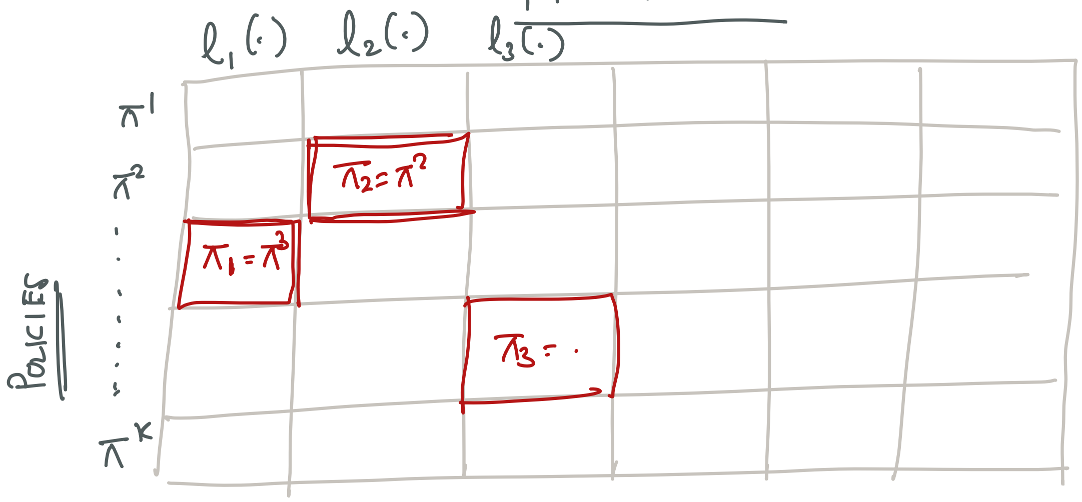
DAGGER PROOF



$$l_i(\pi) = \mathop{\mathbb{E}}_{s \sim d_{\pi_i}} \mathbb{1}\left(\pi(s) \neq \pi^*(s)\right)$$

GOAL: FIND A POLICY $\pi \in \{\pi^1, \dots \pi^k\}$ S.T.

$$\boxed{\mathop{\mathbb{E}}_{s \sim d_{\pi}} \mathbb{1}\left(\pi(s) \neq \pi^*(s)\right)} \leq \varepsilon$$

IS THERE A ROUND i WHERE $\boxed{l_i(\pi_i) \leq \varepsilon}$?

$$\mathop{\mathbb{E}}_{s \sim d_{\pi_i}} \mathbb{1}\left(\pi_i(s) \neq \pi^*(s)\right) \leq \varepsilon$$

$$\min_{i=1,\dots N} l_i(\pi_i) \leq \frac{1}{N}\sum_{i=1}^{N} l_i(\pi_i)$$

$$\leq \left[\frac{1}{N}\sum_{i=1}^{N} l_i(\pi_i) - \min_{\pi \in \{\pi^1, \dots \pi^K\}} \frac{1}{N}\sum_{i=1}^{N} l_i(\pi)\right]$$

$$+ \min_{\pi \in \{\pi^1, \ldots, \pi^k\}} \frac{1}{N} \sum_{i=1}^{N} l_i(\pi)$$

$\leq$ ~~Avg Regret~~ ~~$\rightarrow 0$ as $N \rightarrow \infty$~~ $+ \min_{\pi \in \{\pi^1, \ldots, \pi^k\}} \frac{1}{N} \sum_{i=1}^{N} l_i(\pi) \leq \varepsilon$

Assume: There is a policy in my policy class that can drive down error on aggregated data to epsilon