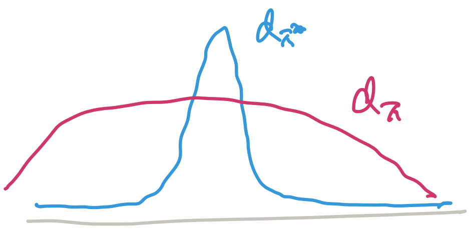
BC in Medium Setting $\quad O(\varepsilon T)$ or $O(\varepsilon T^2)$

$$J(\pi) - J(\pi^*) = \sum_{t=1}^{T} \mathbb{E}_{s\sim d_\pi^t} \underline{A^{\pi^*}(s_t, \pi(s_t))}$$

$\left\| \max_a A^{\pi^*}(s,a) \right\|_\infty$

$$\lesssim M \sum_{t=1}^{T} \mathbb{E}_{s\sim d_\pi^t} \mathbb{1}\left(\pi(s_t) \neq \pi^*(s_t)\right)$$



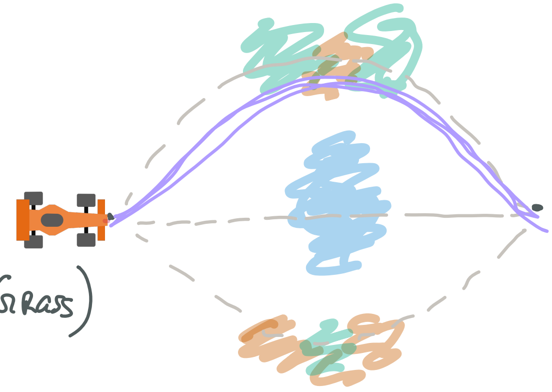
$$\lesssim M \sum_{t=1}^{T} \frac{d_\pi^t(s)}{d_{\pi^*}^t(s)} \underbrace{\boxed{\sum_s d_{\pi^*}^t(s)\, \mathbb{1}\left(\pi(s_t) \neq \pi^*(s_t)\right)}}_{\varepsilon}$$

$$\lesssim M \left\| \frac{d_\pi^t(s)}{d_{\pi^*}^t(s)} \right\|_\infty \varepsilon T$$

$$\lesssim \underline{(\varepsilon T M)} \cdot C$$

$$\boxed{\min\left(\varepsilon T^2 \cdot M,\ \varepsilon T \cdot M C\right)}$$

MDP: $\langle \underbrace{S, A, T}_{\text{known}}, C \rangle$ $\to$ unknown



$C_\theta(s, a, \Phi)$

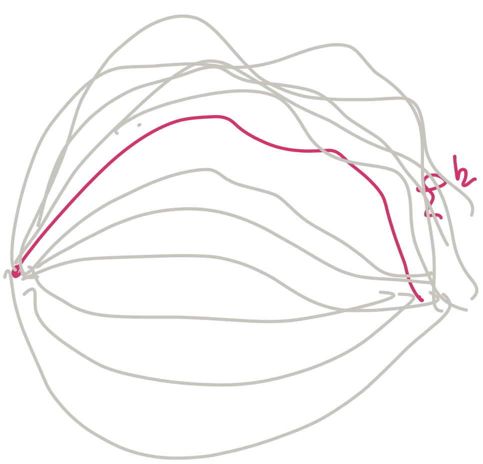
$$= w_1 \underbrace{\mathbb{1}(s \in \text{Water})}_{f^1(s)} + w_2 \underbrace{\mathbb{1}(s \in \text{Grass})}_{f^2(s)} + w_3 \underbrace{\mathbb{1}(s \in \text{Rocks})}_{f^3(s)}$$

$$= w_1 f^1(s) + w_2 f^2(s) + w_3 f^3(s)$$

$$= \underline{w}^T f(s)$$

$w, \Phi \to$ **Planner**: $\min_{\xi} C_\theta(\xi, \Phi) = \sum_{s,a \in \xi} C_\theta(s, a, \Phi) = \sum_{s,a \in \xi} w^T f(s)$ $\to \xi \leftrightarrow \xi^h$

Φ.



$$C_\theta(\xi^h, \Phi) \leq C_\theta(\xi, \Phi) \quad \forall \xi$$

$$w^T f(\xi^h) \leq w^T f(\xi) \quad \forall \xi$$

Margin: $\gamma(\xi, \xi^h) = \begin{cases} 0 & \text{if } \xi = \xi^h \\ 1 & \end{cases}$

$$\underset{1e6}{\underline{w}}^T f(\xi^h) \le \underset{1e6}{\underline{w}}^T f(\xi) - \underset{\in\{0,1\}}{\underline{\gamma}}(\xi, \xi^h) \quad \forall \xi$$

$$\le \min_{\xi} \left[w^T f(\xi) - \gamma(\xi, \xi^h) \right]$$

MAXIMIZE MARGIN (- - - -)

$$\min_{w} \|w\|^2 + \sum_{i=1}^{N} \eta_i$$

s.t. $$w^T f_i(\xi_i^h) \le \min_{\xi} \left[w^T f_i(\xi) - \gamma(\xi, \xi^h) \right] + \eta_i$$

$$\min_{w} \lambda \|w\|^2 + \sum_{i=1}^{N} \left[w^T f_i(\xi_i^h) - \min_{\xi} \left[w^T f_i(\xi) - \gamma(\xi, \xi^h) \right] \right]$$