Lecture 12
TensorPlan and eluder sequences (TensorPlan II.)

can only measure approximately !

$$\Delta(s,a,h,\theta) = \langle \overline{r_a(s)\,(P_a(s)\phi_{h+1} - \phi_h(s))}, \overline{1\,\theta} \rangle$$

Now, recall that the tensor product $\otimes$ of vectors satisfies the following property:

$$\prod_a \langle x_a, y_a \rangle = \langle \otimes_a x_a, \otimes_a y_a \rangle\,,$$

we see that (5), and thus local consistency, is equivalent to

$$\langle \underbrace{\otimes_a \overline{r_a(s)\,(P_a(s)\phi_{h+1} - \phi_h(s))}}_{D(s,h)}, \underbrace{\otimes_a \overline{1\,\theta}}_{F(\theta)} \rangle = 0\,.$$

$$\left| \langle \overbrace{\otimes_a \frac{1}{n} \sum_{i=1}^{n} \overline{R_{a,i} \left(\phi(s'_{a,i}) - \phi(s) \right)}}^{\tilde{D}(s,h)}, \otimes_a \overline{1\,\theta} \rangle \right| \overset{?}{\leq} \varepsilon$$

n rewards/transitions observed for s, ∀a.

For $a = a^*$,

$$\operatorname*{arg\,max}_{\theta \in B_2(B)} \{ \langle \theta, \varphi(s_0) \rangle |$$

$$\left| \left\langle \overline{\frac{1}{n} \sum_{i=1}^{n} R_{a,i} \left(\varphi(s'_{a,i}) - \varphi(s) \right)}, \overline{\theta} \right\rangle \right| \leq \delta$$

whp (Hoeffding)

for $a \neq a^*$

$$\left| \left\langle \overline{\frac{1}{n} \sum_{i=1}^{n} R_{a,i} \left(\varphi(s'_{a,i}) - \varphi(s) \right)}, \overline{\theta} \right\rangle \right| \leq \delta + H$$

So

$$\tilde{D}(s,h) \leq \delta \cdot (\delta + H)^{A-1} \leq \varepsilon / \sqrt{E_d} \quad \text{for} \quad \delta \approx \varepsilon / H^A \sqrt{E_d}$$

$$n \approx \frac{1}{\delta^2}$$

if $|\langle \tilde{D}(s,h), F(\theta) \rangle| > \varepsilon$

detect inconsistency,

add requirement of $|\langle \tilde{D}(s,h), F(\theta) \rangle| > \varepsilon / \sqrt{E_d}$

How long can a sequence go s.t.

θ_i chosen s.t. $\sum_{j<i} \left| \langle \tilde{D}_j , F(\theta_i) \rangle \right| \leq \varepsilon$

but still, $\left| \langle \tilde{D}_i , F(\theta_i) \rangle \right| > \varepsilon$

?

Eluder Dimension and the Sample Complexity of Optimistic Exploration

Daniel Russo
Stanford University
Stanford, CA 94305
djrusso@stanford.edu

Benjamin Van Roy
Stanford University
Stanford, CA 94305
bvr@stanford.edu

$\ell = 3$

x_1, x_2, x_3

class of functions $\mathcal{X} \to \mathbb{R}$

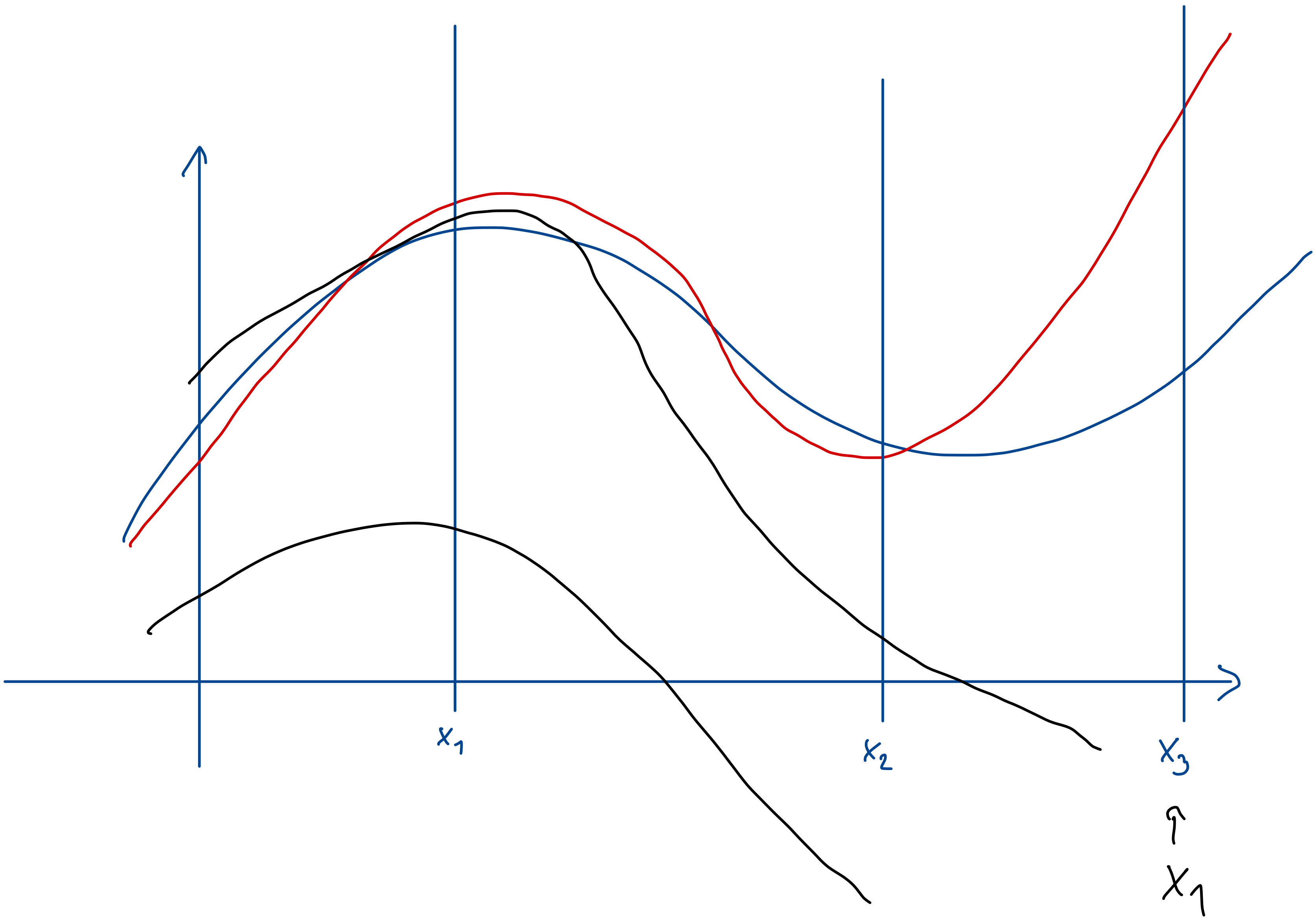
scale

$\dim_E(\mathcal{F}, \varepsilon)$: length τ of the longest "eluder" sequence, of $x_1, \ldots, x_\tau \in \mathcal{X}$ st. for some $\varepsilon' \geq \varepsilon$, for each $\ell = 1, \ldots, \tau$,

$$\sup \left\{ |f_1(x_\ell) - f_2(x_\ell)| \;:\; \sqrt{\sum_{i=1}^{\ell-1} (f_1(x_i) - f_2(x_i))^2} \leq \varepsilon', \; f_1, f_2 \in \mathcal{F} \right\} > \varepsilon'$$

$\dim_E(\mathcal{F}, \varepsilon) = \tilde{O}(d)$ hides $\log(BL/\varepsilon)$

fn class of $x \mapsto \langle x, \theta \rangle$ for $\theta \in \mathbb{R}^d$, $\|\theta\|_2 \leq B$
$x \in \mathbb{R}^d$, $\|x\|_2 \leq L$



x_1
x_2
x_3
X_1

Let $V_i = \lambda I + \sum_{t=1}^{i} x_i x_i^T$, $\quad \lambda = \varepsilon^2 / B^2 \quad$ (s.t. $\lambda \|\theta\|_2^2 \le \varepsilon^2$)

$\exists$ new eluder sequence element x_{i+1} implies:

$$\sqrt{\sum (\langle x, \theta \rangle)^2} < \varepsilon$$

$$\|\theta\|_{V_{i-1}} < \varepsilon$$

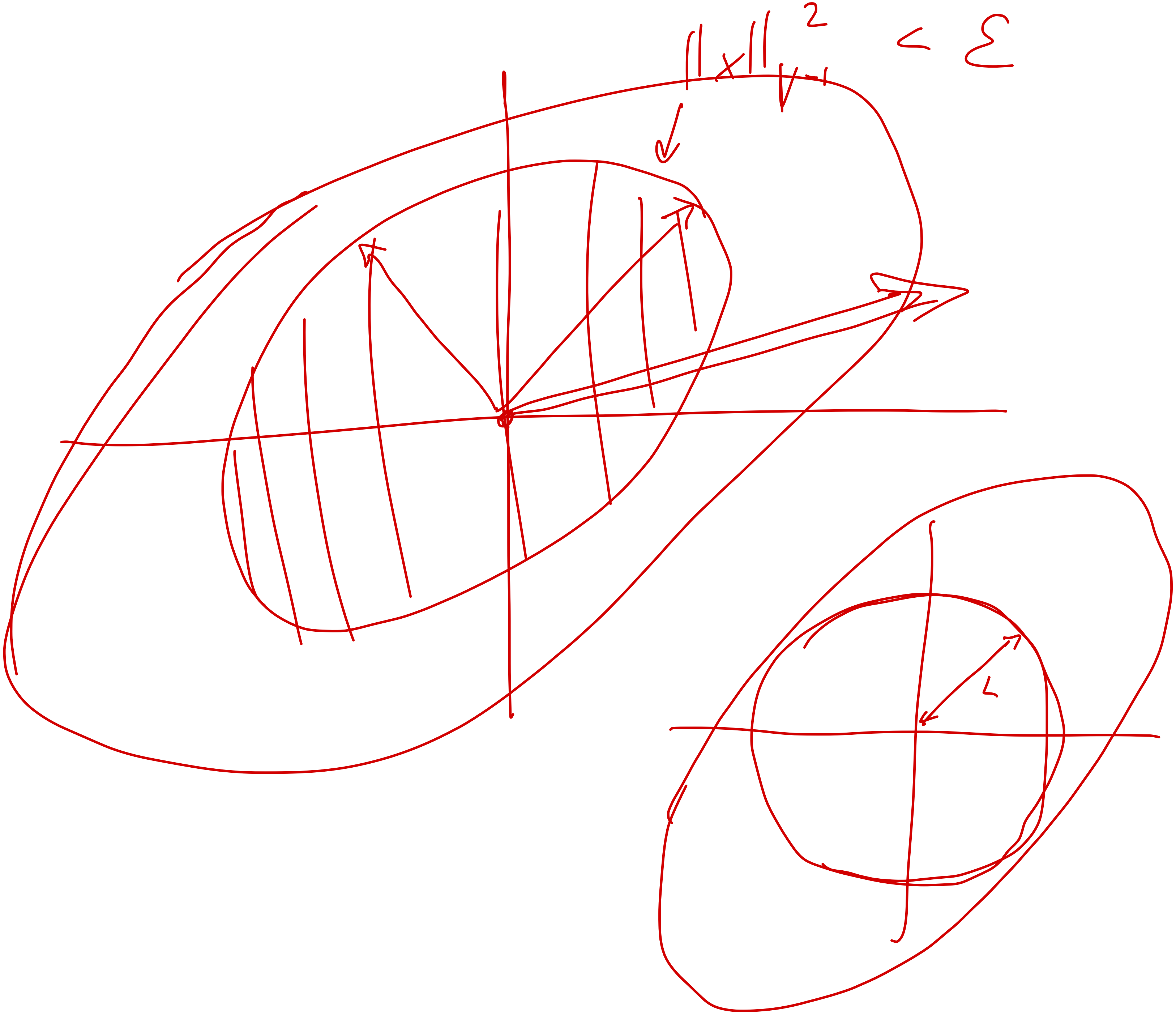
$$\|x_{i+1}\|^2_{V_i^{-1}} \ge \frac{1}{2} \quad \checkmark$$

How long can this go?

$$\det V_i = \det V_{i-1} \cdot \left(1 + \|x_i\|^2_{V_{i-1}^{-1}}\right) \ge \underset{\det V_0}{\lambda^d} + \left(\frac{3}{2}\right)^i$$

AM-GM on product/sum of eigenvals

but also $\quad \det V_i \le \left(\frac{\operatorname{tr} V_i}{d}\right)^d \le \left(\frac{\lambda d + i \cdot L^2}{d}\right)^d$


$\|x\|_{V_1}^2 < \varepsilon$

Connection : elliptical potential lemma (bandits book) :

LEMMA 19.4. *Let* $V_0 \in \mathbb{R}^{d\times d}$ *be positive definite and* $a_1, \dots, a_n \in \mathbb{R}^d$ *be a sequence of vectors with* $\|a_t\|_2 \le L < \infty$ *for all* $t \in [n]$, $V_t = V_0 + \sum_{s\le t} a_s a_s^\top$. *Then,*

$$\sum_{t=1}^{n} \left(1 \wedge \|a_t\|^2_{V_{t-1}^{-1}}\right) \le 2\log\left(\frac{\det V_n}{\det V_0}\right) \le 2d\log\left(\frac{\operatorname{trace} V_0 + nL^2}{d\det(V_0)^{1/d}}\right).$$

$$\left(1 + \frac{nL^2}{d\lambda}\right)$$

Eluder Dimension and Generalized Rank

Gene Li[1] gene@ttic.edu
Pritish Kamath[1] pritish@ttic.edu
Dylan J. Foster[2] dylanf@mit.edu
Nathan Srebro[1] nati@ttic.edu

[1]Toyota Technological Institute at Chicago, [2]MIT

Slightly different Edim:

Definition 1. *For any function class $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$, $f^\star \in \mathcal{F}$, and scale $\varepsilon \geq 0$, the* ***exact eluder dimension*** $\underline{\mathsf{Edim}}_{f^\star}(\mathcal{F}, \varepsilon)$ *is the largest m such that there exists $(x_1, f_1), \ldots, (x_m, f_m) \in \mathcal{X} \times \mathcal{F}$ satisfying:*

$$\forall i \in [m] \quad : \quad |f_i(x_i) - f^\star(x_i)| > \varepsilon, \quad \textit{and} \quad \sum_{j<i} (f_i(x_j) - f^\star(x_j))^2 \leq \varepsilon^2. \tag{1}$$

Then for all $\varepsilon > 0$:

- *the* ***eluder dimension*** *is* $\mathsf{Edim}_{f^\star}(\mathcal{F}, \varepsilon) = \sup_{\varepsilon' \geq \varepsilon} \underline{\mathsf{Edim}}_{f^\star}(\mathcal{F}, \varepsilon')$.
- $\underline{\mathsf{Edim}}(\mathcal{F}, \varepsilon) := \sup_{f^\star \in \mathcal{F}} \underline{\mathsf{Edim}}_{f^\star}(\mathcal{F}, \varepsilon)$ *and* $\mathsf{Edim}(\mathcal{F}, \varepsilon) := \sup_{f^\star \in \mathcal{F}} \mathsf{Edim}_{f^\star}(\mathcal{F}, \varepsilon)$.

This definition is never larger than the original definition of Russo and Van Roy (2013), which asks for a witnessing *pair* of functions $f_i, f_i' \in \mathcal{F}$ (the above restricts $f_i' = f^\star$). Hence, all lower bounds on our

Sdim: like Edim but $\sum_{j<i} \to \sum_{j \neq i}$

Definition 3. *For any $\sigma : \mathbb{R} \to \mathbb{R}$, the* **$\sigma$-rank** $\sigma\text{-}\mathsf{rk}(\mathcal{F}, R)$ *of a function class $\mathcal{F} \subseteq (\mathcal{X} \to \mathbb{R})$ at ~~scale~~ $R > 0$ is the smallest dimension d for which there exists mappings $\phi : \mathcal{X} \to \mathcal{B}_d(1)$ and $w : \mathcal{F} \to \mathcal{B}_d(R)$ such that*[3]

$$\text{for all } (x, f) \in \mathcal{X} \times \mathcal{F} \quad : \quad f(x) \;=\; \sigma(\langle w(f), \phi(x) \rangle), \tag{2}$$

B

or ∞ if no such d exists. For a collection of activation functions $\Sigma \subseteq (\mathbb{R} \to \mathbb{R})$, the **$\Sigma$-rank** *is*

$$\Sigma\text{-}\mathsf{rk}(\mathcal{F}, R) \;:=\; \min_{\sigma \in \Sigma} \sigma\text{-}\mathsf{rk}(\mathcal{F}, R).$$

Examples : rk : identity-rk

$\mathcal{M}_\mu$: σ diff.able, $\mu \leq \sigma' \leq 1$

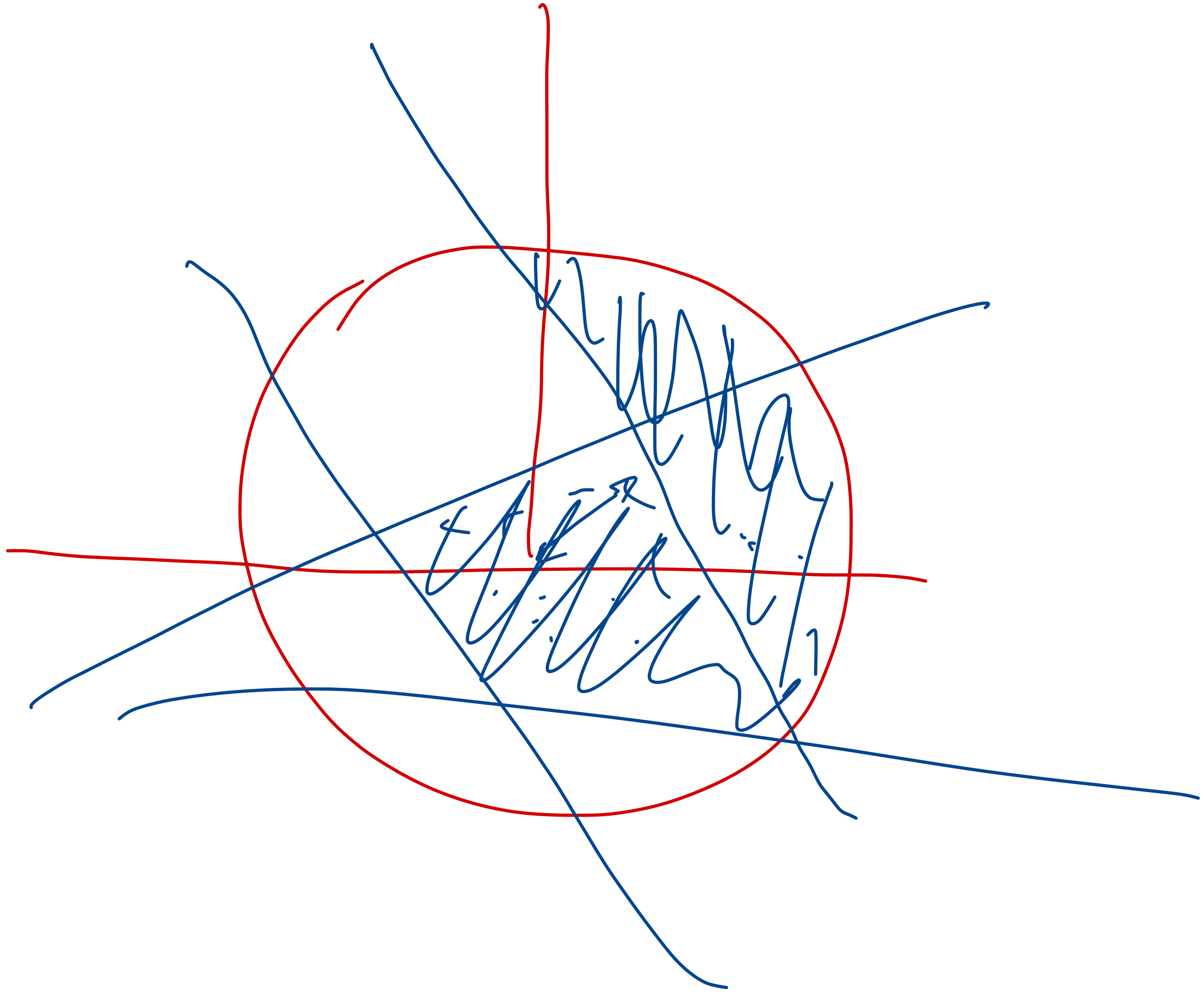
$$\begin{tikzcd}
 & \mathsf{rk} & \\
 & \mathcal{M}_\mu\text{-}\mathsf{rk} \arrow[u] & \\
\mathsf{Edim} \arrow[ur] & & \mathcal{M}_0\text{-}\mathsf{rk} \arrow[ul] \\
\mathsf{Sdim} \arrow[u] & &
\end{tikzcd}$$

(a)

<table>
<tr><th></th><th>rk</th><th>$\mathcal{M}_\mu$-rk</th><th>$\mathcal{M}_0$-rk</th><th>Edim</th><th>Sdim</th></tr>
<tr><td>rk</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>$\mathcal{M}_\mu$-rk</td><td>$\mathcal{F}^{\text{exp}}$</td><td></td><td></td><td></td><td></td></tr>
<tr><td>$\mathcal{M}_0$-rk</td><td colspan="2">$\mathcal{F}^{\text{relu}}$</td><td></td><td colspan="2">$\mathcal{F}^{\text{relu}}$</td></tr>
<tr><td>Edim</td><td colspan="3" rowspan="2">$\mathcal{F}^{\oplus}$</td><td></td><td></td></tr>
<tr><td>Sdim</td><td>$\mathcal{F}^{\text{th}}$</td><td></td></tr>
</table>

(b)

Figure 1: (a) Each arrow $M_1 \to M_2$ indicates that $M_1(\mathcal{F}) \lesssim M_2(\mathcal{F})$ for all $\mathcal{F}$, where the dependence on R and ε is suppressed for clarity (see Propositions 1, 4, 5 for precise bounds). Whenever $M_2 \to M_1$ arrow is missing, there is an example of a class $\mathcal{F}$ where $M_1(\mathcal{F}) \ll M_2(\mathcal{F})$. (b) An entry $\mathcal{F}$ in (M_1, M_2) means that $M_1(\mathcal{F}) \ll M_2(\mathcal{F})$. Green cells indicate that $M_1(\mathcal{F}) \gtrsim M_2(\mathcal{F})$ for all $\mathcal{F}$.



Homayoon Farrahi 5 hours ago

For q* or v* realizable features, there exists a parameter vector that can produce q* or v* using dot product with the features. Is it possible to come up with q* or v* realizable features for an MDP without having to solve for and find the optimal (action-)value function first?

+ 4

$$(0, \dots, \underbrace{\qquad\qquad})$$

$$\theta = (1, \dots, \dots)$$

Gábor Mihucz 5 hours ago

According to this paper (from lecture note 8):
Wang, Ruosong, Dean P. Foster, and Sham M. Kakade. 2020. "What Are the Statistical Limits of Offline RL with Linear Function Approximation?"

"for off-policy evaluation, even if we are given a representation that can perfectly represent the value function of the given policy and the data distribution has good coverage over the features, any provable algorithm still requires an exponential number of samples to non-trivially approximate the value of the given policy."

And the summary of TensorPlan is that it would need queries exponential in the number of actions (but polynomial in dH^A).

Is there any hope that we can avoid the exponential realm with linear function approximation? If not really, in what way could one rigorously reason about "stronger" (non-linear) representations? (edited)

+ 4

$\to$ v^π q^π is linear $\quad \varepsilon \to \varepsilon \cdot \sqrt{d}$

q^*

Yilin Wang 2 hours ago

When we're actually implementing the offline planner, is it worthwhile to spend more efforts in the "preparation step" working out a better feature map, such as choosing an optimal setting of the feature vector dimension d or bounding it? (cuz the size of d has great influence on the computational cost or query cost) If so, are there any nice developed methodologies or research work regarding this?

+ 3

$\varphi = (\cdot \ \cdot \ - \ - \ - \ -)$

$\theta = (0\ 0\ ?\ 0\ 0\ ?\ 0\ 0\ 0\ 0)$

s-sparse θ