Day 8 2/9/2015

Activity:

Let $X_i \overset{iid}{\sim} \mathrm{Exp}(\lambda)$ $\qquad p(x) = \lambda e^{-\lambda x}$

What roughly is $\max_{i=1\ldots n} X_i$? $\sim \frac{\log n}{\lambda}$

Write a high probability ~~concentration~~ bound for $\max_{i \in [n]} X_i$.

$$P(X_i < t \ \forall i) = 1 - P(X_i > t \text{ for some } i)$$
$$\geq 1 - \sum_i P(X_i > t)$$
$$= 1 - n e^{-\lambda t}$$

$$P\left(\max X_i > \frac{2\log n}{\lambda}\right) \leq n e^{-\lambda \frac{2\log n}{\lambda}}$$
$$= n e^{-2\log n}$$
$$= n \frac{1}{n^2} = \frac{1}{n}.$$

So $P\left(\max X_i > \frac{2\log n}{\lambda}\right) < \frac{1}{n}$

$P\left(\max X_i > \frac{c\log n}{\lambda}\right) < \frac{1}{n^{c-1}}$

Introduction to Subgaussians:

Recall: random matrices w/ iid $N(0,1)$ entries have RIP

Used concentration bound on singular values of tall submatrices

If $A_{N \times n}$, $\sqrt{N} - C\sqrt{n} - t \leq \sigma_{min}(A) \leq \sigma_{max}(A) \leq \sqrt{N} + C\sqrt{n} + t$ (*)

w/ prob $1 - e^{-t^2/2}$

In proof of (*), we used the fact that A has iid $N(0,1)$ entries in following ways

Step 1: To show $\cancel{\|\frac{1}{N}A^tA - I\|}$

$$\max_{x \in \mathcal{N}} \left| \frac{1}{N}\|Ax\|_2^2 - 1 \right| \leq \varepsilon/2$$ w/ high prob.

$\|Ax\|_2^2 \sim \chi_N^2$

$P(|\chi_N^2 - N| \geq 2\sqrt{N}t + 2t^2) \leq 2e^{-t^2}$ – tail bound on sum of squares of Gaussian.

Other distributions have similar (or better tails)

Can get results for class of random vars with tails at least as good as Gaussians. This class is Subgaussians

Properties of Gaussians

$$X \sim N(0,1) \qquad p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$P(|X| > t) = \frac{2}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2}\,dx \le \frac{2}{\sqrt{2\pi}} e^{-t^2/2} \quad \text{for } t > 1$$

$$\left(\mathbb{E}|X|^p\right)^{1/p} = \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\right)^{1/p} = O(\sqrt{p}) \quad \text{for } p \ge 1$$

$$\mathbb{E}\, e^{tX} = e^{t^2/2} \quad \text{for } t \in \mathbb{R}$$

Subgaussians will have similar behavior as these.

Lemma. Let X be a random variable. TFAE

1) $P(|X| > t) \le e^{1 - t^2/K_1^2} \quad \forall t \ge 0$ (tail bound)
2) $(\mathbb{E}|X|^p)^{1/p} \le K_2 \sqrt{p} \quad \forall p \ge 1$ (moment bound)
3) $\mathbb{E}(X^2/K_3^2) \le e$ (Superexponential moment bound)

If $\mathbb{E}X = 0$, (1)-(3) Equiv to following:

4) $\mathbb{E}(e^{tX}) \le e^{t^2 K_4^2} \quad \forall t \in \mathbb{R}$ (moment generating function bound)

Subgaussians

A r.v. X that satisfies any of (1)-(3) is subgaussian. Subgaussian norm, $\|X\|_{\psi_2}$ is

$$\|X\|_{\psi_2} = \sup_{p \geq 1} \frac{\left(\mathbb{E}(|X|^p)\right)^{1/p}}{\sqrt{p}}$$

(best constant possible in property 2)

Examples:

If $N(0,1)$ has subgaussian norm C

Gaussians $N(0,\sigma^2)$ has subgaussian norm $C\sigma$

Bernoulli $X = \begin{cases} 1 & \text{w/ } p=1/2 \\ -1 & \text{w } p=1/2 \end{cases}$ $\|X\|_{\psi_2} = 1$

Any Bounded variable X ($X \leq M$ almost surely) $\|X\|_{\psi_2} \leq M$

Subexponential:

A r.v. is subexponential if it satisfy any of following conditions

- $P(|X| > t) \leq e^{1 - t/K_1} \quad \forall t \geq 0$
- $(\mathbb{E} |X|^p)^{1/p} \leq K_2 p \quad \forall p \geq 1$
- $\mathbb{E} e^{X/K_3} \leq e$

Subexponential norm

$$\|X\|_{\psi_1} = \sup_{p \geq 1} p^{-1} (\mathbb{E} |X|^p)^{1/p}$$

Lemma: X is subgaussian iff X^2 subexponential.

Subexponential Bernstein inequality:

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

for X_i indep, centered, subexponential rvs w/ $K = \max_i \|X_i\|_{\psi_1}$

Pf: WLOG take $K=1$

Let $S = \sum_{i=1}^{N} X_i$

$$P(S \geq t) = P(e^{\lambda S} \geq e^{\lambda t}) \leq e^{-\lambda t} \mathbb{E} e^{\lambda S} = e^{-\lambda t} \prod_{i=1}^{N} E e^{\lambda X_i}$$

$$P(S \geq t) \leq e^{-\lambda t} \prod_i e^{c_1 \lambda^2} = e^{-\lambda t + c_1 \lambda^2 N}$$

choose $\lambda = \min\left(\frac{t}{2c_1 N}, c \right)$

$$P(S \geq t) \leq e^{-\min\left(\frac{t^2}{4cN}, \frac{ct}{2} \right)}$$

Similar for $P(S \leq -t)$.

Corollary

Let $X_1 \dots X_N$ be ind cent subexp rv. $K = \max \|X_i\|_{\psi_1}$.

& $\varepsilon \geq 0$

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

Theorem:

Let A be $N \times n$ matrix w/ independent subgaussian isotropic random rows A_i then $\forall t$

$$\sqrt{N} - C\sqrt{n} - t \leq \sigma_{min}(A) \leq \sigma_{max}(A) \leq \sqrt{N} + C\sqrt{n} + t$$

with prob at least $1 - 2e^{-ct^2}$.

Note: C, c depend only on subgaussian norm $\max_i \|A_i\|_{\psi_2}$

Proof: Let $\varepsilon = \delta \vee \delta^2$ $\quad \delta = C\sqrt{\frac{n}{N}} + \frac{t}{\sqrt{N}}$

Suffices to show

$$\max_{x \in \mathcal{N}_{1/4}} \left| \frac{1}{N} \|Ax\|_2^2 - 1 \right| \leq \varepsilon/2 \quad \text{with high probability} \qquad (*)$$

$$\|Ax\|_2^2 = \sum_{i=1}^{N} \underbrace{\langle A_i, x \rangle^2}_{Z_i^2}$$

Let $Z_i := \langle A_i, x \rangle$ independent subgaussian r.v.s

$\mathbb{E} Z_i^2 = 1$ and $\|Z_i\|_{\psi_2} \leq K$

So $Z_i^2 - 1$ are indep, centered, subexponential rv

$$\|Z_i^2 - 1\|_{\psi_1} \leq 2\|Z_i^2\|_{\psi_1} \leq 4\|Z_i\|_{\psi_2}^2 \leq 4K^2$$

Also, note $K \geq \|Z_i\|_{\psi_2} \geq \frac{1}{\sqrt{2}} (\mathbb{E}|Z_i|^2)^{1/2} = \frac{1}{\sqrt{2}}$.

By subexponential deviation bound

$$P\left(\left| \frac{1}{N} \sum_{i=1}^{N} (Z_i^2 - 1) \right| \geq \varepsilon/2 \right) \leq 2e^{-c \min\left(\frac{\varepsilon^2}{\|Z_i^2 - 1\|_{\psi_1}^2}, \frac{\varepsilon}{\|Z_i^2 - 1\|_{\psi_1}} \right) N}$$

$$\leq 2e^{-c_1 \frac{\min(\varepsilon^2, \varepsilon)}{K^4} N} = 2e^{-c_1 \frac{\delta^2 N}{K^4}}$$

$$\leq 2e^{-\frac{c_1}{K^4}(C^2 n + t^2)}$$

This gets high prob control $(*)$