Suppose $\mathcal{X}$ is uncountable. Can we still bound

Choose $x_1, \dots, x_T$ measure $y_t = \langle x_t, \theta_* \rangle + \eta_t$

(G-optimal)

η_t: 1-sub-Gaussian

$$\hat{\theta} = \left(\sum_t x_t x_t^T \right)^{-1} \sum_t x_t y_t$$

$$= \theta_* + \left(\sum_t x_t x_t^T \right)^{-1} \sum_t x_t \eta_t \qquad A = \sum_t x_t x_t^T$$

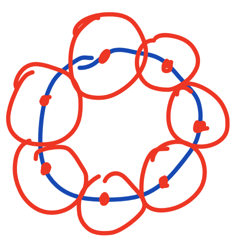
$$\max_{x \in \mathcal{X}} |\langle \hat{\theta} - \theta_*, x \rangle| = \max_x |(\hat{\theta} - \theta_*)^T A^{1/2} A^{-1/2} x|$$

$$\leq \max_x \|A^{1/2}(\hat{\theta} - \theta_*)\|_2 \cdot \|A^{-1/2} x\|_2$$

$$= \|\hat{\theta} - \theta_*\|_A \cdot \underbrace{\max_x \|x\|_{A^{-1}}}_{\leq \sqrt{d/T}} \lesssim \sqrt{d + 2\log(1/\delta)} \cdot \sqrt{d/T} \approx \sqrt{\frac{d^2}{T}}$$

$$\|\hat{\theta} - \theta_*\|_A = \|A^{1/2}(\hat{\theta} - \theta_*)\|_2 = \sup_{u : \|u\|_2 \leq 1} u^T A^{1/2} (\hat{\theta} - \theta_*)$$

Lemma $\exists$ cover $C \subset \mathbb{R}^d$: $|C| \leq (3/\varepsilon)^d$ and $x \in C$, $\|x\|_2 = 1$ and $\forall x$ w/ $\|x\|_2 = 1$ we have $\|y - x\|_2 \leq \varepsilon$ for some $y \in C$.



In other words $\bigcup_{x \in C} B(x, \varepsilon) \supset S^{d-1}$

sphere in d-dim

$$\|\hat{\theta} - \theta_*\|_A = \sup_{u : \|u\|_2 \leq 1} u^T A^{1/2} (\hat{\theta} - \theta_*)$$

$$= \sup_u \min_{y \in C} (u - y)^T A^{1/2} (\hat{\theta} - \theta_*) + y^T A^{1/2} (\hat{\theta} - \theta_*)$$

$$\leq \sup_u \underbrace{\min_{y \in C} \|u - y\|_2}_{\leq \varepsilon} \|\hat{\theta} - \theta_*\|_A + \underbrace{y^T A^{1/2} (\hat{\theta} - \theta_*)}_{\leq \sqrt{2\log\left(\frac{|C|}{\delta}\right)}}$$

$$\leq \varepsilon \|\hat{\theta}-\theta_*\|_A + \sqrt{2\log\left(\frac{|C|}{\delta}\right)}$$

$$\Longrightarrow \|\hat{\theta}-\theta_*\|_A \leq \frac{1}{1-\varepsilon}\cdot\sqrt{2\log\left(\frac{|C|}{\delta}\right)} \leq \frac{1}{1-\varepsilon}\cdot\sqrt{2d\log(3/\varepsilon)+2\log(1/\delta)}.$$

$$\mathbb{E}\left[\left(y^T A^{1/2}(\hat{\theta}-\theta_*)\right)^2\right] = \mathbb{E}\left[y^T A^{1/2} A^{-1}\left(\sum_t x_t z_t\right)\left(\sum_s x_s z_s\right)^T A^{-1} A^{1/2} y\right]$$

$$= y^T A^{-1/2}\, \underbrace{\mathbb{E}\left[\sum_t x_t x_t^T z_t^2\right]}_{=A} A^{-1/2} y$$

$$= y^T y$$

$$= 1$$

$$\mathbb{E}[z_t^2]=1$$

Stochastic Processes

$$\left\| \mathbb{E}\left[\sum_t x_t x_t^T z_t^2\right]^{1/2} A^{-1/2} y \right\|_2$$

$$\leq$$

Flip two coins

Sample/outcome space $\Omega = \{HH, HT, TH, TT\}$

σ-algebra of a set Ω called $\mathcal{F}$ satisfies

1. $\Omega \in \mathcal{F}$,
2. $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$
3. $F_n \in \mathcal{F}\ \forall n$ then $\bigcup_n F_n \in \mathcal{F}$.

Probability measure $\mathbb{P}$ assign numbers to $\mathcal{F}$ ($\mathbb{P}: \mathcal{F} \to \mathbb{R}$)

$\mathbb{P}(F) \geq 0 \quad \forall F \in \mathcal{F}$

$\mathbb{P}(\emptyset) = 0 \quad \mathbb{P}(\Omega) = 1$,

F_n are disjoint $\forall n$

$$\mathbb{P}\left(\bigcup_n F_n\right) = \sum_n \mathbb{P}(F_n)$$

Random variable X is a function mapping

$\Omega \to \mathbb{R}$.

Conditional expectation

$$\mathbb{P}(Y|X) = \frac{\mathbb{P}(Y, X)}{P(X)}$$

If X is finite then

$$\mathbb{E}[Y|X=x] = \sum_y y\,\mathbb{P}(Y=y|X=x]$$

but otherwise

$\mathbb{E}[Y|X](\omega) = g(\omega)$ where g is any function

satisfying

$$\int_{\omega \in H} g(\omega)\, d\mathbb{P}(\omega) = \int_{\omega \in H} Y(\omega)\, d\mathbb{P}(\omega)$$

for all $H \in \sigma(X)$

$X_1, X_2, X_3, \ldots$

$\mathcal{F}_t = \sigma(\{X_s\}_{s \le t})$. $\quad \mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \ldots$

Filtration : $\{\mathcal{F}_t\}_{t=1}^{\infty} = \mathbb{F}$

X_t is $\mathbb{F}$-adapted if X_t is $\mathcal{F}_t$ measurable.

$$\mathbb{E}[X_t \mid \mathcal{F}_t] = X_t$$

An $\mathbb{F}$-adapted sequence of R.V.s is a martingale if $\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] = X_t \ \forall t$, and $\mathbb{E}[|X_t|] < \infty$.

If $\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] \le X_t$ we say its supermartingale

If $\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] \ge X_t$ " sub-martingale.

Ex. $Z_t \sim N(0,1)$, then $X_t = \sum_{s=1}^{t} Z_s$ is a martingale.

$$\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] = \mathbb{E}\left[\sum_{s=1}^{t+1} Z_s \mid \mathcal{F}_t\right] = \mathbb{E}[Z_{t+1} \mid \mathcal{F}_t] + \sum_{s=1}^{t} Z_s = 0 + X_t$$

Ex. $Z_t = \begin{cases} 2 & \text{w.p. } 1/2 \\ 0 & \text{w.p. } 1/2 \end{cases}$, $X_t = \prod_{s=1}^{t} Z_s$, $X_0 = 1$

Ex. $Z_t \sim N(0,1)$, $X_t = \prod_{s=1}^{t} \exp(\lambda Z_s)$, $\lambda > 0$

$$\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] = \prod_{s=1}^{t} \exp(\lambda Z_s)\, \mathbb{E}[\underbrace{\exp(\lambda Z_{t+1}) \mid \mathcal{F}_t}_{e^{\lambda^2/2}}]$$

$$= X_t \cdot e^{\lambda^2/2}$$

$$\geq X_t \quad \Longrightarrow \quad X_t \text{ is sub-martingale}$$

$$X_t = \prod_{s=1}^{t} \exp(\lambda Z_s - \lambda^2/2) \quad \Longrightarrow \quad X_t \text{ is super-martingale}$$

$X(\omega)$ $\quad \omega \in \Omega$

$$\mathbb{E}[X] = \sum_{\omega} X(\omega)\, \mathbb{P}(\omega)$$