Sub-gaussian random vectors

Last Week

- Sub-gaussian r.vs. Hoeffding's Ineq.
- Sub-exponential r.vs. Bernstein's Ineq.

Concentration of the norm — Application of Bernstein's Ineq.

Lemma Let $X = (X_1, \dots X_n) \in \mathbb{R}^n$ be a random vector with independent, sub-gaussian coordinates X_i's that satisfy $\mathbb{E} X_i^2 = 1$. Then,

$$\| \|X\|_2 - \sqrt{n} \|_{\psi_2} \leq C K^2$$

with $K := \max_i \|X_i\|_{\psi_2}$

$\boxed{\text{Proof:}}$ Apply Bernstein's ineq. to $\frac{1}{n}\|X\|^2 - 1 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 1)$

Assume $K \geq 1$

(Ind., Mean Zero, SubExp)

$\|X_i^2 - 1\|_{\psi_1} \leq c\|X_i^2\|_{\psi_1}$ (centering)

$\leq c\|X_i\|_{\psi_2}^2$ (square of SubG is SubExp)

$\leq C K^2$

Bernstein's Ineq. gives us

$$\mathbb{P}_r\left\{ \left| \frac{1}{n}\|X\|^2 - 1 \right| > u \right\} \leq 2 \exp\left[-cn \min\left\{ \frac{u^2}{K^4}, \frac{u}{K^2} \right\} \right] \quad (K^4 \geq K^2 \geq 1)$$

$$\leq 2 \exp\left[\frac{-cn}{K^4} \min\{u^2, u\} \right]$$

- Numeric Ineq. For all $z \geq 0$ $\quad |z-1| \geq \delta$ implies $|z^2 - 1| \geq \max\{\delta, \delta^2\}$

(Proof via case analysis $z \geq 1+\delta$ $\quad z < 1-\delta$)

$$\mathbb{P}_r\left\{ \left| \frac{1}{\sqrt{n}}\|X\| - 1 \right| \geq \delta \right\} \leq \mathbb{P}_r\left\{ \left| \frac{1}{n}\|X\|^2 - 1 \right| \geq \max\{\delta, \delta^2\} \right\}$$

Set $u = \max\{\delta, \delta^2\}$

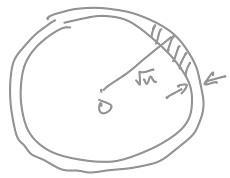
$$\leq 2 \exp\left(\frac{-cn}{K^4} \cdot \delta^2 \right)$$

Set $t = \delta\sqrt{n}$

$$\mathbb{P}_r\left\{ \left| \|X\| - \sqrt{n} \right| \geq t \right\} \leq 2 \exp\left(\frac{-ct^2}{K^4} \right) \quad \text{for all } t \geq 0$$

$\square$

Sharp Concentration: With high prob, $\|X\|$ is within $\sqrt{n} \pm$ Constant



Sub-gaussian random vectors

Defn: A random vector $X \in \mathbb{R}^n$ is called sub-gaussian iff the one-dimensional marginals $\langle X, x \rangle$ are sub-gaussian r.v.s for all $x \in \mathbb{R}^n$.

The sub-gaussian norm of X is defined as

$$\|X\|_{\psi_2} := \sup_{x \in S^{n-1}} \|\langle X, x \rangle\|_{\psi_2}$$

(i.e., $\|x\| = 1$)

Examples

① Vector X with independent, sub-gaussian coordinates is subgaussian. $\|X\|_{\psi_2} \le \max_i \|X_i\|_{\psi_2}$

Pf: fix $x \in S^{n-1}$

$$\Big\| \sum_i x_i X_i \Big\|_{\psi_2}^2 \le \sum_{i=1}^n x_i^2 \|X_i\|_{\psi_2}^2$$

(Hoeffding's ineq.)

$$\le C \cdot \max_i \|X_i\|_{\psi_2}^2$$

② Multivariate Normal

$$g \sim N(\underbrace{0}_{\text{Mean}}, \underbrace{I_n}_{\text{Covariance.}})$$

Density function

$$f(x) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_i^2}{2}\right)$$

$$= \frac{1}{(2\pi)^{n/2}} \cdot e^{-\|x\|^2/2}$$

$N(\mu, \Sigma)$

$\mu \in \mathbb{R}^n$

$\Sigma \in \mathbb{R}^{n \times n}$

Recall

Covariance Matrix $\Sigma := \mathbb{E}\, X X^T$

(PSD)

③ $X \sim \mathrm{Unif}(\sqrt{n}\, S^{n-1})$

$X \sim \mathrm{Unif}(\sqrt{n}\, B_2^n)$

Defn: A random vector $X \in \mathbb{R}^n$ is said to be isotropic iff

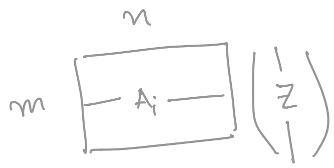
$$\mathbb{E}\, X X^T = I_n$$

Corollary of Norm Concentration

THM: Let A be an $m \times n$ random matrix, with independent, sub-gaussian, and isotropic rows. Then, for all $z \in S^{n-1}$

$$\Big\| \|Az\|_2 - \sqrt{m} \Big\|_{\psi_2} \le CK^2$$

$$K := \max_i \|A_i\|_{\psi_2}$$



Pf:

$$Az = \begin{pmatrix} A_1^T z \\ \vdots \\ A_i^T z \\ \vdots \\ A_m^T z \end{pmatrix} = X$$

Components of X are independent and satisfy $\mathbb{E}\, X_i^2 = 1$

$$\mathbb{E}\, z^T A_i \cdot A_i^T z = z^T (\mathbb{E} A_i A_i^T) z \overset{\text{(isotropy)}}{=} z^T \cdot I_n z = 1. \quad (z \in S^{n-1})$$

Since A_i is sub-gaussian with $\|A_i\|_{\psi_2} \le K$

Therefore,

$$\|A_i^T z\|_{\psi_2} \le K.$$

Applying norm concentration we get

$$\Big\| \|Az\| - \sqrt{m} \Big\|_{\psi_2} \le CK^2$$

$\square$

Johnson - Lindenstrauss Lemma

Thm: Let X be a set of N points in $\mathbb{R}^n$ and $0<\varepsilon<1$. Consider an $m \times n$ matrix A whose rows are independent, mean-zero, isotropic, and sub-gaussian rand. vectors in $\mathbb{R}^n$.

Rescale A by defining (random projection) $P := \frac{1}{\sqrt{m}} \cdot A$

Assume that $m \geq C\varepsilon^{-2} \log N$, where C is a large enough constant (which depends on the sub g norm of A_i-s).

Then, whp, the map P preserves all pairwise distances in X, up to ε relative error,

$$(1-\varepsilon)\,\|x-y\| \leq \|Px - Py\|_2 \leq (1+\varepsilon)\,\|x-y\| \qquad \text{for all } x,y \in X.$$

$\boxed{\text{Proof:}}$ Write $T := \left\{ \frac{x-y}{\|x-y\|} : \text{distinct } x,y \in X \right\} \subseteq S^{n-1}$

$$|T| = \Theta(N^2)$$

For each $z \in T$, we have

$$\big\| \|Az\| - \sqrt{m} \big\|_{\psi_2} \leq C K^2$$

sub-g r.v., for each $z \in T$

$$\mathbb{E} \max_{z \in T} \big| \|Az\| - \sqrt{m} \big| \leq C K^2 \sqrt{\log |T|}$$

$$\leq \varepsilon \sqrt{m} \qquad (\text{Since } m \geq \varepsilon^{-2} K^4 \log |T|$$

Scaling by $\sqrt{m}$

$$\mathbb{E} \max_{z \in T} \big(\|Pz\| - \|z\| \big) \leq \varepsilon$$

Via Markov's Ineq., with constant prob.

$$(1-\varepsilon)\,\|z\| \leq \|Pz\| \leq (1+\varepsilon)\,\|z\|$$

$\square$