$\forall\, 0 \leq \zeta \leq 1:$

$$0 \leq X_\zeta \leq \boxed{\frac{1}{1-\gamma}} \qquad \text{wp} \quad 1-\zeta \quad , \quad \underbrace{X_\zeta \leq \frac{\varepsilon}{2}}$$

$$G_\zeta = \left\{ X_\zeta \leq \frac{\varepsilon}{2} \right\}$$

$\exists \zeta > 0:$ $\mathbb{E}[X_\zeta] \leq \varepsilon$

$\forall \zeta \in [0,1]:$

$$\mathbb{E}[X_\zeta] = \mathbb{E}[X_\zeta \mathbb{I}_{G_\zeta}] + \mathbb{E}[X_\zeta \mathbb{I}_{G_\zeta^c}] \qquad (\mathbb{P}(G_\zeta^c) \leq \zeta$$

$$X_\zeta \mathbb{I}_{G_\zeta} \leq \frac{\varepsilon}{2} \qquad X_\zeta \mathbb{I}_{G_\zeta^c} \leq \frac{1}{1-\gamma} \mathbb{I}_{G_\zeta^c}$$

$$\Rightarrow \mathbb{E}[X_\zeta] \leq \mathbb{E}\left[\frac{\varepsilon}{2}\right] + \mathbb{E}\left[\frac{1}{1-\gamma} \mathbb{I}_{G_\zeta^c}\right]$$

$$= \frac{\varepsilon}{2} + \frac{1}{1-\gamma} \underbrace{\mathbb{P}(G_\zeta^c)}_{\leq \zeta} \leq \frac{\varepsilon}{2} + \frac{1}{1-\gamma} \cdot \zeta$$

$$\zeta = (1-\gamma)\frac{\varepsilon}{2} \qquad \Rightarrow \mathbb{E}[X_\zeta] \leq \varepsilon.$$

[Feb 23]

$|A| = O(1)$

Fixed-horizon $H > 0$

local planning

v^* realizability

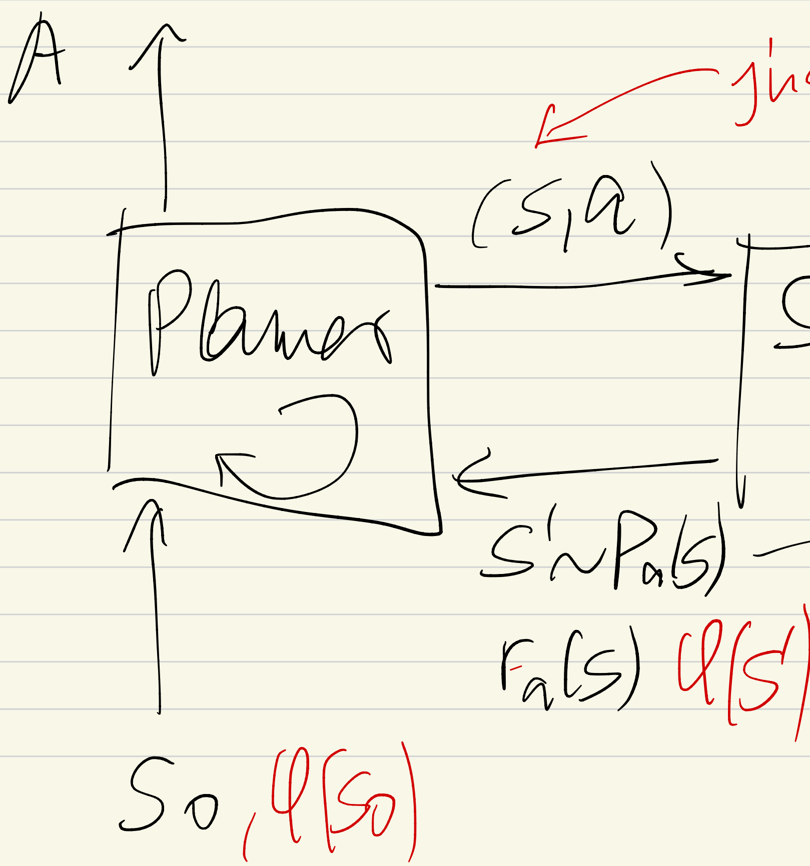
Gellért Weisz!

Thm: $\exists$ planner $\mathcal{P}$ s.t. δ-suboptimal policy w. local planning $\# q = \mathrm{poly}\left[\left(\frac{dH}{\delta}\right)^A, B\right]$

compute cost $= ?$

$$\boxed{v_h^*(s) = \varphi_h(s)^\top \theta_*} \quad h \in [H]$$

$$\|\theta_*\|_2 \le B$$



just states it encountered

No global info about feature map!

Idea of Alg: $\Theta \subseteq \mathbb{R}^d$

$[\Theta = \{ \theta \in \mathbb{R}^d \mid \|\theta\|_2 \le B \}]$

Phases: $s_0 \in S$ $\qquad \theta_* \in \Theta$

$$\boxed{\hat\theta = \operatorname*{argmax}_{\theta \in \boxed{\Theta}} \varphi_1(s_0)^T \theta :}$$

$v_1^*(s) = \varphi_1(s)^T \theta_*$

Local consistency

$\forall s \ \exists a : v_h^*(s) = r_a(s) + \langle P_a(s), v_{h+1}^* \rangle$

B.O.E. $\quad v_h^*(s) = \max_a r_a(s) + \langle P_a(s), v_{h+1}^* \rangle$

$v_{H+1}^* = 0$

$h = H$

$$\varphi_h(s)^T \hat\theta \overset{?}{=} \max_a r_a(s) + \underbrace{\sum_{s'} P_a(s)\, \varphi_{h+1}(s')^T}_{P_a(s)\varphi_{h+1} \in \mathbb{R}^d} \hat\theta$$

$$= \max_a r_a(s) + \langle P_a(s)\varphi_{h+1}, \hat\theta \rangle$$

$$\boxed{\Delta(\theta, h, s, a) = r_a(s) + \langle P_a(s)\varphi_{h+1}, \theta \rangle - \varphi_h(s)^T \theta}$$

Loc. cons: $\forall h, s \ \exists a$ s.t. $\Delta(\theta^*, h, s, a) = 0$.

$\Updownarrow$

$$\boxed{\forall h, s \qquad \underbrace{\prod_{a} \Delta(\theta^A, h, s, a)}_{} = 0}$$

$$\prod_{\underline{a \in A}} \Delta(\theta, h, s, a) = 0 \qquad \text{linear in } f(\theta)$$

$$"\theta^2" = \theta\theta^T = (\theta_i \theta_j)_{i,j}$$

$$"\theta^3" = \theta \otimes \theta \otimes \theta = (\theta_i \theta_j \theta_k)_{i,j,k}$$

$$\Delta(\theta, h, s, a) = \langle \underbrace{\underline{r_a}(s)\,(P_a(s)\psi_{h+1} - \psi_h(s))}_{v_a}, \; \underbrace{1 \; \theta}_{\begin{pmatrix} 1 \\ \theta \end{pmatrix}} \rangle$$

$$\prod_a \langle v_a, u_a \rangle =$$

$$= \langle \otimes v_a, \; \otimes u_a \rangle \qquad f(s, h)$$

$$\prod_a \Delta(\theta, h, s, a) = \langle \overline{\bigotimes_a r_a(s)\,[P_a(s)\psi_{h+1} - \psi_h(s)]}, \; \underbrace{(1\,\theta)^{\otimes A}}_{g(\theta)} \rangle$$

$$= f(s, h)^T g(\theta)$$

$$v^{\otimes A} = \bigotimes_{a=1}^{A} v$$

Loc. cons:

$$\Theta = \left\langle \bigotimes_{a} r_a(s)\; P_a(s)\varphi_{n+1} - \varphi_n(s) \,\middle|\, \begin{pmatrix}1\\ \Theta\end{pmatrix}^{\otimes A} \right\rangle$$

$\in \mathbb{R}^{(d+1)^A}$

$\hat{\Theta}$ given:

$s := s_0$

$$\langle \varphi_n(s), \hat{\Theta}\rangle = r_a(s) + \langle P_a(s)\varphi_{n+1}, \hat{\Theta}\rangle$$

$\exists a \longrightarrow$ take it.

$s := s' \sim P_a(s)$.

FAIL

$\varphi_1(s_0)^T \hat{\Theta}$ verify this!

FAIL

Dealing with noise?

→ Flatter dim.