Low-rank MDP:

$$\underset{S\times A}{\begin{bmatrix} & & \\ & P & \\ & & \end{bmatrix}} = \underset{S\times A}{\begin{bmatrix} \Phi \\ \text{---}\ \phi(s,a)\ \text{---} \end{bmatrix}}_{d} \; \underset{d}{\begin{bmatrix} \Psi \end{bmatrix}}_{S}$$

$$R(s,a) = \underline{\phi(s,a)^T}\,\underline{\theta_R}.$$

Linear MDP = Low-rank MDP + Φ is known.

$\forall f: S\times A \to \mathbb{R}$. $\forall \pi$.

$\mathcal{T}^\pi f$, $\mathcal{T} f$ are linear in ϕ.

Martingale Azuma's Ineq.

Let $\{S_k, k=0,1,2,\ldots\}$ be a martingale.

w/ $|S_k - S_{k-1}| \le c_k$ a.s. Then fix N.

w.p. $\ge 1-\delta$, $|S_N - S_0| \le \sqrt{2\left(\sum_{k=1}^N c_k^2\right)\log\frac{2}{\delta}}$.

$$\mathbb{E}[S_{k+1} - S_k \mid \underline{s_1,\ldots,s_k}] = 0.$$

i.i.d $X_1, X_2, \cdots, X_N \in [a,b]$.

$$\left|\underbrace{\sum_{i=1}^{N}\left(X_i - \mathbb{E}[X_i]\right)}_{S_N}\right| \le O\left((b-a)\sqrt{N}\right)$$

$k=3$.

MAB.

a_1	a_2	a_3	a_T
r_1	r_2		r_T

$r_t \sim R(\cdot \mid a_t)$.

$R(\cdot \mid a)$ mean μ_a.

$$\underbrace{\sum_{t=1}^{T}\left(r_t - \mu_{a_t}\right)}_{S_T}$$

$$\mathbb{E}[S_k - S_{k-1} \mid S_0, \cdots, S_{k-1}] \overset{?}{=} 0$$

$$= \mathbb{E}\left[\mathbb{E}[r_k - \mu_{a_k} \mid S_0, \cdots, S_{k-1}, a_k]\right]$$

$a_k \mid S_0, \cdots, S_{k-1}$.

$$= \mathbb{E}_{a_k | s_0, \dots, s_{k-1}} \mathbb{E}[r_k - \mu_{a_k} | a_k] = 0.$$
Δ

MAB. $A = \{a, a'\}$. Alg: toss a coin
$$\begin{cases} p = 1/2, & a_t = a \ \forall t. \\ q = 1/2, & a_t = a' \ \forall t. \end{cases}$$

$$\begin{matrix} a_1 & a_2 & a_3 & \cdots & a_T \\ r_1 & r_2 & r_3 & & r_T. \end{matrix}$$

① $\sum_{t=1}^{T} (r_t - \mu_{a_t})$. $\nearrow \frac{\mu_a + \mu_{a'}}{2}$.

② $\sum_{t=1}^{T} (r_t - \mathbb{E}[r_t])$.

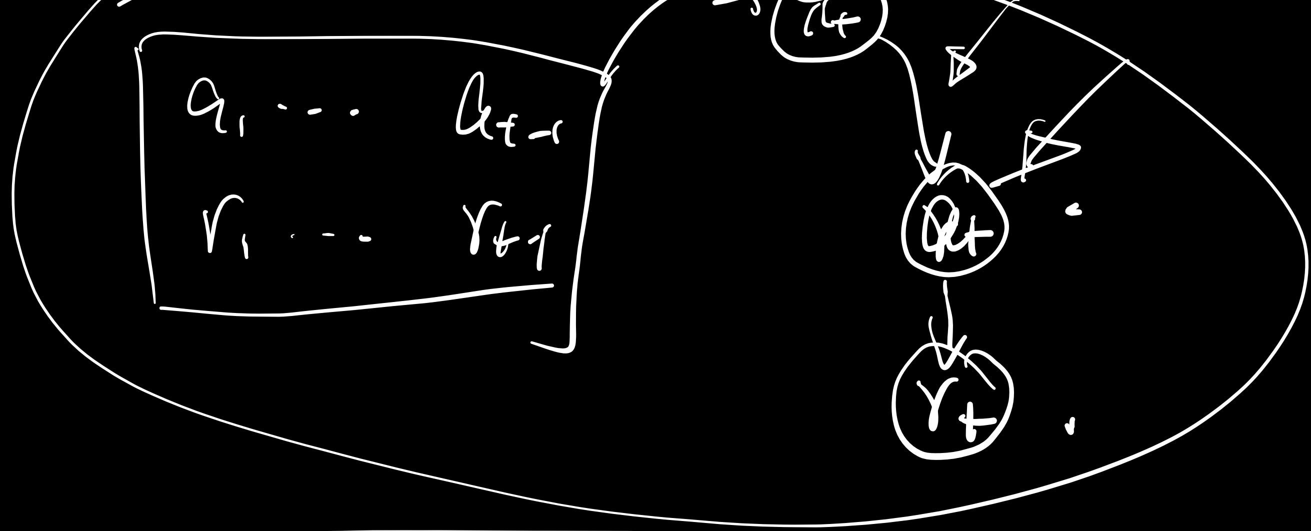
Δ.

$$\left(\sum_{t=1}^{T} (r_t - \underset{\Delta}{\boxed{\mathbb{E}}}) \right)$$

MAB. every t, Alg: $(a_{1:t-1}, r_{1:t-1}) \mapsto \pi_t \in \Delta(A)$.

① $\sum_{t=1}^{T} (r_t - \mu_{a_t})$ ✗ ✓

② $\sum_{t=1}^{T} (r_t - \mathbb{E}_{a \sim \pi_t}[r])$.



Ridge Regression for $t = 1, 2, \cdots, T$.

- Nature chooses $x_t \in \mathbb{R}^d$. $\|x_t\| \le 1$.
- Receive $y_t = \underbrace{x_t^T \theta^* + \varepsilon_t}$. $\in [0, V_{max}]$.

ε_t: 0-mean noise.

Goal 1: At t, $x_{1:t-1}$, $y_{1:t-1}$

predict $y(x) = x^T \theta^*$.

Alg: $\hat{\theta}_t = \arg\min_{\theta \in \mathbb{R}^d} \sum_{i=1}^{t-1} (x_i^T \theta - y_i)^2 + \|\theta\|^2$.

$\hat{y}_t(x) = x^T \hat{\theta}_t$.

$\|x\|_A := \sqrt{x^T A x}$.

Claim: fix t. w.p. $\ge 1-\delta$, $\forall x \in \mathbb{R}^d$,

$\triangleright \; |\hat{y}_t(x) - y(x)| \le \|x\|_{\Lambda^{-1}} \cdot O(V_{max}\sqrt{d + \log\frac{1}{\delta}}$

$$\Lambda_t := I + \sum_{i=1}^{t-1} x_i x_i^T \in \mathbb{R}^{d\times d}$$

$x_1 = x_2 = \cdots = x_t = x^0.$

$y_1 \quad y_1 \quad \cdots \quad y_t.$

$$(a x^0)^T \theta^* \approx \frac{\sum_{i=1}^{t} y_i}{t} \times a.$$

$x_1 = \cdots = x_{t/2} = x^0.$

$x_{t/2+1} = \cdots = x_t = x^1.$

$$(a x^0 + b x^1)^T \theta^*$$

$t = 1, 2, 3, \cdots, T.$

- Nature picks $x_t \in \mathbb{R}^d$. $\|x_t\| \le 1$
- Learner predicts $\hat{y}_t = \hat{y}_t(x_t).$
- Observe y_t. $x^T\theta^*$.

$$\text{Regret}_T := \sum_{t=1}^{T} \left| \hat{y}_t - y(x_t) \right| = \tilde{O}(\sqrt{T})$$ poly(d).

Lemma: (Elliptical potential lemma). $\|x_t\| \le 1$.

$$\sum_{t=1}^{T} x_t^T \Lambda_t^{-1} x_t \le 2\log\det(\Lambda_{T+1}) \le \boxed{2d\log(T+1)}$$

$$\text{Regret}_T = \sum_{t=1}^{T} \left| \hat{y}_t - y(x_t) \right|$$

Const.

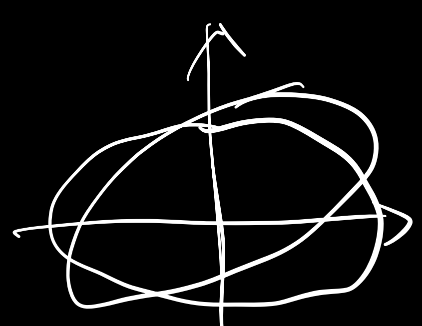
$$\le \sum_{t=1}^{T} \|x_t\|_{\Lambda_t^{-1}} \cdot \left(V_{max} \sqrt{\cdots} \right).$$

$$= \sum_{t=1}^{T} \sqrt{x_t^T \Lambda_t^{-1} x_t} \cdot (\cdots).$$

$$= T \cdot \frac{\sum_{t=1}^{T} \sqrt{x_t^T \Lambda_t^{-1} x_t}}{T} \le \sqrt{T} \cdot \sqrt{\sum_{t=1}^{T} x_t^T \Lambda_t^{-1} x_t}.$$

$$= \sqrt{T} \cdot \sqrt{2d\log(T+1)}.$$

$$\{x : x^T A x \le \beta\}.$$



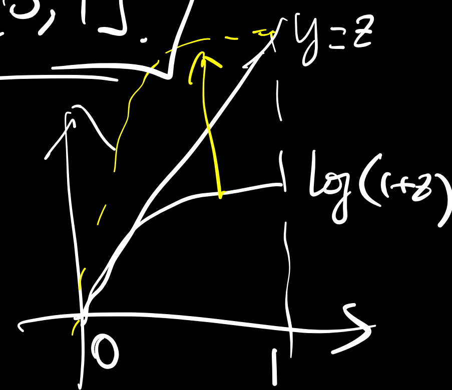
$$\left\{ x : [x_1 \; x_2] \begin{bmatrix} a^{-1} & \\ & b^{-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le 1 \right\}.$$

Lemma: (Elliptical potential lemma). $\|x_t\| \le 1$.

$$\underbrace{\sum^{T} x_t^T \Lambda_t^{-1} x_t}_{\Delta} \le 2\log\det(\Lambda_{T+1}) \le 2d\log(T+1)$$

Proof: $\Lambda_t^{-1} \preceq I \Rightarrow x_t^T \Lambda_t^{-1} x_t \le 1.$

$$\boxed{z \le 2\log(1+z) \quad \forall z \in [0,1]}$$



$$\log\left(\frac{\det(\Lambda_{t+1})}{\det(\Lambda_t)}\right) \approx x_t^T \Lambda_t^{-1} x_t.$$

$$\det(\Lambda_{t+1}) = \det(\Lambda_t + x_t x_t^T).$$

$$= \det\left(\Lambda_t^{1/2}\left(I + \Lambda_t^{-1/2} x_t x_t^T \Lambda_t^{-1/2}\right)\Lambda_t^{1/2}\right).$$

$$= \det(\Lambda_t)\ \det\left(I + \underbrace{\Lambda_t^{-1/2} x_t}_{A} x_t^T \Lambda_t^{-1/2}\right).$$

$$= \det(\Lambda_t)\ \left(1 + x_t^T \Lambda_t^{-1} x_t\right).$$

$\det(I + AA^T) = \det(I + A^T A)$.

$$\log \frac{\det(\Lambda_{t+1})}{\det(\Lambda_t)} = \log(1 + \dots) \ge \frac{x_t^T \Lambda_t^{-1} x_t}{2}.$$

Claim: fix t. w.p. $\geq 1-\delta$, $\forall x \in \mathbb{R}^d$.

$$\triangleright \; \boxed{|\hat{y}_t(x) - y(x)|} \leq \boxed{\|x\|_{\Lambda_t^{-1}}} \cdot O\left(V_{max}\sqrt{d + \log\frac{1}{\delta}}\right)$$

$$\Lambda_t := I + \sum_{i=1}^{t-1} x_i x_i^T \in \mathbb{R}^{d \times d}$$

$$\sqrt{y^T \Lambda_{t-1}^{-1} y} \leq O(\ldots)$$

Lemma: w.p. $\geq 1-\delta$. $\left\|\sum_{i=1}^{t-1} x_i \varepsilon_i\right\|_{\Lambda_t^{-1}} \leq O\left(V_{max}\sqrt{d \log\frac{1}{\delta}}\right)$

$$\hat{\theta}_t = \Lambda_t^{-1} \sum_{i=1}^{t-1} x_i y_i.$$

$$|\hat{y}_t(x) - y(x)|$$
$$= \left|x^T \Lambda_t^{-1} \sum_{i=1}^{t-1} x_i y_i - x^T \theta^*\right|.$$

$$\begin{bmatrix} X \end{bmatrix} \begin{bmatrix} \theta \end{bmatrix} = \begin{bmatrix} y \end{bmatrix}$$

$$(XX^T)^{-1} XY.$$

$$\leq \left|x^T \Lambda_t^{-1} \sum_{i=1}^{t-1} x_i (y_i - x_i^T \theta^*)\right|$$
$$+ \left|x^T \Lambda_t^{-1} \sum_{i=1}^{t-1} x_i (x_i^T \theta^*) - x^T \theta^*\right|.$$

$$= \left| \boxed{x^T \Lambda_t^{-1/2}} \Lambda_t^{-1/2} \sum_{i=1}^{t-1} x_i \varepsilon_i \right|$$

PSD: A.

$= U^T S U$.

$A^{1/2} = U^T S^{1/2} U$

$$\leq \| x^T \Lambda_t^{-1/2} \| \cdot \| \Lambda_{t-1}^{-1/2} \sum_{i=1}^{t-1} x_i \varepsilon_i \|$$

$$= \|x\|_{\Lambda_t^{-1}} \cdot \left\| \sum_{i=1}^{t-1} x_i \varepsilon_i \right\|_{\Lambda_t^{-1}}$$

$\|x\|_A := \sqrt{x^T A x}$

$x^T A^{1/2} A^{1/2} x$

$= \|x^T A^{1/2}\|$

$$\left| x^T \Lambda_t^{-1} \sum_{i=1}^{t-1} x_i (x_i^T \theta^*) - x^T \theta^* \right|$$

$$= \left| \boxed{x^T \Lambda_t^{-1/2}} \Lambda_t^{-1/2} \sum_{i=1}^{t-1} (x_i x_i^T) \theta^* \right.$$

$$\left. - \boxed{x^T \Lambda_t^{-1/2} \Lambda_t^{-1/2}} \Lambda_t \theta^* \right|$$

I.

$$= \left| \underline{x^T \Lambda_t^{-1/2} \Lambda_t^{-1/2}} \left(\sum_{i=1}^{t-1} (x_i x_i^T) - \Lambda_t \right) \theta^* \right|$$

$$= \left| \underbrace{x^T \Lambda_t^{-1/2}} \underline{\Lambda_t^{-1/2} \theta^*} \right| .$$

$$\leq \|x\|_{\Lambda_t^{-1}} \cdot \|\theta^*\| .$$

$\varepsilon_i X_i$

Special case: $X_i \in \mathbb{R}$. $(d=1)^{\Delta}$.

ε_i'.

$[a_i, b_i]$.

$\sum (b_i - a_i)^2$

$$\frac{\left(\sum_{i=1}^{t-1} \varepsilon_i'\right)^2}{\sum X_i^2} \le O\left(V_{max}\sqrt{d \log \frac{1}{\delta}}\right).$$

$|X_i| \approx c$.

$$\frac{\left|\sum_{i=1}^{t-1} \varepsilon_i'\right|}{t-1} \le O\left(\frac{\sqrt{c^2 \cdot V_{max}^2 \log \frac{1}{\delta}}}{t-1}\right)$$

1