n arms, sub-Gaussian distributions

means $\theta_1 > \theta_2 \geq \dots \geq \theta_n$ $\qquad \Delta_i = \theta_1 - \theta_i$

Best arm ID: w.p. $\geq 1-\delta$ identify $\arg\max_i \theta_i$

Elim Alg accomplishes this in $\sum_i \Delta_i^{-2} \log\left(\frac{n \log(\Delta_i^{-2})}{\delta}\right)$

Regret: $\max_i \mathbb{E}\left[\sum_{t=1}^T X_{t,i} - X_{t,I_t}\right] \leq \min\left\{\sqrt{nT\log(T)}, \sum_i \Delta_i^{-1} \log(T)\right\}$

How do we argue that these results

are nearly unimprovable?

Suppose $X_1, \dots, X_n \sim \mathcal{N}(\mu, 1)$ where

$H_0: \mu = 0$

$H_1: \mu = \varepsilon$

Let ψ be a decision function that takes $X_1, \dots, X_n$

as input and outputs $\psi \in \{0,1\}$.

$$\inf_\psi \max\left\{ \mathbb{P}_0(\psi = 1), \mathbb{P}_1(\psi = 0)\right\} \geq \frac{1}{4}\exp(-n\varepsilon^2/2)$$

$\Rightarrow$ Any test that succeeds w/ prob $1-\delta$

requires $n \geq 2\varepsilon^{-2}\log\left(\frac{1}{4\delta}\right)$.

Note: consider the decision

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\psi(X_1, \dots, X_n) = \underset{i=0,1}{\operatorname{argmin}} \left(\hat{\mu}_n - \mu_i\right)^2$$

$\mu_0 = 0$

$\mu_1 = \Delta$

Succeeds if $|\hat{\mu}_n - \mu| \le \frac{\Delta}{2}$

By Chernoff $\mathbb{P}\left(|\hat{\mu}_n - \mu| > \frac{\Delta}{2}\right) \le 2\exp\left(-n\left(\frac{\Delta}{2}\right)^2 \cdot \frac{1}{2}\right)$

$$= 2\exp\left(-\frac{n\Delta^2}{8}\right)$$

$$= \delta$$

If $n \ge 8\Delta^{-2}\log(2/\delta)$ then this decision function succeeds w/ prob $\ge 1-\delta$.

Lets show $\inf_{Alg} \max_{\theta \in \mathbb{R}^n} R_T(\theta) \geq c\sqrt{nT}$

$$\inf_{Alg} \max_{\theta \in \mathbb{R}^n} R_T(\theta) \geq \inf_{Alg} \max_{i=0,1} R_T(\theta^{(i)})$$

$$\theta^{(0)} = (\Delta, 0, \ldots, 0)$$

$$\theta^{(1)} = (\Delta, 0, \ldots, 0, 2\Delta, 0, \ldots, 0)$$

$\hat{i}$ location

Fix any alg, set $\hat{i} = \underbrace{\arg\min_{i=1,\ldots,n} \mathbb{E}_0[T_i]}_{\leq T/n}$

$\underline{\Delta \simeq \sqrt{\frac{n}{T}}}$

If $\hat{i}$ is played fewer than $\simeq \Delta^{-2} = \frac{T}{n}$ times, then any algorithm cannot tell if its playing against θ^0 or $\theta^{(1)}$

$\Downarrow$

$$\underline{\mathbb{E}_0[T_i] \simeq \mathbb{E}_1[T_i]}$$

If $\mathbb{E}_0[T_1] \le T/2 \Rightarrow R^T(\theta^{(0)}) \ge \Delta \cdot T/2 = \sqrt{nT}$

$$\max\{R^T(\theta^{(0)}), R^T(\theta^{(1)})\}$$

Case 1

$$\ge R^T(\theta^{(0)}) \ge \sqrt{nT}$$

If $\mathbb{E}_0[T_1] > T/2$ and $\mathbb{E}_0[T_1] \approx \mathbb{E}_1[T_1]$

$\Rightarrow \mathbb{E}_1[T_1] > T/2 \Rightarrow \mathbb{E}_1[T_{\hat{i}}] \le T/2$

$$R(\theta^{(1)}) \ge \Delta\, T/2 = \sqrt{nT}$$

Case 2

$$\max\{R^T(\theta^{(0)}), R^T(\theta^{(1)})\}$$

$$\ge R^T(\theta^{(1)})$$

$$\ge \sqrt{nT}$$

$$\text{Fix } \overset{\text{known}}{x_1, \dots, x_n} \in \mathbb{R}^d \text{ and } \overset{\text{unknown}}{\theta_*} \in \mathbb{R}^d.$$

$\forall i$ Observe

$$y_i = x_i^T \theta_* + \varepsilon_i, \quad \{\varepsilon_i\}_i \text{ are IID mean-zero 1-sub-Gaussian R.V.}$$

ex: Best-arm ID $\operatorname{argmax}_{i=1,\dots,n} \langle x_i, \theta_* \rangle$

$$\hat{\theta} = \operatorname{argmin}_\theta \sum_{i=1}^n (\langle x_i, \theta \rangle - y_i)^2$$

$$= \left(\sum_{i=1}^n x_i x_i^T\right)^{-1} \sum_{i=1}^n x_i y_i$$

$$= \left(\sum_{i=1}^n x_i x_i^T\right)^{-1} \left(\sum_{i=1}^n x_i x_i^T\right)\theta_* + \left(\sum_{i=1}^n x_i x_i^T\right)^{-1} \sum_{i=1}^n x_i \varepsilon_i$$

$$= \theta_* + \left(\sum_{i=1}^n x_i x_i^T\right)^{-1} \sum_{i=1}^n x_i \varepsilon_i$$

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}\left[\theta_* + \left(\sum_{i=1}^n x_i x_i^T\right)^{-1} \sum_{i=1}^n x_i \varepsilon_i\right] = \theta_* + \left(\sum_{i=1}^n x_i x_i^T\right) \sum_i x_i \underbrace{\mathbb{E}[\varepsilon_i]}_{=0 \text{ by assumption}} = \theta_*$$

Motivate by best arm ID, estimate $\langle z, \theta_* \rangle$ for any fixed z.

How much error is in $\langle z, \hat{\theta} \rangle$?

$$\langle z, \hat{\theta} - \theta_* \rangle = z^T \left(\sum_{i=1}^n x_i x_i^T \right)^{-1} \sum_i x_i \varepsilon_i$$
$$= z^T (X^T X)^{-1} X^T \varepsilon$$
$$=: w^T \varepsilon \quad \text{where } \underline{w = X(X^T X)^{-1} z}$$

$$X = \begin{bmatrix} - & x_1^T & - \\ & \vdots & \\ - & x_n^T & - \end{bmatrix} \in \mathbb{R}^{n \times d} \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\hat{\theta} = (X^T X)^{-1} X^T Y = \theta_* + (X^T X)^{-1} X^T \varepsilon$$

$$\implies \langle z, \hat{\theta} - \theta_* \rangle = w^T \varepsilon = \sum_{i=1}^n w_i \varepsilon_i$$

Show $\langle z, \hat{\theta} - \theta_* \rangle$ is sub-Gaussian to construct tail bound.

Recall: R.V. $U \in \mathbb{R}$ is σ^2-sub-Gaussian if

$$\mathbb{E}[\exp(\lambda U)] \leq \exp(\sigma^2 \lambda^2 / 2) \quad \forall \lambda \in \mathbb{R}$$

$$\mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n} w_i \varepsilon_i\right)\right] = \mathbb{E}\left[\prod_{i=1}^{n} \exp(\lambda w_i \varepsilon_i)\right]$$

$$= \prod_{i=1}^{n} \mathbb{E}\left[\exp(\lambda w_i \varepsilon_i)\right] \quad (\varepsilon_i \text{ are independant})$$

$$\leq \prod_{i=1}^{n} \exp\left(\lambda^2 w_i^2 / 2\right) \quad (\varepsilon_i \text{ are 1-sub Gaussian})$$

$$= \exp\left(\left(\sum_{i=1}^{n} w_i^2\right) \lambda^2 / 2\right)$$

$$= \exp\left(\|w\|_2^2 \lambda^2 / 2\right)$$

$\Longrightarrow$ $\langle z, \hat{\theta} - \theta_* \rangle = \sum_{i=1}^{n} w_i \varepsilon_i$ is $\|w\|_2^2$-sub-Gaussian.

Chernoff Bound

$\Longrightarrow$

$$\mathbb{P}\left(\langle z, \hat{\theta} - \theta_* \rangle > t\right) = \mathbb{P}\left(\sum_i w_i \varepsilon_i > t\right)$$

$$\leq \exp\left(-\frac{t^2}{2\|w\|_2^2}\right).$$

$$\|w\|_2^2 = w^T w = z^T (X^T X)^{-1} X^T \cdot X (X^T X)^{-1} z$$

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

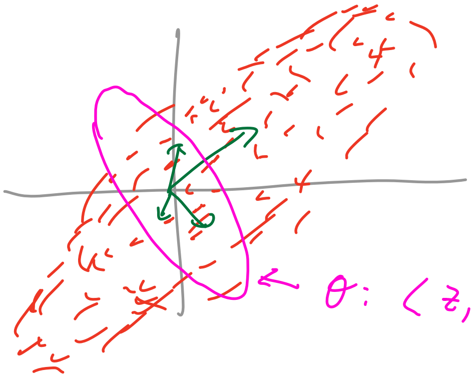
$$=: \|z\|_{(X^T X)^{-1}}^2$$

For any PSD matrix A define $\|z\|_A^2 = z^T A z$.

A is PSD (positive semi definite) if $x^T A x \geq 0 \quad \forall x$

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

$x_1, \dots, x_n$

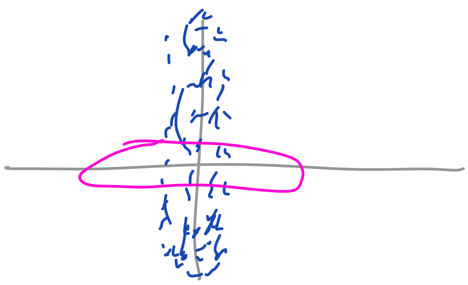


ith dof is $x_i \in \mathbb{R}^2$

pick direction z then

$\langle z, \hat{\theta} - \theta_* \rangle \lesssim \|z\|_{(X^TX)^{-1}}$

$\theta : \langle z, \theta - \theta_* \rangle \le \|z\|_{(X^TX)^{-1}}$



Different $x_1, \dots, x_n$ results in different uncertainty of $\hat{\theta} - \theta_*$

$$\mathrm{Cov}(\hat{\theta} - \theta_*) = (X^TX)^{-1}$$

The uncertainty of $\hat{\theta} - \theta_*$ in any direction only depends on $x_1, \dots, x_n$ and <u>not</u> $y_1, \dots, y_n$ or θ_*

Preview:
$$R_T = \max_i \mathbb{E}\left[\sum_{t=1}^T y_{t,i} - y_{t,I_t} \right]$$
$$= \max_i \mathbb{E}\left[\sum_{t=1}^T \langle x_i, \theta_* \rangle - \langle x_{I_t}, \theta_* \rangle \right]$$
$$= \max_i \mathbb{E}\left[\sum_t \langle x_i - x_{I_t}, \theta_* \rangle \right]$$

We will show an alg. that achieves

$$R_T \le \sqrt{dT\log(nT)}\ .$$

What if the model is wrong?

Suppose expected reward from arm $x \in \mathcal{X}$

is $\langle \theta_*, x \rangle + \underbrace{\xi_x}_{\text{model-specification}} = \mu_x$

If $\max_{x\in\mathcal{X}} |\xi_x| \le h$ then the same alg

satisfies $\max_i \mathbb{E}\left[\sum_{t=1}^{T} \mu_{x_i} - \mu_{x_{I_t}}\right] \le \sqrt{dT\log(nT)} + Th\sqrt{d}$

Observe $\mu_x + \varepsilon_t$