Feb 25

Finishing off TensorPlan

$\hat{\theta} \in \Theta$

$S_1 = s_0$ ← state where we need action

$(S_t, A_t)_{t=1}^{H}$; A_t cons. with $\hat{\theta}$ ← action

$S_{t+1} \sim P_{A_t}(S_t)$

$$\underbrace{\varphi_1(s_0)^T\hat{\theta}}_{\hat{v}_1(s_0)} - \delta \le \mathbb{E}\left[\sum_{t=1}^{H} r_{A_t}(S_t)\right]$$

$\hat{v}_h(s) = \varphi_h(s)^T\hat{\theta}$

$$\mathbb{E}\left[\sum_{t=1}^{H} r_{A_t}(S_t)\right] = \dots$$

$\Delta(S_t, A_t, t, \hat{\theta}) = -\text{RHS} + \text{LHS}$

$$\boxed{\hat{v}_t(S_t) = r_{A_t}(S_t) + \langle P_{A_t}(S_t), \hat{v}_{t+1} \rangle}$$

$$\longrightarrow \; = \mathbb{E}\left[\underline{r_{A_t}(S_t)} + \underline{\hat{v}_{t+1}(S_{t+1})} \mid S_t, A_t\right]$$

$$\hat{v}_1(s_0) = \hat{v}_1(S_1) =$$

$$= \mathbb{E}\left[\hat{v}_1(S_1)\right] = \mathbb{E}\left[r_{A_1}(S_1) + \underline{\hat{v}_2(S_2)}\right]$$

$$= \mathbb{E}\left[r_{A_1}(S_1) + r_{A_2}(S_2) + \hat{v}_3(S_3)\right] = \dots$$

$$= \dots = \mathbb{E}\left[r_{A_1}(S_1) + \dots + r_{A_H}(S_H) + \hat{v}_{H+1}(S_{H+1}) \right]$$

$$\hat{v}_{H+1}(S_{H+1}) = 0$$

$$\Delta(S_t, A_t, t, \hat{\theta}) = \left(r_{A_t}(S_t) + \langle P_{A_t}(S_t), \hat{v}_{t+1} \rangle \right) - \hat{v}_t(S_t)$$

$$\hat{v}_1(S_0) = \hat{v}_1(S_1) = \mathbb{E}[\hat{v}_1(S_1)] = \mathbb{E}\left[\Delta(S_1, A_1, 1, \hat{\theta}) + r_{A_1}(S_1) + \hat{v}_2(S_2) \right]$$

$$= \dots = \mathbb{E}\left[r_{A_1}(S_1) + \dots + r_{A_H}(S_H) \right] + \mathbb{E}\left[\Delta(S_1, A_1, 1, \hat{\theta}) + \dots + \Delta(S_H, A_H, H, \hat{\theta}) \right]$$

$$\geq \mathbb{E}\left[r_{A_1}(S_1) + \dots + r_{A_H}(S_H) \right] - \delta$$

(*) $$|\Delta(S_i, A_i, i, \hat{\theta})| \leq \frac{\delta}{H} \quad \forall i \in [H]$$

$$\prod_{a \in A} |\Delta(s_i, a, i, \hat{\theta})| \leq \frac{\delta}{H} H^{A-1}$$

$[-H, H]$

Phase $i = 1, 2, \ldots$

Stop: $|\Delta(S_i, A_i, i, \hat{\Theta})| > \frac{\delta}{H}$

$$\hat{\Theta}_{i+1} = \left\{ \Theta \in \hat{\Theta}_i \;\middle|\; \prod_{a \in A} |\Delta(S_i, a, i, \Theta)| \leq \left(\frac{\delta}{H} \cdot H^{A-1}\right)^? \right\}$$

$\left(\frac{\delta}{H} \cdot H^{A-1}\right) = \varepsilon$

How long can this go on?

$\mathcal{X}$: data, e.g. for us S_i, i

Θ : (initial) hypothesis $\left(\{\theta \in \mathbb{R}^d \mid \|\theta\|_2 \leq B\} \right)$

$\Delta : \mathcal{X} \times \Theta \to \mathbb{R}$ $\quad [f(S_i, i)^T g(\theta) = \Delta(S_i, i, \theta)]$

discriminant fn.

$x_{1:n} = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$ s.t.

$\forall i \in [n]$ it holds that $\exists \theta \in \Theta$

$$|\Delta(x_i, \theta)| > \varepsilon \text{ s.t. } \sum_j |\Delta(x_j, \theta)|^2 \leq \varepsilon$$

~~$\forall j \leq i-1$~~.

<u>Def:</u> ε-eluder seq. for Δ / Θ

$\varepsilon > 0$

Def: $\dim_E^{\Delta}(\Theta, \varepsilon) =$

$= \max\{n \mid \exists x_{1:n} \in \mathcal{X}^n$ s.t. $x_{1:n}$ ε-eluder seq. for $\Delta / \Theta\}$

$\mathcal{X} \subseteq B_2(\gamma)$, $\Theta = \underbrace{\{\theta \in \mathbb{R}^d \mid \|\theta\|_2 \leq S\}}_{B_2(S)}$

$S, \gamma > 0$

$\Delta(x, \theta) = x^T\theta$ linear discriminant

$\dim_E^{\Delta}(\Theta, \varepsilon)$

Prop: $\dim_E^{\Delta}(\Theta, \varepsilon) = O\left(d \ln\left(\frac{\gamma S}{\varepsilon}\right)\right)$

Proof: Let $x_{1:n} \in \mathcal{X}^n$ ε-eluder sequence.

Fix $i \in [n]$. $\exists \theta_i \in \Theta$

s.t.

$\varepsilon < |\Delta(x_i, \theta_i)| = |x_i^T \theta_i|$

~~$\forall 1 \le j \le i-1$:~~ $\sum_j |\Delta(x_j, \theta_i)|^2 \le \varepsilon^2$

$$\varepsilon < |x_i^T \theta_i| \le \max\{|x_i^T \theta| \mid \sum_j \Delta(x_j, \theta)^2 \le \varepsilon^2, \; \|\theta\|_2^2 \le S^2\}$$

$$\le \max\{|x_i^T \theta| \mid \sum_{j=1}^{i-1} \Delta(x_j, \theta)^2 \le \varepsilon^2, \; \|\theta\|_2^2 \le S^2\}$$

$$= \max\{|x_i^T \theta| \mid \theta^T \underbrace{\sum_{j=1}^{i-1} x_j x_j^T}_{X_{i-1}} \theta \le \varepsilon^2, \; \|\theta\|_2^2 \le S^2\}$$

$\theta^T \theta \le S^2$

$$\le \max\{|x_i^T \theta| \mid \theta^T (X_{i-1} + \lambda I)\theta \le 2\varepsilon^2\}$$

$$\theta^T (X_{i-1} + \lambda I)\theta \le 2\varepsilon^2$$

$$\lambda \|\theta\|_2^2 \le \varepsilon^2$$

$$\boxed{\lambda = \frac{\varepsilon^2}{S^2}}$$

$$V_{i-1} = X_{i-1} + \lambda I$$

$$= \max \left\{ |x_i^T \theta| \;\middle|\; \theta^T V_{i-1} \theta \leq 2\varepsilon^2 \right\}$$

$$v = \frac{V_{i-1}^{1/2}}{\sqrt{2} \cdot \varepsilon} \theta \qquad \|\theta\|_2^2 \leq 1$$

$$\|v\|_2^2 \leq 1 \qquad \max \left\{ x^T \theta \;\middle|\; \|\theta\|_2 \leq 1 \right\}$$

$$\theta = \frac{x}{\|x\|_2}$$

$$\theta = \sqrt{2}\varepsilon V_{i-1}^{-1/2} v$$

$$= \frac{x^T x}{\|x\|_2} = \|x\|_2 .$$

$$= \max \left\{ \sqrt{2}\varepsilon |x_i^T V_{i-1}^{-1/2} v| \;\middle|\; \|v\|_2 \leq 1 \right\}$$

$$(V_{i-1}^{-1/2} x_i)^T v$$

$$= \sqrt{2} \cdot \varepsilon \left(\|V_{i-1}^{-1/2} x_i\|_2^2 \right)^{1/2} = \sqrt{2} \cdot \varepsilon \, \|x_i\|_{V_{i-1}^{-1}}$$

$$x_i^T V_{i-1}^{-1/2} V_{i-1}^{-1/2} x_i$$

$$\varepsilon < \dots \leq \sqrt{2}\,\varepsilon \, \|x_i\|_{V_{i-1}^{-1}}$$

$$\|x_i\|_{V_{i-1}^{-1}}^2 \geq \frac{1}{2}$$

$$V_i = X_i + \lambda I$$

$$\ldots \le \det(V_i) \le \ldots$$

$$\det(V_i) \le \left(\frac{\operatorname{tr} V_i}{d}\right)^d \qquad \text{AM-GM} \quad \text{Jensen}$$

$$\le \left(\frac{\lambda d + i \cdot \gamma^2}{d}\right)^d$$

$$\det V_i = \det(V_{i-1})\left(1 + \|x_i\|^2_{V_{i-1}^{-1}}\right)$$

$$\ge \det(V_{i-1}) \frac{3}{2} \ge \ldots \ge \det(\lambda I)\left(\frac{3}{2}\right)^i$$

V_0, λ^d

$$\lambda^d \left(\frac{3}{2}\right)^i \le \det(V_i) \le \left(\frac{\lambda d + i\gamma^2}{d}\right)^d$$

/ log

$$d\log\lambda + i\log\left(\frac{3}{2}\right) \le \ldots d\log\left(1 + \frac{i\gamma^2}{d}\right)$$

$\lambda = \frac{\varepsilon^2}{S^2}$

$\max\{i \mid \text{s.t. } \uparrow \text{ holds}\}$

$$O\left(d \log \frac{S^2\gamma^2}{\varepsilon^2}\right)$$