Best arm identification for Linear Bandits.

Input: $\mathcal{X} \subset \mathbb{R}^d$, $\delta \in (0,1)$

for $t = 1, 2, \ldots$

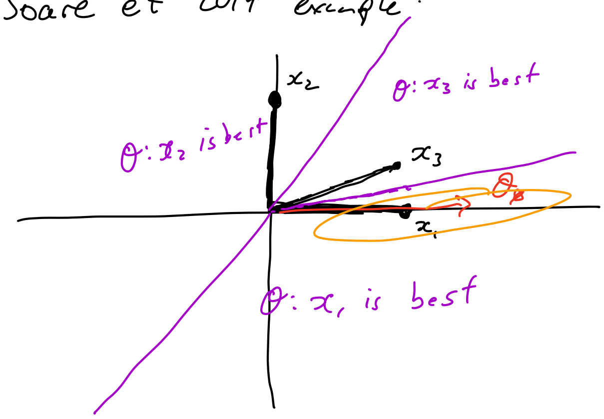
  Player chooses $x_t \in \mathcal{X}$

  Nature reveals $y_t = \langle x, \theta_* \rangle + \xi_t$ ← 1-sub-Gaussian

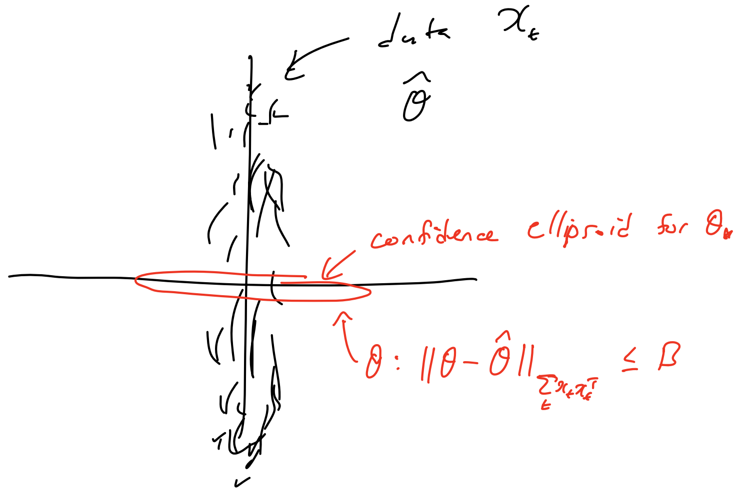
  Player optionally stops and outputs $\hat{x} \in \mathcal{X}$.

Goal: With prob $\geq 1-\delta$ output $\hat{x} = x_*$ after as few samples as possible.

Soare et 2014 example:



$$x_3 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix}$$



Suppose we take samples $x_1, \ldots, x_T$ and observe $\{y_t\}_{t=1}^T$.

Compute $\hat{\theta} = (X^T X)^{-1} X^T y$.

$$x_* = \underset{x \in \mathcal{X}}{\operatorname{argmax}} \langle x, \theta_* \rangle$$

If $\hat{x} = \underset{x \in \mathcal{X}}{\operatorname{argmax}} \langle x, \hat{\theta} \rangle$ and $x_* = \hat{x}$

then $\langle x_* - x, \hat{\theta} \rangle > 0 \quad \forall x.$

$$\langle x_* - x, \hat{\theta} \rangle = \langle x_* - x, \hat{\theta} - \theta_* \rangle + \langle x_* - x, \theta_* \rangle$$

$$\geq \|x_* - x\|_{(X^T X)^{-1}} \sqrt{2\log(1/\delta)} + \langle x_* - x, \theta_* \rangle$$

> 0 if

$$\equiv \quad \frac{\|x_* - x\|_{(X^T X)^{-1}} \sqrt{2\log(1/\delta)}}{\langle x_* - x, \theta_* \rangle} < 1 \qquad \forall x \neq x_*$$

$$\equiv \quad \max_{x \neq x_*} \frac{\|x_* - x\|^2_{(X^T X)^{-1}}}{\langle x_* - x, \theta_* \rangle^2} \cdot 2\log(1/\delta) < 1$$

$$X^T X = \sum_{t=1}^{T} x_t x_t^T$$

$$\min_{T, x_1, \ldots, x_T} T$$

$$\text{s.t.} \quad \max_{x \neq x_*} \frac{\|x_* - x\|^2_{(\sum_{t=1}^{T} x_t x_t^T)^{-1}}}{\langle x_* - x, \theta_* \rangle^2} \cdot 2\log(1/\delta) < 1$$

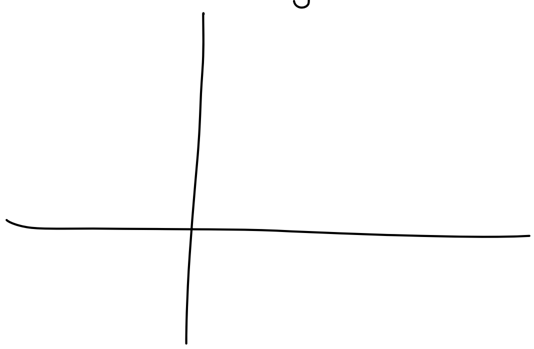
$$A(\lambda) = \sum_{x \in \mathcal{X}} \lambda_x x x^T \qquad \lambda \in \Delta_{\mathcal{X}}, \qquad T A(\lambda) \approx \sum_{t=1}^{T} x_t x_t^T$$

$$\min_{\lambda \in \Delta_{\mathcal{X}}, T} T$$

$$\text{s.t.} \quad \max_{x \neq x_*} \frac{\|x_* - x\|^2_{A(\lambda)^{-1}}}{\langle x_* - x, \theta_* \rangle^2} \cdot 2\log(1/\delta) < T$$

$$\Longrightarrow \quad T \geq \underbrace{\min_{\lambda \in \Delta_{\mathcal{X}}} \max_{x \neq x_*} \frac{\|x_* - x\|^2_{A(\lambda)^{-1}}}{\langle x_* - x, \theta_* \rangle^2}}_{:= \rho_*} \cdot 2\log(1/\delta).$$

<u>Claim</u> There exists an algorithm that identifies x_* w.p. $> 1-\delta$ and takes fewer than $\rho_* \log(|\mathcal{X}|/\delta)$ total pulls.



Input: Finite set $\mathcal{X} \subset \mathbb{R}^d$, confidence level $\delta \in (0,1)$.
Let $\mathcal{X}_1 \leftarrow \mathcal{X}, t \leftarrow 1$
while $|\mathcal{X}_\ell| > 1$ **do**
Let $\widehat{\lambda}_\ell \in \triangle_{\mathcal{X}}$ be a $\frac{d(d+1)}{2}$-sparse minimizer of $f(\lambda; \mathcal{X}_\ell)$ where
$$f(\mathcal{V}) = \inf_{\lambda \in \mathcal{X}} f(\lambda; \mathcal{V}) = \inf_{\lambda \in \mathcal{X}} \max_{x, x' \in \mathcal{V}} \|x - x'\|^2_{(\sum_{x \in \mathcal{X}} \lambda_x x x^\top)^{-1}}$$
Set $\epsilon_\ell = 2^{-\ell}$, $\tau_\ell = 2\epsilon_\ell^{-2} f(\mathcal{X}_\ell) \log(4\ell^2|\mathcal{X}|/\delta)$
Pull arm $x \in \mathcal{X}$ exactly $\lceil \tau_\ell \widehat{\lambda}_{\ell,x} \rceil$ times and construct $\widehat{\theta}_\ell$
$\mathcal{X}_{\ell+1} \leftarrow \mathcal{X}_\ell \setminus \{x \in \mathcal{X}_\ell : \max_{x' \in \mathcal{X}_\ell} \langle x' - x, \widehat{\theta}_t \rangle > \epsilon_\ell\}$
$t \leftarrow t + 1$
Output: $\mathcal{X}_{t+1}$

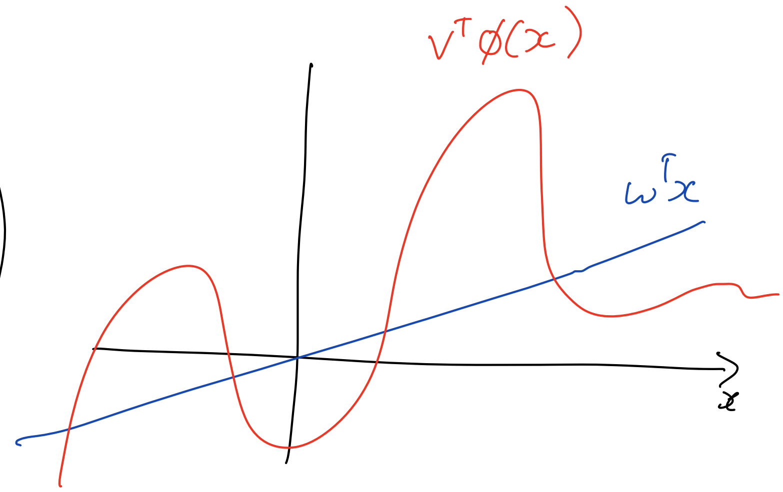
$$\begin{aligned}
\max_x \|x - x_0\| \leq \max_{x,x'} \|x - x'\| &= \max_{x,x'} \|(x - x_0) - (x' - x_0)\| \\
&\leq \max_{x,x'} \|x - x_0\| + \|x' - x_0\| \\
&= \max_x 2\|x - x_0\|
\end{aligned}$$

$x \longmapsto \phi(x) \in \mathbb{R}^p \quad \forall x \in \mathcal{X} \subset \mathbb{R}^d$

$x \in \mathbb{R}$
$\phi(x) \in \mathbb{R}^{20}$

$p >> d$

$$\phi\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \\ x_1 \\ x_2 \\ 1 \end{pmatrix}$$



For any $\phi : \mathbb{R}^d \to \mathbb{R}^p$ define $k(x, x') := \langle \phi(x), \phi(x') \rangle$.

RBF kernel $k(x, x') = e^{-\|x\|^2 / 2\gamma^2}$

$$\langle w, x \rangle \equiv \sum_i \alpha_i k(x, x_i)$$

<u>UCB</u> $\quad \hat{\theta} = (X^T X + \lambda I)^{-1} X^T y \quad \leftarrow$ from data up to time T

$$= V_t^{-1} \sum_{s=1}^{t} x_s y_s$$

$$\operatorname*{argmax}_{x \in \mathcal{X}} \langle \hat{\theta}, x \rangle + \|x\|_{V_t^{-1}} \cdot \sqrt{\beta_t}$$

$$\langle \hat{\theta}, x \rangle = x^T (X^T X + \lambda I)^{-1} X^T y$$

<u>Claim</u> $\quad (X^T X + \lambda I)^{-1} X^T = X^T (X X^T + \lambda I)^{-1}$

$$\langle \hat{\theta}, x \rangle = x^T X^T (XX^T + \lambda I)^{-1} y$$
$$= k_t(x)^T (K_t + \lambda I)^{-1} y$$

$$k_t(x) = \begin{bmatrix} k(x, x_1) \\ \vdots \\ k(x, x_t) \end{bmatrix}$$

$$[K_t]_{i,j} = k(x_i, x_j)$$

$$\|x\|^2_{V_t^{-1}} = x^T V_t^{-1} x = x^T (X^T X + \lambda I)^{-1} x$$
$$= \frac{1}{\lambda} k(x,x) - \frac{1}{\lambda} k_t(x)^T (K_t + \lambda I)^{-1} k_t(x)$$

$$(X^T X + \lambda I) b = X^T X b + \lambda b$$

$$b = (X^T X + \lambda I)^{-1} X^T X b + \lambda (X^T X + \lambda I)^{-1} b$$
$$= X^T (XX^T + \lambda I)^{-1} X b + \lambda (X^T X + \lambda I)^{-1} b$$

$$\Rightarrow \quad a^T (X^T X + \lambda I)^{-1} b = \frac{1}{\lambda} a^T b - \frac{1}{\lambda} a^T X^T (XX^T + \lambda I)^{-1} X b$$