \gamma^k \|f_0 - Q^*\|_\infty$$



$$\|Tf - Tf'\|_\infty \le \gamma \|f - f'\|_\infty$$

$\Leftarrow \forall s, a. \quad |(Tf)(s,a) - (Tf')(s,a)|$

$= |R(s,a) + \gamma \mathbb{E}_{s' \sim p(\cdot|s,a)}[\max_{a'} f(s',a')]$

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

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

$$\leq \gamma \max_{s'} \left| \max_{a'} f(s', a') - \max_{a'} f'(s', a') \right|.$$

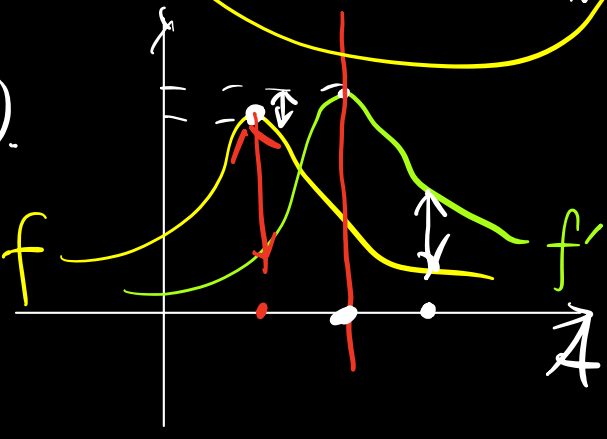
$$\overset{?}{\leq} \gamma \max_{s', a'} \left| f(s', a') - f'(s', a') \right|.$$

Suffices to show: fix s'.

$$\left| \max_{a'} f(s', a') - \max_{a'} f'(s', a') \right| \leq \max_{a'} \left| f(s', a') - f'(s', a') \right|.$$

w.l.o.g. assume.

$$\max_{a'} f(s', a') \geq \max_{a'} f'(s', a').$$



$$a^* = \operatorname*{argmax}_{a'} f(s', a')$$

$$\text{LHS} = f(s', a^*) - \max_{a'} f'(s', a')$$

$$\leq f(s', a^*) - f'(s', a^*).$$

---

1. Sol is unique: $Q_1^* = TQ_1^*$

$Q_2^* = TQ_2^*$.

$$\Rightarrow \|Q_1^* - Q_2^*\|_\infty \le \gamma \|Q_1^* - Q_2^*\|_\infty$$

2. Stopping: ① $V_{max} = \frac{R_{max}}{1-\gamma}$.

$$\|f_k - Q^*\|_\infty \le \gamma^k \|f_0 - Q^*\|_\infty$$

$f_0 \sim 0 \le V_{max}$.

$\le \varepsilon$.

$$\gamma^k V_{max} \le \varepsilon.$$

$$\left(\frac{1}{\gamma}\right)^k \ge \frac{V_{max}}{\varepsilon} \iff k \ge \log_{\frac{1}{\gamma}} \frac{V_{max}}{\varepsilon}$$

$$\le O\left(\frac{1}{1-\gamma}\right).$$

$\|f_{k+1} - f_k\|_\infty$.

$= \|Tf_k - f_k\|_\infty$. $\le \varepsilon'$.

$\|Tf_k - f_k\|_\infty$ → Bellman error.

$\|f_k - Q^*\|_\infty = \|f_k - Tf_k + Tf_k - Q^*\|_\infty$

ε.

$\le \|f_k - Tf_k\|_\infty + \|Tf_k - TQ^*\|_\infty$.

$$\leq \varepsilon' + \gamma \|f_k - Q^*\|_\infty$$

$$\|f_k - Q^*\| \leq \frac{\varepsilon'}{1-\gamma}.$$

---

3. ① $Q^* = TQ^*$ .. ② $f_k \leftarrow T f_{k-1}$ ③ γ-contraction.

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

① ↑

$\forall f \in \mathbb{R}^S$, $(Tf)(s) =$ (f.)

$V^* = TV^*$ ② $f_k \leftarrow T f_{k-1}$

$$V^\pi = R^\pi + \gamma P^\pi V^\pi.$$

$\forall f \in \mathbb{R}^S$. $T^\pi = R^\pi + \gamma P^\pi f$.

$$Q^\pi(s,a) = \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^{t-1} r_t \,\middle|\, s_1 = s,\ a_1 = a,\ a_{2:\infty} \sim \pi \right]$$

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

$\max_a Q^*(s,a) = V^*(s)$ | $Q^\pi(s,\pi) = V^\pi(s)$

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

$$V^{\pi^*} = \mathcal{T}^{\pi^*} V^{\pi^*}.$$

$$\mathbb{E}_{a \sim \pi(\cdot|s)} [Q^\pi(s,a)]$$

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

where $\pi_f(s) = \underset{a \in A}{\operatorname{argmax}} f(s,a)$

$$\|f_K - Q^*\|_\infty \leq \varepsilon.$$

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

$$\text{subopt}(\pi_f) \leq \text{func}(\|f - Q^*\|_\infty)$$

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