An: Step (triggered) that H = \omega(t)

What if no absorbing (terminating) states in discounted

Convergence:

Finite horizon:

Convergence Speed Value Iteration:

\[ \| f - q \|_\infty \leq 1 \]

Convergence

Effective horizon:

\[ e^{-\frac{t-1}{c \log \frac{3}{\epsilon}}} = O\left(\frac{1}{\epsilon}\right) \]

Why \( \epsilon \approx \frac{1}{\log \frac{3}{\epsilon}} \)
from each \((s,a)\), we have sampled \(\{r_i, s'_i\}\) for \(i = 1, \ldots, n\).

\[
\hat{R}(s,a) = \frac{1}{n} \sum_{i=1}^{n} r_i
\]

\[
\hat{P}(s'|s,a) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(s'_i = s')
\]

\[
(\mathcal{J}f)(s,a) = \hat{R}(s,a) + \gamma \mathbb{E}_{s' \sim \hat{P}(.|s,a)} \left[ \max_{a'} f(s',a') \right].
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} r_i + \gamma \sum_{s' \in S} \hat{P}(s'|s,a) \left( \max_{a'} f(s',a') \right).
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} r_i + \frac{1}{n} \sum_{s' \in S} \mathbb{I}(s'_i = s') \max_{a'} f(s'_i,a')
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} r_i + \frac{1}{n} \sum_{i=1}^{n} \max_{a'} f(s'_i,a')
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} r_i + \frac{1}{n} \sum_{i=1}^{n} \max_{a'} f(s'_i,a') \approx (\mathcal{J}f)(s,a)
\]

\[
\mathbb{E} \left[ r_i + \gamma \max_{a'} f(s'_i,a') \right] = \hat{R}(s,a) + \gamma \mathbb{E}_{s' \sim \hat{P}(.|s,a)} \left[ \max_{a'} f(s',a') \right]
\]

\[
= (\mathcal{J}f)(s,a)
\]

"empirical Bellman update"