Stochastic Linear Bandits

Input n arms $\mathcal{X} = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$

for $t = 1, 2, \ldots, T$

    Player chooses $I_t \in [n] = \{1, \ldots, n\}$

    Nature reveals $\langle x_{I_t}, \theta^* \rangle + z_t$ $\left(\mathbb{E}[z] = 0, \right.$
$\left. z - \text{sub-Gaussian} \right)$

$$x_* = \underset{x \in \mathcal{X}}{\text{argmax}} \langle x, \theta^* \rangle$$

$$
\begin{aligned}
\text{Regret}(T) &= \underset{x \in \mathcal{X}}{\max} \sum_{t=1}^{T} \langle x, \theta^* \rangle - \langle x_{I_t}, \theta^* \rangle \\
&= \sum_{t=1}^{T} \langle x_*, \theta^* \rangle - \langle x_{I_t}, \theta^* \rangle \\
&= \sum_{t=1}^{T} \langle x_* - x_{I_t}, \theta^* \rangle \\
&= \sum_{x \in \mathcal{X}} \langle x_* - x, \theta^* \rangle \, T_x \\
&= \sum_{x \in \mathcal{X}} \Delta_x \, T_x
\end{aligned}
$$

$$T_x = \sum_{t=1}^{T} \mathbb{1}\{x_{I_t} = x\} \qquad \Delta_x = \langle x_* - x, \theta_* \rangle$$

Least squares $\quad X = \begin{bmatrix} - x_1^T - \\ \vdots \\ - x_J^T - \end{bmatrix} \in \mathbb{R}^{J \times d} \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix} \in \mathbb{R}^J$

Suppose I am given a dataset $\{(x_i, y_i)\}_{i=1}^{J}$, $z = \begin{bmatrix} z_1 \\ \vdots \\ z_J \end{bmatrix} \in \mathbb{R}^J$

$$y_i = \langle x_i, \theta^* \rangle + z_i, \qquad z_i \overset{iid}{\sim} \mathcal{N}(0, 1)$$

MLE (maximum likelihood estimator) is the LS estimator:

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \sum_{t=1}^{J} (y_i - \langle x_i, \theta \rangle)^2 \qquad z \sim \mathcal{N}(0, I)$$

$$= \underset{\theta}{\text{argmin}} \ \| y - X\theta \|_2^2 \qquad \left(y = X\theta + z \right)$$

$$= (X^T X)^{-1} X^T y$$

$$= (X^T X)^{-1} X^T (X\theta_* + z)$$

$$= \theta_* + (X^T X)^{-1} X^T z$$

$\left(\text{Note } \mathbb{E}[z] = 0 \right.$

so $\mathbb{E}[\hat{\theta}] = \theta_* \Big)$

$$\mathbb{E}\left[(\hat{\theta} - \theta_*)(\hat{\theta} - \theta_*)^T \right] = \mathbb{E}\left[(X^T X)^{-1} X^T z z^T X (X^T X)^{-1} \right]$$

$$= (X^T X)^{-1} X^T \mathbb{E}[z z^T] X (X^T X)^{-1}$$
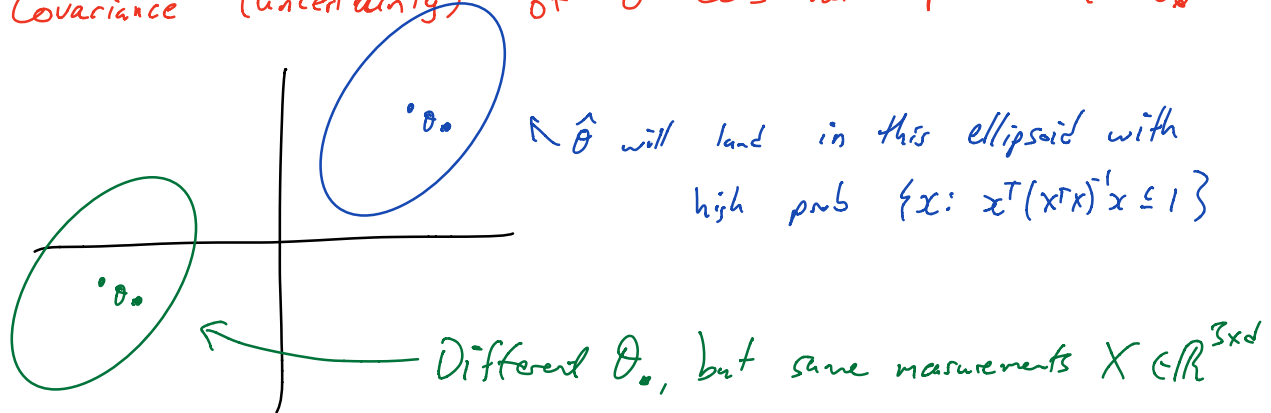
Any linear transformation of Gaussian vector is also a Gaussian vector $= (X^T X)^{-1} X^T X (X^T X)^{-1}$

$$= (X^T X)^{-1}$$

$\longrightarrow \quad \hat{\theta} \sim \mathcal{N}(\theta_*, (X^T X)^{-1}), \qquad \hat{\theta} - \theta_* \sim \mathcal{N}(0, (X^T X)^{-1})$

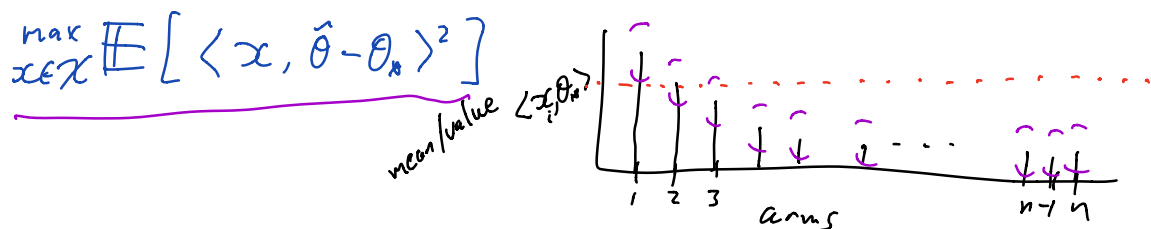Covariance (uncertainty) of $\hat{\theta}$ does not depend on θ_*

$\hookleftarrow \hat{\theta}$ will land in this ellipsoid with high prob $\{x: x^T(X^TX)^{-1}x \le 1\}$

Different θ_*, but same measurements $X \in \mathbb{R}^{3\times d}$

Since the "shape" of uncertainty <u>only</u> depends on the measurement vectors $X = \begin{bmatrix} -x_1^T- \\ \vdots \\ -x_3^T- \end{bmatrix}$, if we have a desired "shape" we can choose X to hit it.

This observation $\uparrow$ is the driving motivation for optimal linear experimental design.

To minimize regret we want to identify those arms $x \in X$ w/ $\langle x_* - x, \theta^* \rangle > 0$ ASAP to to stop playing them and reducing regret.

We don't know θ_*, but estimate gap w/ $\langle x_* - x, \hat{\theta} \rangle$

Idea: choose $X = (x_1, \ldots, x_3)$ in order to minimize

$$\max_{x \in X} \mathbb{E}\left[\langle x, \hat{\theta} - \theta_* \rangle^2 \right]$$

mean/value $\langle x_i, \theta_* \rangle$

arms
1 2 3   n-1 n

Confidence intervals

$$\hat{\theta} - \theta_* \sim \mathcal{N}(0, (X^TX)^{-1})$$

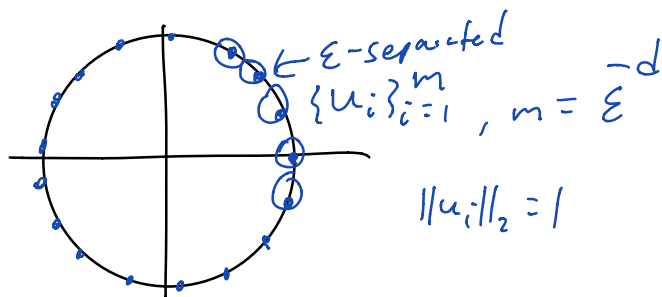For any vector $x \in \mathcal{X}$, $\langle x, \hat{\theta} - \theta_* \rangle$ is

Gaussian w/ mean 0 and variance $x^T(X^TX)^{-1}x$.

$$=: \|x\|^2_{(X^TX)^{-1}}$$

By sub-Gaussian tail-bound

$$\mathbb{P}\left(\langle x, \hat{\theta} - \theta_* \rangle \geq \|x\|_{(X^TX)^{-1}} \sqrt{2\log(1/\delta)} \right) \leq \delta.$$

(see lecture notes, ch. 2)

Any $x \in \mathbb{R}^d$.

$\leftarrow \varepsilon$-separated $\{u_i\}_{i=1}^m$, $m = \varepsilon^{-d}$

$\|u_i\|_2 = 1$

$$\mathbb{P}\left(\exists u, \|u\|_2 = 1, \langle u, \hat{\theta} - \theta_* \rangle \geq \|u\|_{(X^TX)^{-1}} \sqrt{2\log(m/\delta)} \right)$$

$$\leq \mathbb{P}\left(\bigcup_{i=1}^m \left\{ \langle u_i, \hat{\theta} - \theta_* \rangle \geq \|u_i\|_{(X^TX)^{-1}} \sqrt{2\log(m/\delta)} + \varepsilon \right\} \right)$$

$$\leq \delta$$

$$\mathbb{P}\left(\langle x, \hat{\theta} - \theta_* \rangle \geq \|x\|_{(X^TX)^{-1}} \sqrt{2\log(1/\delta)} \right) \leq \delta. \quad \textcircled{1}$$

$$\max_{x \in \mathcal{X}} \mathbb{E}\left[\langle x, \hat{\theta} - \theta_* \rangle^2 \right] = \max_{x \in \mathcal{X}} \|x\|^2_{(X^TX)^{-1}}$$

Choose $X = (x_1, \ldots, x_g)$
to minimize

For any set $(x_1, \ldots, x_g) \in \mathcal{X}$

$$\exists \ \lambda \in \Delta_{\mathcal{X}} = \left\{ z \in \mathbb{R}^{|\mathcal{X}|} : \sum_{x \in \mathcal{X}} z_x = 1, \ z_x \geq 0 \right\}$$

$$\text{w/} \quad X^TX = \sum_{t=1}^{g} x_t x_t^T = g \sum_{x \in \mathcal{X}} \lambda_x \, x x^T$$

$\Longrightarrow$ motivates relaxing to finding the

continuous / fractional design problem

$$\min_{\lambda \in \Delta_{\mathcal{X}}} \max_{x \in \mathcal{X}} \|x\|^2_{\left(\sum_{x \in \mathcal{X}} \lambda_x \, x x^T \right)^{-1}} = \min_{\lambda} f(\lambda)$$

$$A(\lambda) = \sum_{x \in \mathcal{X}} \lambda_x x x^T$$

$$f(\lambda) = \max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}} \qquad g(\lambda) = \log \det(A(\lambda))$$

Lemma 4 (Kiefer-Wolfowitz (1960)). *For any $\mathcal{X}$ with $d = dim(span(\mathcal{X}))$, there exists a $\lambda^* \in \triangle_{\mathcal{X}}$ that*

- $\max_\lambda g_D(\lambda) = g_D(\lambda^*)$

- $\min_\lambda f_G(\lambda) = f_G(\lambda^*)$

- $f_G(\lambda^*) = d$

- $support(\lambda^*) = (d+1)d/2$

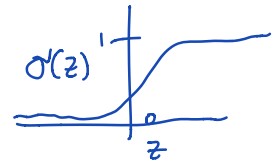$\lambda^* = \underset{\lambda}{argmin}\, f(\lambda)$

Proposition 2. *If λ^* is the G-optimal design for $\mathcal{X}$ then if we pull arm $x \in \mathcal{X}$ exactly $\lceil \tau \lambda_x^* \rceil$ times for some $\tau > 0$ and compute the least squares estimator $\widehat{\theta}$. Then for each $x \in \mathcal{X}$ we have with probability at least $1 - \delta$*

$$\langle x, \widehat{\theta} - \theta^* \rangle \leq \|x\|_{(\sum_{x \in \mathcal{X}} \lceil \tau \lambda_x^* \rceil xx^\top)^{-1}} \sqrt{2 \log(1/\delta)} \quad \text{①}$$

$$\leq \frac{1}{\sqrt{\tau}} \|x\|_{(\sum_{x \in \mathcal{X}} \lambda_x^* xx^\top)^{-1}} \sqrt{2 \log(1/\delta)}$$

$$\leq \sqrt{\frac{2d \log(1/\delta)}{\tau}}$$

and we have taken at most $\tau + \frac{d(d+1)}{2}$ pulls. Thus, for any $\delta' \in (0,1)$ we have $\mathbb{P}(\bigcup_{x \in \mathcal{X}}\{|\langle x, \widehat{\theta} - \theta^* \rangle| > \sqrt{\frac{2d \log(2|\mathcal{X}|/\delta')}{\tau}}\}) \leq \delta'$.

$$y_t \sim \text{Bernoulli}\left(\sigma(\langle x_t, \theta_* \rangle)\right), \qquad \sigma'(z) = \frac{1}{1 + e^{-z}}$$

$$\widehat{\theta}_3 \; \text{MLE of} \; \{(x_t, y_t)\}_{t=1}^3$$

$\sigma(z)$

As $3 \to \infty$ we have by CLT

$$\widehat{\theta}_3 \longrightarrow \mathcal{N}\left(\theta_*, \left(\sum_{t=1}^3 \sigma(\langle x_t, \theta_* \rangle)(1 - \sigma(\langle x_t, \theta_* \rangle)) x_t x_t^\top\right)^{-1}\right)$$

Input: Finite set $\mathcal{X} \subset \mathbb{R}^d$, confidence level $\delta \in (0,1)$.
Let $\widehat{\mathcal{X}}_1 \leftarrow \mathcal{X}, \ell \leftarrow 1$
while $|\mathcal{X}_\ell| > 1$ **do**
  Let $\widehat{\lambda}_\ell \in \triangle_{\mathcal{X}_\ell}$ be a $\frac{d(d+1)}{2}$-sparse minimizer of $f(\lambda) = \max_{x \in \mathcal{X}_\ell} \|x\|^2_{(\sum_{x \in \mathcal{X}_\ell} \lambda_x x x^\top)^{-1}}$ ⎤ *Choosing*
                                                                                                                                                              ⎦ *a design*
  $\epsilon_\ell = 2^{-\ell}$, $\tau_\ell = 2d\epsilon_\ell^{-2} \log(4\ell^2 |\mathcal{X}|/\delta)$
  Pull arm $x \in \mathcal{X}$ exactly $\lceil \widehat{\lambda}_{\ell,x}\tau_\ell \rceil$ times and construct the least squares estimator $\widehat{\theta}_\ell$ using only the observations of this round
  $\mathcal{X}_{\ell+1} \leftarrow \mathcal{X}_\ell \setminus \{x \in \mathcal{X}_\ell : \max_{x' \in \mathcal{X}_\ell} \langle x' - x, \widehat{\theta}_\ell \rangle > 2\epsilon_\ell\}$
  $\ell \leftarrow \ell + 1$
Output: $\mathcal{X}_\ell$

Lemma 5. *Assume that* $\max_{x \in \mathcal{X}} \langle x^\star - x, \theta^* \rangle \leq 4$. *With probability at least* $1 - \delta$, *we have* $x^\star \in \mathcal{X}_\ell$ *and* $\max_{x \in \mathcal{X}_\ell} \langle x^\star - x, \theta^* \rangle \leq 8\epsilon_\ell$ *for all* $\ell \in \mathbb{N}$.