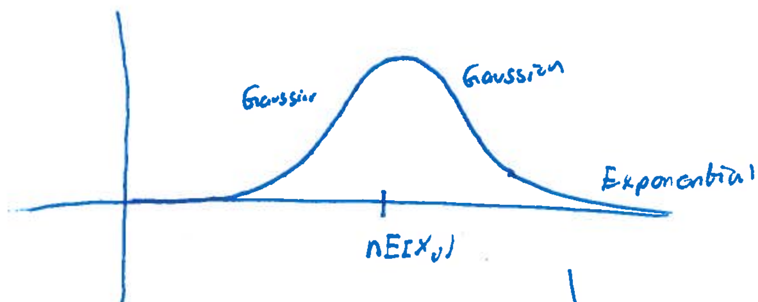
Day 10 2/18/2015

Activity:

Consider sum of n exponential r.v. X_i

Draw density function for $\sum_{i=1}^{n} X_i$. Label decay rates



tail has two parts
- Gaussian decay (central limit thm)
- Exponential tail (due to tails of each term)

Reconcile with central limit thm.

Bernstein type inequality: Let $X_1 \cdots X_N$ indep, centered, subexponential rv w/ $K = \max_i \|X_i\|_{\psi_1}$

$$P\left(\left|\sum_{i=1}^{N} X_i\right| > t\right) \le 2e^{-c \min\left(\frac{t^2}{K^2 N}, \frac{t}{K}\right)}$$

Bernstein ~~Rando~~ Inequality

Thm: Let X_i be indep, centered rv w/ $|X_i| \le K$ a.s.

Let $\sum_{i=1}^{N} \mathbb{E} X_i^2 \le \sigma^2$

$$P\left(\sum_{i=1}^{N} X_i > t \right) \le e^{-\frac{t^2/2}{\sigma^2 + \frac{1}{3}Kt}} \le$$

$$\le e^{-c \min\left(\frac{t^2}{\sigma^2}, t/K\right)}$$

Connection to Bernstein type ineq for subexponentials

X_i is subgaussian w/ $\|X_i\|_{\psi_1} \le K$.

$\sum_{i=1}^{N} \mathbb{E} X_i^2 \le N K^2$.

So this thm follows from Bernstein for subgaussians

$$P\left(\left| \sum_{i=1}^{N} X_i \right| > t \right) \le 2 e^{-c \min\left(\frac{t^2}{K^2 N}, t/K\right)}$$ ~~same as~~

Non commutative Bernstein-type inequality

Let X_i indep centered self-adjoint random $n \times n$ matrices.

Let $\|X_i\| \leq K$ almost surely. $\|\sum_{i=1}^{N} \mathbb{E} X_i^2\| \leq \sigma^2$

Then for all $t \geq 0$,

$$P(\|\sum_i X_i\| \geq t) \leq 2n e^{-\frac{t^2/2}{\sigma^2 + Kt/3}} \leq 2n e^{-c \min\left(\frac{t^2}{\sigma^2}, \frac{t}{K}\right)}$$

mixture of Gaussian & exponential tails

Singular values of tall matrices w/ heavy tailed rows

Thm: Let A be $N \times n$ w/ rows A_i indep isot rv in $\mathbb{R}^n$
Let $\|A_i\|_2 \leq \sqrt{m}$ almost surely $\forall i$.
For all $t \geq 0$

$$\sqrt{N} - t\sqrt{m} \leq \sigma_{min}(A) \leq \sigma_{max}(A) \leq \sqrt{N} + t\sqrt{m} \quad \text{w/ prob } 1 - 2n e^{-ct^2}$$

Pf: By matrix Bernstein inequality

Suffices to show $\left\| \frac{1}{N} A^t A - I_n \right\| \leq \underbrace{\max(\delta, \delta^2)}_{\text{let this be } \varepsilon}$ w $\delta = t\sqrt{\frac{m}{N}}$

Note $\frac{1}{N} A^t A - I_n = \frac{1}{N} \sum_{i=1}^{N} A_i^t A_i - I_n = \sum_{i=1}^{N} \underbrace{\frac{1}{N}(A_i^t A_i - I_n)}_{X_i} = \sum_{i=1}^{N} X_i$

To apply matrix Bernstein,

$\|X_i\| \leq \frac{1}{N}(\|A_i^t A_i\| + 1) = \frac{1}{N}(\|A_i\|_2^2 + 1) \leq \frac{1}{N}(m+1) = \frac{2m}{N} =: K.$

$\|\mathbb{E}X_i^2\|$ can be shown to be $\leq \frac{2m}{N^2} = \cancel{\sigma^2}$ (Expand X_i^2, take EXP, bound)

So $\left\| \sum_{i=1}^{N} \mathbb{E}X_i^2 \right\| \leq \frac{2m}{N} =: \sigma^2.$

So $P\left(\left\|\frac{1}{N} A^t A - I_n\right\| > \varepsilon\right) = P\left(\left\|\sum X_i\right\| > \varepsilon\right) \leq 2n e^{-c \min\left(\frac{\varepsilon^2}{\sigma^2}, \frac{\varepsilon}{K}\right)}$

$\leq 2n e^{-c\min(\varepsilon^2, \varepsilon)\frac{N}{2m}} = 2n e^{-c\delta^2 \frac{N}{2m}} = 2n e^{-ct^2/2}$

Comment on probability bound. for heavy tailed case.

$$\sqrt{N} - t\sqrt{n} \le \sigma_i \le \sqrt{N} + t\sqrt{n} \quad \text{w/ prob } 1 - 2n e^{-ct^2}$$

this n can't be removed.

Prob bound nontrivial only if $t \gtrsim \sqrt{\log n}$

So bound really says

$$\sqrt{N} - c\sqrt{n \log n} \le \sigma \le \sqrt{N} + c\sqrt{n \log n} \quad \text{whp.}$$

log necessary by Coupon collector problem

Let $A = \begin{array}{cc} & n \\ N & \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \end{array}$ each row is random $\{e_i\}_{i=1}^n \in \mathbb{R}^n$

For $\sigma_{min}(A) > 0$, no column may be 0.
Must select each of e_i at least once
Need $N \gtrsim n \log n$.

Application of sing value concentration to RIP by union bound.

Subgaussian case: A is $m \times n$ w/ $n > m$

A_T is $m \times k$ submatrix

$\sqrt{m} - C\sqrt{k} - \hat{t} \leq \sigma(A_T) \leq \sqrt{m} + C\sqrt{k} + \hat{t}$ w/ prob $1 - 2e^{-C_k \hat{t}^2}$

with $\binom{n}{k} \lesssim \left(\frac{n}{k}\right)^k$ subsets, want tail prob small

$\left(\frac{n}{k}\right)^k e^{-c\hat{t}^2}$ should be small $\Rightarrow \hat{t} \gtrsim \sqrt{k \log n/k}$

So, can achieve high prob lower bound on all $\sigma(A_T)$

$\sqrt{m} - \sqrt{k \log n/k} \lesssim \sigma_{min}$ whp.

Need $m \gtrsim k \log n/k$. Can use union bound

Heavy tailed case: A is $m \times n$ w/ $n > m$

A_T is $m \times k$

$\sqrt{m} - t\sqrt{k} \leq \sigma(A_T) \leq \sqrt{m} + t\sqrt{k}$ w prob $1 - 2ne^{-ct^2}$

High prob control over all $A_T \Rightarrow t \gtrsim \sqrt{k \log n/k}$

Need $\sqrt{m} - \sqrt{k \log n/k}\,\sqrt{k} \geq 0$

$\sqrt{m} \gg k\sqrt{\log n/k} \Rightarrow m \gg k^2 \log n/k.$

Bad.

Can't use union bound.

Theorem: Uniform control of singular values

Let A be $N \times d$ w/ $N \leq d$
with rows A_i ind. isotropic r v in $\mathbb{R}^d$
Let $|a_{ij}| \leq K$ almost surely. $\forall\ 1 < n \leq d$

$$\mathbb{E} \max_{|T| \leq n} \max_j |\sigma_j(A_T) - \sqrt{N}| \leq C_K \log n \sqrt{\log d} \sqrt{\log N}$$

Thm: Heavy tailed RIP

Let A be $m \times n$
rows A_i indep isotropic rv in $\mathbb{R}^n$
$|a_{ij}| \leq K$ a.s.
Let $\bar{A} = \frac{1}{\sqrt{m}} A$. $\forall\ 1 < k \leq n$ $\forall\ \delta \in (0,1)$

if $m \gtrsim \frac{C_K}{\delta^2} k \log n \log^2 k \log(\frac{1}{\delta^2} k \log n \log^2 k)$ then $\mathbb{E}\ \delta_k(\bar{A}) \leq \delta$.

Corollary: If $k \gtrsim \log n$, $k \gtrsim \frac{1}{\delta^2}$, Thm reduces to

if $m \gtrsim \frac{C_K}{\delta^2} k \log n \log^3 k$ then $\mathbb{E}(\delta_k(\bar{A})) \leq \delta$.

Conclude: W/ random Fourier measurements, need

$m \gtrsim C k \log n \log^3 k$ measurements