Stochastic Linear Bandits

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

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

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

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

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

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

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

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

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

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

MLE (maximum likelihood estimator) is the LS estimator:

$$\hat{\theta} = \arg\min_\theta \sum_{i=1}^S (y_i - \langle x_i, \theta \rangle)^2 \qquad z \sim \mathcal{N}(0, I)$$

$$= \arg\min_\theta \| y - X\theta \|_2^2 \qquad (y = X\theta + z)$$

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

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

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

(Note $\mathbb{E}[z] = 0$ so $\mathbb{E}[\hat{\theta}] = \theta_*$)

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

$$= (X^TX)^{-1} X^T \mathbb{E}[z z^T] X (X^TX)^{-1}$$

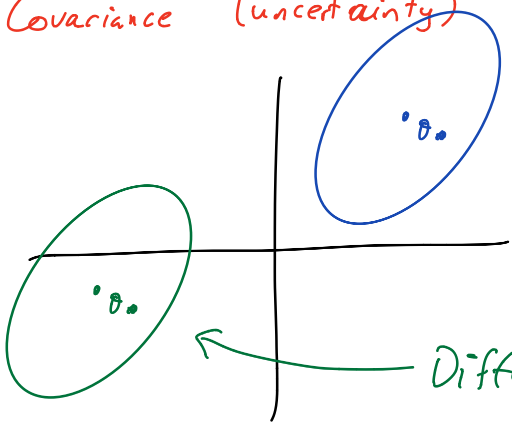
$$= (X^TX)^{-1} X^T X (X^TX)^{-1}$$

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

Any linear transformation of Gaussian vector is also a Gaussian vector

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

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



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

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

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

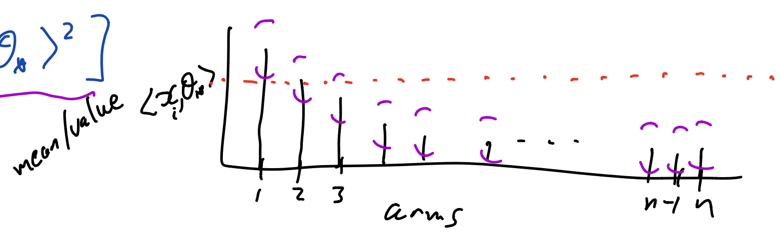
This observation $\uparrow$ is the driving motivation for optimal linear experimental design.

To minimize regret we want to identify those arms $x \in \mathcal{X}$ 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 $X = (x_1, \dots, x_J)$ in order to minimize

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



Confidence intervals

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

For any vector $x \in \mathcal{X}$, $\langle x, \hat{\theta} - \theta_* \rangle$ is Gaussian w/ mean 0 and variance $x^T (X^T X)^{-1} x$.

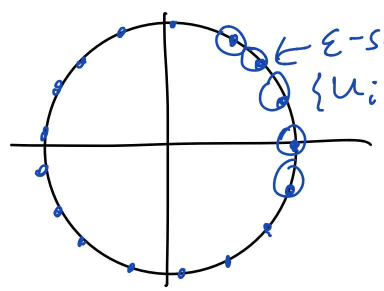
$$=: \|x\|^2_{(X^T X)^{-1}}$$

By sub-Gaussian tail-bound

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

(see lecture notes, ch. 2)

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



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

$\|u_i\|_2 = 1$

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

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

$$\leq \delta$$

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

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

Choose $X = (x_1, \dots, x_\tau)$ to minimize

For any set $(x_1, \dots, x_\tau) \in \mathcal{X}$

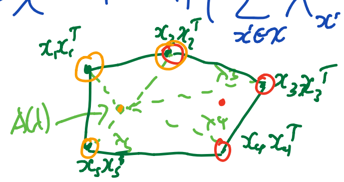
$\exists \; \lambda \in \Delta_{\mathcal{X}} = \{ z \in \mathbb{R}^{|\mathcal{X}|} : \sum_{x\in\mathcal{X}} z_x = 1, \; z_x \geq 0 \}$

w/ $X^TX = \sum_{t=1}^{\tau} x_t x_t^T = \tau \sum_{x\in\mathcal{X}} \lambda_x x x^T$

$\Rightarrow$ motivates relaxing to finding the continuous / fractional design problem

$$\min_{\lambda\in\Delta_{\mathcal{X}}} \max_{x\in\mathcal{X}} \|x\|^2_{\left(\sum_{x'\in\mathcal{X}} \lambda_{x'} x' x'^T\right)^{-1}} = \min_{\lambda} f(\lambda)$$

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



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

$$g(\lambda) = \log\det(A(\lambda))$$

$Tr(AB) = Tr(BA)$

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

$$f(\lambda) = \max_{x \in \mathcal{X}} \|x\|^2_{A(\lambda)^{-1}} \geq \sum_{x\in\mathcal{X}} \lambda_x \|x\|^2_{A(\lambda)^{-1}} = \sum_{x\in\mathcal{X}} \lambda_x \, x^T \Big(\sum_{x'} \lambda_{x'} x' x'^T\Big)^{-1} x$$

$$= \sum_{x\in\mathcal{X}} \lambda_x Tr\Big(x x^T \Big(\sum_{x'} \lambda_{x'} x' x'^T\Big)^{-1}\Big)$$

$$= Tr\Big(\Big(\sum_{x'} \lambda_{x'} x' x'^T\Big)\Big(\sum_{x'} \lambda_{x'} x' x'^T\Big)^{-1}\Big) = Tr(I) = d$$

- $\max_\lambda g_D(\lambda) = g_D(\lambda^*)$
- $\min_\lambda f_G(\lambda) = f_G(\lambda^*)$
- $f_G(\lambda^*) = d$
- *support*$(\lambda^*) \leq (d+1)d/2$. $\lambda^* = \arg\min_\lambda f(\lambda)$

Caratheodory Theorem

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 $1 - \delta$*

$$X^T X = \sum_{x\in\mathcal{X}} \lceil \lambda^*_x \tau \rceil x x^T$$

$$\begin{aligned} \langle x, \widehat{\theta} - \theta^* \rangle &\leq \|x\|_{(\sum_{x\in\mathcal{X}} \lceil \tau\lambda^*_x \rceil x x^\top)^{-1}} \sqrt{2\log(1/\delta)} \quad (1) \\ &\leq \frac{1}{\sqrt{\tau}} \|x\|_{(\sum_{x\in\mathcal{X}} \lambda^*_x x x^\top)^{-1}} \sqrt{2\log(1/\delta)} \\ &\leq \sqrt{\frac{2d\log(1/\delta)}{\tau}} \end{aligned}$$

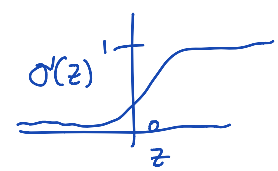
$$= z^T\Big(\sum_x \lceil \lambda_x \tau \rceil x x^T\Big)^{-1} z \leq z^T\Big(\sum_x \lambda_x \tau x x^T\Big)^{-1} z$$

$$= \frac{1}{\tau} z^T\Big(\sum \lambda_x x x^T\Big)^{-1} z$$

$$= \frac{1}{\tau} \|z\|^2_{(\sum \lambda_x x x^T)}$$

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

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



$\widehat{\theta}_\tau$ MLE of $\{(x_t, y_t)\}_{t=1}^{\tau}$

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

$$\widehat{\theta}_\tau \to \mathcal{N}\Big(\theta_*, \Big(\sum_{t=1}^{\tau} \sigma(\langle x_t, \theta_* \rangle)(1 - \sigma(\langle x_t, \theta_* \rangle)) x_t x_t^T\Big)^{-1}\Big)$$

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 $\underline{\widehat{\lambda}_\ell \in \triangle}_{\mathcal{X}_\ell}$ 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 $x \in \mathcal{X}_\ell$ 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: $\mathcal{X}_\ell$

Theorem | w.p. $\geq 1-\delta$ we have after T total pulls

$$\max_{x \in \mathcal{X}} \sum_{t=1}^{T} \langle x - x_{I_t}, \theta^* \rangle \underset{\substack{\uparrow \\ \text{ignores loglog() factors and constants}}}{\lesssim} \min \left\{ \sqrt{dT \log(|\mathcal{X}|/\delta)}, \frac{d \log(|\mathcal{X}|/\delta)}{\Delta} \right\}$$

$$\Delta = \min_{x \neq x_*} \Delta_x$$

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* $\max_{x \in \mathcal{X}_\ell} \underbrace{\langle x^\star - x, \theta^* \rangle}_{=\Delta_x} \leq 8\epsilon_\ell$ *for all* $\ell \in \mathbb{N}$.

$$\mathcal{E}_{x,\ell} = \left\{ |\langle x, \hat{\theta}_\ell - \theta_* \rangle| \leq \varepsilon_\ell \right\}$$

$$\mathbb{P}\left(\bigcup_{\ell=1}^{\infty} \bigcup_{x \in \mathcal{X}_\ell} \mathcal{E}^c_{x,\ell} \right) \leq \sum_{\ell=1}^{\infty} \mathbb{P}\left(\bigcup_{x \in \mathcal{X}_\ell} \mathcal{E}^c_{x,\ell} \right)$$

$$= \sum_{\ell=1}^{\infty} \sum_{V \subset \mathcal{X}} \mathbb{P}\left(\bigcup_{x \in V} \mathcal{E}^c_{x,\ell}(V), \mathcal{X}_\ell = V \right)$$

$$= \sum_{\ell=1}^{\infty} \sum_{V \subset \mathcal{X}} \underbrace{\mathbb{P}\left(\bigcup_{x \in V} \mathcal{E}^c_{x,\ell}(V) \right)}_{\leq \delta/2\ell^2} \mathbb{P}(\mathcal{X}_\ell = V)$$

$$\leq \sum_{\ell=1}^{\infty} \frac{\delta}{2\ell^2} \underbrace{\sum_{V} \mathbb{P}(\mathcal{X}_\ell = V)}_{=1}$$

$$\leq \delta$$

Assume $\mathcal{E}_{x,\ell}$ for all ℓ, $x \in \mathcal{X}_\ell$.

1st step show $x^* \in \mathcal{X}_\ell \quad \forall \ell$

For any x'

$$\langle x' - x^*, \hat{\theta}_\ell \rangle = \underbrace{\langle x' - x^*, \theta^* \rangle}_{\leq 0} + \underbrace{\langle x' - x^*, \hat{\theta}_\ell - \theta^* \rangle}_{\leq \varepsilon_\ell + \varepsilon_\ell}$$

$$\leq 2\varepsilon_\ell$$

On the other hand, if $\Delta_x > 4\varepsilon_\ell$

then

$$\max_{x' \in \mathcal{X}_\ell} \langle x' - x, \hat{\theta}_\ell \rangle \geq \langle x^* - x, \hat{\theta}_\ell \rangle$$

$$= \underbrace{\langle x^* - x, \theta^* \rangle}_{> 4\varepsilon_\ell} + \underbrace{\langle x^* - x, \hat{\theta}_\ell - \theta^* \rangle}_{\geq -2\varepsilon_\ell}$$

$$> 2\varepsilon_\ell$$

Pick $\nu \geq 0$

$\lceil z \rceil \leq z+1$

$$\sum_{x\in\mathcal{X}\setminus x^\star} \Delta_x T_x = \sum_{x\in\mathcal{X}\setminus x^\star:\Delta_x\leq\nu} \Delta_x T_x + \sum_{x\in\mathcal{X}\setminus x^\star:\Delta_x>\nu} \Delta_x T_x$$

$$\leq \nu T + \sum_{\ell=1}^{\infty} \sum_{x\in\mathcal{X}_\ell\setminus x^\star:\Delta_x>\nu} \Delta_x \lceil \tau_\ell \widehat{\lambda}_{\ell,x} \rceil$$

$$\sum_{x\in\mathcal{X}_\ell} \hat{\lambda}_{x,\ell} = 1$$

$$\leq T\nu + \sum_{\ell=1}^{\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil} 8\epsilon_\ell(|\mathrm{support}(\widehat{\lambda}_\ell)| + \tau_\ell)$$

$$\sum_{\ell=1}^{L} 2^\ell \leq 2^{L+1}$$

$$= T\nu + \sum_{\ell=1}^{\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil} 8\epsilon_\ell(\tfrac{(d+1)d}{2} + 2d\epsilon_\ell^{-2}\log(4\ell^2|\mathcal{X}|/\delta))$$

$$\leq T\nu + 4(d+1)d\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil + \sum_{\ell=1}^{\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil} 16d\epsilon_\ell^{-1}\log(4\ell^2|\mathcal{X}|/\delta)$$

$$\leq T\nu + 4(d+1)d\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil + 16d\log(4\log_2^2(16(\Delta\vee\nu)^{-1})|\mathcal{X}|/\delta) \sum_{\ell=1}^{\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil} 2^\ell$$

$$\leq T\nu + 4(d+1)d\lceil \log_2(8(\Delta\vee\nu)^{-1})\rceil + 512d(\Delta\vee\nu)^{-1}\log(4\log_2^2(16(\Delta\vee\nu)^{-1})|\mathcal{X}|/\delta)$$

If $\nu = 0 \implies \dfrac{d\log(|\mathcal{X}|/\delta)}{\Delta}$

$\nu = \sqrt{\dfrac{d\log(|\mathcal{X}|/\delta)}{T}} \implies \sqrt{dT\log(|\mathcal{X}|/\delta)}$.

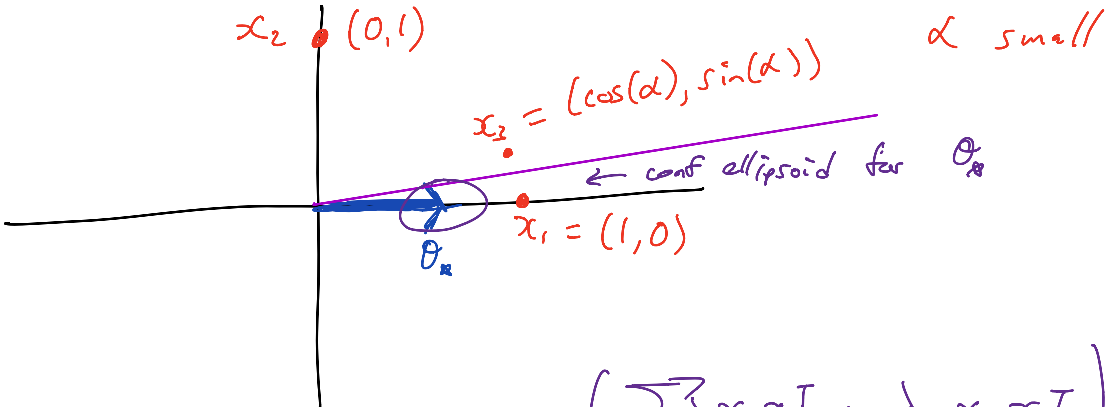
Is Tight? Consider $\mathcal{X} = \{e_i\}_{i=1}^d$ $\quad e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

ith position

For this set we showed for MAB

$$R(T) \leq \sum_{i=2}^{d} \frac{1}{\Delta_i}\log(d/\delta) \qquad \Delta_i = \theta_1^* - \theta_i^*$$

Assume $\theta_1^* > \theta_2^* \geq \dots \geq \theta_d^*$



$$\left(\sum \lambda_1 x_1 x_1^T + \lambda_3 x_3 x_3^T \right)$$

Trying to estimate $\langle \theta^*, x_1 - x_3 \rangle$