Solutions to Selected Exercises in *High-Dimensional Probability*

Yuze Han Liangyu Zhang

October 2020

1 Appetizer and preliminaries on random variable

Exercise 0.0.5.

$$\left(\frac{n}{m}\right)^m \leq \prod_{i=0}^{m-1} \frac{n-i}{m-i} = \binom{n}{m} \leq \sum_{k=0}^{m} \binom{n}{k} \leq \left(\frac{n}{m}\right)^m \sum_{k=0}^{m} \binom{n}{k} \left(\frac{m}{n}\right)^k \leq \left(\frac{n}{m}\right)^m \left(1+\frac{m}{n}\right)^n \leq \left(\frac{en}{m}\right)^m.$$

Exercise 0.0.6. Let $k := \lceil 1/\epsilon^2 \rceil$. The number of ways to choose k elements from an N-element set with repetitions is $\binom{N+k-1}{k}$. Exercise 2.2.5 implies

$$\binom{N+k-1}{k} \leq \left(\frac{e(N+k-1)}{k}\right)^k \leq \left(e + eN\epsilon^2\right)^{\lceil 1/\epsilon^2 \rceil}.$$

Exercise 1.3.3. Let $\sigma^2 = \operatorname{Var}(X_1)$.

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right| \leq \sqrt{\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right)^2} \leq \sqrt{\frac{\sigma^2}{N}}.$$

Note that for a sequence of r.v.s $\{X_n\}$, $X_n \xrightarrow{d} X$ does not necessarily imply $\mathbb{E}\, X_n \to \mathbb{E}\, X$. And a sufficient condition for $\mathbb{E}\, X_n \to \mathbb{E}\, X$ is uniform integrability.

2 Concentration of sums of independent random variables

Exercise 2.1.4.

$$\begin{aligned}\mathbb{E}\, g^2 \mathbf{1}_{\{g>t\}} &= \int_t^\infty \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_t^\infty \frac{x}{\sqrt{2\pi}} d\left(-e^{-x^2/2}\right) = t \cdot \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \mathbb{P}\ \{g > t\} \\ &\le \left(t + \frac{1}{t}\right) \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \quad \text{(by Proposition 2.1.2)}\end{aligned}$$

Exercise 2.2.7. We first prove the following lemma.

Lemma. *If $\mathbb{E}\, X = 0$ and $a \le X \le b$, we have $\mathbb{E}\, e^{\lambda X} \le e^{\lambda^2 (b-a)^2/8}$.*

Proof. The convexity of e^x implies

$$\mathbb{E}\, e^{\lambda X} \le \mathbb{E}\, \left(\frac{b - X}{b - a} e^{\lambda a} + \frac{x - a}{b - a} e^{\lambda b}\right) = e^{\phi(\lambda)},$$

where $\phi(x) = \ln\left[\left(\frac{b}{b-a} e^{ax} + \frac{-a}{b-a} e^{bx}\right)\right]$. One can prove that $\phi(0) = 0$, $\phi'(0) = 0$ and $\phi''(x) \le \frac{(b-a)^2}{4}$. Taylor expansion gives the desired result. An alternative proof can be found in [1, Lemma 2.2] □

For an arbitrary $\lambda > 0$,

$$\begin{aligned}\mathbb{P}\ \left\{\sum_{i=1}^N (X_i - \mathbb{E}\, X_i) \ge t\right\} &\le \exp(-\lambda t) \prod_{i=1}^N \mathbb{E}\ \exp(\lambda(X_i - \mathbb{E}\, X_i)) \\ &\le \exp(-\lambda t) \prod_{i=1}^N \exp(\lambda^2 (M_i - m_i)^2/8) \\ &= \exp\left(\frac{1}{8} \sum_{i=1}^N (M_i - m_i)^2 \lambda^2 - t\lambda\right)\end{aligned}$$

The minimum is attained for $\lambda = \frac{4t}{\sum_{i=1}^N (M_i - m_i)^2}$. This complete the proof of Hoeffding's inequality.

Exercise 2.2.8. Let $X_i = \mathbf{1}_{\{\text{The algorithm returns the wrong answer at the i-th time}\}}$. Then $X_i \sim \mathit{Bernoulli}\left(1, \frac{1}{2} - \delta\right)$. As long as $N \ge \frac{1}{2\delta^2} \ln(1/\epsilon)$, Hoeffding's inequality implies

$$\mathbb{P}\ \left\{\sum_{i=1}^N X_i > \frac{1}{2}\right\} = \mathbb{P}\ \left\{\sum_{i=1}^N (X_i - \mathbb{E}\, X_i) > \delta N\right\} \le \exp\left(-2N\delta^2\right) \le \epsilon.$$

Exercise 2.2.9.

(a) Let $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$. For $N = 4\sigma^2/\epsilon^2$,

$$\mathbb{P}\ \{|\hat{\mu} - \mu| \geq \epsilon\} \leq \frac{\mathrm{Var}\,(\hat{\mu})}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2} = \frac{1}{4}$$

(b) We repeat part (a) K times and obtain K independent sapmle means $\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_K$ with $K = 8\ln(1/\delta)$. Denote their median by $\tilde{\mu}$. Let $Y_i = \mathbf{1}_{|\hat{\mu}_1 - \mu| \geq \epsilon}$. Then $Y_i \sim Bernoulli(1, p_i)$ with $p_i \leq 1/4$. Hoeffding's inequality implies

$$\mathbb{P}\ \{|\tilde{\mu} - \mu| \geq \epsilon\} \leq \mathbb{P}\ \left\{\sum_{i=1}^{K} Y_i \geq \frac{K}{2}\right\} \leq \mathbb{P}\ \left\{\sum_{i=1}^{K} (Y_i - \mathbb{E}\,Y_i) \geq \frac{K}{4}\right\} \leq \exp\left(-\frac{K}{8}\right) \leq \delta.$$

Exercise 2.3.2. For an arbitrary $\lambda > 0$,

$$\mathbb{P}\ \{S_N \leq t\} = \mathbb{P}\ \{-\lambda S_N \geq -\lambda t\} \leq e^{\lambda t} \prod_{i=1}^{N} \mathbb{E}\,\exp(-\lambda X_i) \leq \exp\left(\lambda t + (e^{-\lambda} - 1)\mu\right).$$

Substituting $\lambda = \ln(\mu/t)$ gives the desired result.

Exercise 2.3.5. By Theorem 2.3.1 and Exercise 2.3.2, we have

$$\begin{aligned}\mathbb{P}\ \{|S_N - \mu| \geq \delta\mu\} &\leq e^{-\mu}\left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} + e^{-\mu}\left(\frac{e}{1-\delta}\right)^{(1-\delta)\mu} \\ &= \exp\left[\delta\mu - (1+\delta)\mu\ln(1+\delta)\right] + \exp\left[-\delta\mu - (1-\delta)\mu\ln(1-\delta)\right].\end{aligned}$$

It suffices to prove

$$\delta - (1+\delta)\ln(1+\delta) \leq -c\delta^2, \quad \delta \in (-1, 1)$$

for some $c > 0$. We can choose $c \leq 2\ln 2 - 1$.

Exercise 2.3.8. The ch.f. of $\frac{X-\lambda}{\lambda}$ converges to $\exp(-t^2/2)$ as $\lambda \to \infty$. Note that $\forall \theta > 0, \lim_{n\to\infty} f(n\theta) = a$ does not imply $\lim_{x\to\infty} f(x) = a$. See [3, Example 3.4.8] for an alternative proof.

Exercise 2.4.2. Suppose $d \leq C \log n$. By Theorem 2.3.1

$$\mathbb{P}\left\{\exists i \leq n : d_i \geq KC \log n\right\} \leq \sum_{i=1}^{n} \mathbb{P}\left\{d_i \geq KC \log n\right\} \leq ne^{-d}\left(\frac{ed}{KC \log n}\right)^{KC \log n}$$
$$\leq ne^{KC \log n} K^{-KC \log n} \leq e^{(KC+1-KC \log K)\log n} \leq 0.1$$

for a sufficient large constant K.

Exercise 2.4.3.

$$\mathbb{P}\left\{\exists i \leq n : d_i \geq K\frac{\log n}{\log\log n}\right\} \leq ne^{-d}\left(\frac{ed \log\log n}{K \log n}\right)^{K \log n/\log\log n}$$
$$\leq e^{-(K-1)\log n + o(\log n)} \leq 0.1$$

for a sufficient large constant K.

Exercise 2.4.4. Take $V' \subset V$ randomly of size $\tilde{n} = n^{1/3}$ with vertex indices $I = \{i_1, i_2, \ldots, i_{\tilde{n}}\}$. Let $\tilde{d}_j$ denote the degree of the vertex $j \in I$ of V' and $\tilde{d}$ denote the expected degree. Then $\tilde{d} = o\left(\frac{\log n}{n^{2/3}}\right)$. It follows that $\log \tilde{d}/\log n \leq -2/3 + o(1)$. We have

$$\mathbb{P}\left\{\exists i \leq \tilde{n} : \tilde{d}_i \geq 1\right\} \leq \tilde{n}e^{-\tilde{d}}e\tilde{d} \leq e^{\log \tilde{n} + \log \tilde{d} + 1} \leq 1 - \sqrt{0.9}$$

when n is sufficiently large. Denote $A = \{\text{There are no edges between vertices in } V'\}$. Then $\mathbb{P}\{A\} \geq \sqrt{0.9}$. Condition on A, $\{d_j, j \in I\}$ are independent. By Poisson approximation (for the accuracy of Poisson approximation, see [5]), we obtain

$$\mathbb{P}\left\{d_i = 10d\right\} = \frac{1}{\sqrt{2\pi 10d}}e^{-d}\left(\frac{ed}{10d}\right)^{10d}(1+o(1)) = \frac{1}{\sqrt{20\pi}}e^{-(10\log 10 + 1/2 - 9)d}(1+o(1)) \geq an^{-c}$$

for some constants $a, c > 0$ and sufficiently large n. Since $d = o(\log n)$, we may assume $c < 1/3$. It follows that

$$\begin{aligned}
\mathbb{P}\left\{\exists i : d_i = 10d\right\} &\geq \mathbb{P}\left\{\exists j \in I : d_j = 10d\right\} \\
&\geq \mathbb{P}\left\{\exists j \in I : d_j = 10d \,|A\right\} \mathbb{P}\left\{A\right\} \\
&= \left[1 - (1 - \mathbb{P}\left\{d_i = 10d\right\})^{\tilde{n}}\right] \mathbb{P}\left\{A\right\} \\
&\geq \left[1 - \exp\left(-an^{1/3-c}\right)\right] \mathbb{P}\left\{A\right\} \\
&\geq 0.9
\end{aligned}$$

when n is sufficiently large.

Exercise 2.4.5. Take $V' \subset V$ randomly of size $\tilde{n} = n^{1/3}$ with vertex indices $I = \{i_1, i_2, \ldots, i_{\tilde{n}}\}$. Let $\tilde{d}_j$ denote the degree of the vertex $j \in I$ of V' and $\tilde{d}$ denote the expected degree. Then $\tilde{d} = O\left(n^{-2/3}\right)$. We have

$$\mathbb{P}\left\{\exists i \leq \tilde{n} : \tilde{d}_i \geq 1\right\} \leq \tilde{n}e^{-\tilde{d}}e\tilde{d} \leq e^{\log \tilde{n} + \log \tilde{d} + 1} \leq 1 - \sqrt{0.9}$$

when n is sufficiently large. Denote $A = \{\text{There are no edges between vertices in } V'\}$. Then $\mathbb{P}\left\{A\right\} \geq \sqrt{0.9}$. Condition on A, $\{d_j, j \in I\}$ are independent. Let $k = \frac{b \log n}{\log \log n}$ with $b < 1/4$. Assume $1 \leq d \leq C$. Then we have

$$\begin{aligned}
\mathbb{P}\left\{d_i = k\right\} &= \binom{n-1}{k}\left(\frac{d}{n-1}\right)^k \left(1 - \frac{d}{n-1}\right)^{n-k-1} \\
&\geq \frac{(n-k)^k}{k!}\left(\frac{1}{n-1}\right)^k \left(1 - \frac{C}{n-1}\right)^{n-k-1} \\
&= \frac{1}{k!}\left(1 - \frac{k-1}{n-1}\right)^k e^{-C}(1 + o(1)) \\
&= \frac{e^k}{k^k\sqrt{2\pi k}} e^{-C}(1 + o(1)) \\
&= \exp\left(k - k\log k - \frac{1}{2}\log(2\pi k) - C\right)(1 + o(1)) \\
&= e^{-b\log n(1+o(1))} \\
&\geq n^{-1.1b}
\end{aligned}$$

when n is sufficiently large. Note that we can also use Poisson approximation to obtain the same inequality. It follows that

$$\begin{aligned}
\mathbb{P}\{\exists i : d_i = k\} &\geq \mathbb{P}\{\exists j \in I : d_j = k\} \\
&\geq \mathbb{P}\{\exists j \in I : d_j = k \,|A\}\,\mathbb{P}\{A\} \\
&= \left[1 - (1 - \mathbb{P}\{d_i = k\})^{\tilde{n}}\right]\mathbb{P}\{A\} \\
&\geq \left[1 - \exp\left(-n^{1/3-1.1b}\right)\right]\mathbb{P}\{A\} \\
&\geq 0.9
\end{aligned}$$

when n is sufficiently large.

Exercise 2.5.1.

$$\begin{aligned}
\mathbb{E}\,|X|^p &= 2\int_0^\infty x^p \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
&= \frac{2^{p/2}}{\sqrt{\pi}} \int_0^\infty t^{\frac{p-1}{2}} e^{-t} dt \quad \text{(by change of variables } t = x^2\text{)} \\
&= 2^{p/2} \frac{\Gamma((1+p)/2)}{\Gamma(1/2)}.
\end{aligned}$$

(2.11) follows from Stirling's formula $\Gamma(x+1) \sim \sqrt{2\pi x}\left(\frac{x}{e}\right)^x$.

Exercise 2.5.4. Assume $K_4 = 1$. By Jensen's inequality,

$$\lambda \mathbb{E}\, X \leq \ln \mathbb{E}\, e^{\lambda X} \leq \lambda^2.$$

Since λ is arbitrary, $\mathbb{E}\, X = 0$.

Exercise 2.5.5.

(a) For $\lambda^2 \geq 1/2$, $\mathbb{E}\,\exp(\lambda^2 X^2) = \infty$.

(b) For $M > \sqrt{K}$, we have

$$\mathbb{P}\ (X > M) \leq \exp(-\lambda^2 M^2)\,\mathbb{E}\,\exp(\lambda^2 X^2) \leq \exp((K - M^2)\lambda^2).$$

Letting $\lambda \to \infty$, $\mathbb{P}\ (X > M) = 0$.

Exercise 2.5.7. We only verify positive definiteness and the triangular inequality.

positive definiteness If $\|X\|_{\psi_2} = 0$, $\mathbb{E}\,\exp\left(X^2/t^2\right) \leq 2$ for all $t > 0$. Then for all $c > 0$, $\mathbb{P}\,(X > c) \leq e^{-c^2/t^2}\mathbb{E}\,e^{X^2/t^2} \leq 2e^{-c^2/t^2} \to 0$ as $t \to 0$. Thus $X = 0$.

triangular inequality Let $a = \|X\|_{\psi_2}$, $b = \|Y\|_{\psi_2}$. Since e^{x^2} is convex, we have

$$\mathbb{E}\,\exp\left(\frac{X+Y}{a+b}\right)^2 \leq \frac{a}{a+b}\mathbb{E}\,\exp\left(\frac{X}{a}\right)^2 + \frac{b}{a+b}\mathbb{E}\,\exp\left(\frac{Y}{b}\right)^2 \leq 2.$$

Thus $\|X + Y\|_{\psi_2} \leq \|X\|_{\psi_2} + \|Y\|_{\psi_2}$.

Exercise 2.5.9.

Poisson If $X \sim Pois(\lambda)$, we have $\mathbb{P}\,(X \geq t) \geq \frac{\lambda^{\lceil t \rceil} e^{-\lambda}}{\lceil t \rceil !} = \Omega(e^{-t\ln t})$.

Exponential If $X \sim Exp(\lambda)$, we have $\mathbb{P}\,(X \geq t) = e^{-\lambda t}$.

Pareto If $X \sim Pa(a, \theta)$, we have $\mathbb{P}\,(X \geq t) = (a/t)^\theta = \Omega(e^{-\theta \ln t})$.

Cauchy The expectation of Cauchy distribution dose not exist.

Exercise 2.5.10. From the proof of Proposition 2.5.2, we can assume $c = 1$. Let $Y_i = \frac{|X_i|}{\sqrt{1+\log i}}$. For $t \geq \sqrt{2}K$,

$$\begin{aligned}
\mathbb{P}\left(\max_i Y_i > t\right) &\leq \sum_i \mathbb{P}\left(|X_i| > t\sqrt{1+\log i}\right) \\
&\leq \sum_i 2\exp\left(-\frac{t^2(1+\log i)}{K^2}\right) \\
&= \sum_i 2\exp\left(-\frac{t^2}{K^2}\right) i^{-t^2/K^2} \\
&\leq C_1 \exp\left(-\frac{t^2}{K^2}\right)
\end{aligned}$$

Then we have

$$\begin{aligned}
\mathbb{E}\max_i Y_i &= \int_0^\infty \mathbb{P}\left(\max_i Y_i > t\right) dt \\
&= \sqrt{2}K + \int_{\sqrt{2}K}^\infty \mathbb{P}\left(\max_i Y_i > t\right) dt \\
&\leq \sqrt{2}K + C_1 \int_0^\infty \exp\left(-\frac{t^2}{K^2}\right) \\
&\leq C_2 K,
\end{aligned}$$

where $C_1, C_2 > 0$ do not depend on any parameter. Finally, for every $N \geq 2$,

$$\mathbb{E}\max_{i \leq N} |X_i| \leq \sqrt{1+\log N}\,\mathbb{E}\max_i Y_i \leq C_3 K\sqrt{\log N}.$$

Exercise 2.5.11 This proof is inspired by [6]. For a tight bound, see [4].

For $N \geq 3$, let $A_N = \{\max_{i \leq N} X_i \geq C_N \sqrt{\log N}\}$ with $C_N = \sqrt{2 - \frac{\log\log N}{\log N}} > 1$. First, we have

$$\begin{aligned}
\mathbb{E} \max_{i \leq N} X_i &= \mathbb{E}\left[\max_{i \leq N} X_i \middle| A_N\right] \mathbb{P}(A_N) + \mathbb{E}\left[\max_{i \leq N} X_i \middle| A_N^c\right] \mathbb{P}(A_N^c) \\
&\geq C_N \sqrt{\log N}\, \mathbb{P}(A_N) + \mathbb{E}\left[X_1 | A_N^c\right] \mathbb{P}(A_N^c) \\
&= C_N \sqrt{\log N}\, \mathbb{P}(A_N) + \mathbb{E}\left[X_1 | X_1 < C_N \sqrt{\log N}\right] (1 - \mathbb{P}(A_N)) \\
&\geq C_N \sqrt{\log N}\, \mathbb{P}(A_N) + \mathbb{E}\left[X_1 | X_1 < 0\right] (1 - \mathbb{P}(A_N)) \\
&\geq \left(\sqrt{\log N} - \sqrt{\frac{2}{\pi}}\right) \mathbb{P}(A_N) - \sqrt{\frac{2}{\pi}}.
\end{aligned}$$

Now we bound $\mathbb{P}(A_N)$.

$$\begin{aligned}
\mathbb{P}(A_N) &= 1 - \left[1 - \mathbb{P}(X_1 \geq C_N \sqrt{\log N})\right]^N \\
&\geq 1 - \left[1 - \frac{1}{C_N \sqrt{2\pi \log N}} \left(1 - \frac{1}{C_N^2 \log N}\right) N^{-C_N^2/2}\right]^N \\
&\geq 1 - \exp\left[-\frac{1}{C_N \sqrt{2\pi \log N}} \left(1 - \frac{1}{C_N^2 \log N}\right) N^{1 - C_N^2/2}\right] \\
&= 1 - \exp\left[-\frac{1}{C_N \sqrt{2\pi}} \left(1 - \frac{1}{C_N^2 \log N}\right)\right] \\
&= 1 - \exp\left[\frac{1}{C_N \sqrt{2\pi}} \left(\frac{1}{2\log N - \log\log N} - 1\right)\right] \\
&\geq 1 - \exp\left[\frac{1}{\sqrt{2\pi}} \left(\frac{1}{\log N} - 1\right)\right].
\end{aligned}$$

For $N \geq 8$, $\mathbb{P}(A_N) \geq 1 - \exp(-1/2\sqrt{2\pi}) \triangleq C$. When $N \geq \exp\left(\frac{8(C+1)^2}{\pi C^2}\right)$, we obtain

$$\mathbb{E} \max_{i \leq N} X_i \geq C\sqrt{\log N} - \sqrt{\frac{2}{\pi}}(C+1) \geq \frac{C}{2}\sqrt{\log N}.$$

Since $\mathbb{E}\, X_1 = 0$ and $\mathbb{E} \max_{i \leq N} X_i > 0$ for $N \geq 2$, there exists $c > 0$ such that $\mathbb{E} \max_{i \leq N} X_i \geq c\sqrt{\log N}$ for all $N \geq 1$.

Exercise 2.6.5. The lower bound follows from

$$\left\| \sum_{i=1}^N a_i X_i \right\|_{L^p} \geq \left\| \sum_{i=1}^N a_i X_i \right\|_{L^2} = \left(\sum_{i=1}^N a_i^2 \right)^{1/2}.$$

The upper bound follows from

$$\left\| \sum_{i=1}^N a_i X_i \right\|_{L^p} \leq C_1\sqrt{p} \left\| \sum_{i=1}^N a_i X_i \right\|_{\psi_2} \leq C_2\sqrt{p}\sqrt{\sum_{i=1}^N \|a_i X_i\|_{\psi_2}} \leq CK\sqrt{p} \left(\sum_{i=1}^N a_i^2 \right)^{1/2}.$$

Exercise 2.6.7. The upper bound follows from

$$\left\| \sum_{i=1}^N a_i X_i \right\|_{L^p} \leq \left\| \sum_{i=1}^N a_i X_i \right\|_{L^2} = \left(\sum_{i=1}^N a_i^2 \right)^{1/2}.$$

To prove the lower bound, we first note that.

$$\|Z\|_{L^2}^2 = \mathbb{E}\,|Z|^{p/2}|Z|^{2-p/2} \leq \||Z|^{p/2}\|_{L^2}\||Z|^{2-p/2}\|_{L^2} = \|Z\|_{L^p}^{p/2}\|Z\|_{L^{4-p}}^{2-p/2}.$$

It follows that

$$\left\| \sum_{i=1}^N a_i X_i \right\|_{L^p} \geq \frac{\left\| \sum_{i=1}^N a_i X_i \right\|_{L^2}^{4/p}}{\left\| \sum_{i=1}^N a_i X_i \right\|_{L^{4-p}}^{4/p-1}} \geq \frac{\left(\sum_{i=1}^N a_i^2 \right)^{2/p}}{\left[CK\sqrt{4-p} \left(\sum_{i=1}^N a_i^2 \right)^{1/2} \right]^{4/p-1}} = c(K) \left(\sum_{i=1}^N a_i^2 \right)^{1/2}$$

with $c(K) = \left(CK\sqrt{4-p}\right)^{1-4/p}$.

Exercise 2.6.9. Let X be $\sqrt{\log 2}$ with probability $1/4$ and $-\sqrt{\log 2}$ with probability $3/4$. We have $\|X\|_{\psi_2} = 1$ but $\mathbb{E}\exp(X - \mathbb{E}\,X)^2 > 2$.

Exercise 2.7.2.

a $\Rightarrow$ b. Assume $K_1 = 1$. We have

$$\mathbb{E}\,|X|^p = \int_0^\infty pt^{p-1}\mathbb{P}\left(|X| \geq t\right)dt = 2p\int_0^\infty t^{p-1}e^{-t}dt = 2p\Gamma(p) \leq 2p^p.$$

b ⇒ c. Assume $K_2 = 1$. For $0 < \lambda < 1/2e$,

$$\mathbb{E}\,\exp(\lambda|X|) = \sum_{p=0}^{\infty} \frac{\lambda^p \mathbb{E}\,|X|^p}{p!} \le \sum_{p=0}^{\infty} \frac{\lambda^p \mathbb{E}\,|X|^p e^p}{p^p} \le \sum_{p=0}^{\infty} \lambda^p e^p = \frac{1}{1-\lambda e} \le \exp(2e\lambda).$$

c ⇒ d is trivial.

d ⇒ a. Assume $K_4 = 1$. $\mathbb{P}\left(|X| \ge t\right) \le e^{-t}\mathbb{E}\, e^{|X|} \le 2e^{-t}$.

Exercise 2.7.3. We first state the proposition.

Proposition. *Let X be a random variable. Then the following properties are equivalent; the parameters $K_i > 0$ appearing in these properties differ from each other by at most an absolute constant factor.*

(a) The tails of X satisfy

$$\mathbb{P}\left\{|X| \ge t\right\} \le 2\exp(-t^\alpha/K_1^\alpha) \quad \textit{for all } t \ge 0.$$

(b) The moments of X satisfy

$$\|X\|_{L^p} = (\mathbb{E}\,|X|^p)^{1/p} \le K_2 \left(\frac{p}{\alpha}\right)^{1/\alpha} \quad \textit{for all } p \ge 1.$$

(c) The MGF of $|X|^\alpha$ satisfies

$$\mathbb{E}\,\exp(\lambda^\alpha|X|^\alpha) \le \exp(K_3^\alpha\lambda^\alpha) \quad \textit{for all } \lambda \textit{ such that } 0 \le \lambda \le \frac{1}{K_3}.$$

(d) The MGF of $|X|^\alpha$ is bounded at some point, namely

$$\mathbb{E}\,\exp(|X|^\alpha/K_4^\alpha) \le 2.$$

Moreover, if $\mathbb{E}\,X = 0$ and $\alpha \ge 2$, then properties a-d are also equivalent to the following one.

(e) The MGF of X satisfies

$$\mathbb{E}\,\exp(\lambda X) \le \exp\left(K_5^{\alpha/(\alpha-1)}|\lambda|^{\alpha/(\alpha-1)}\right) \quad \textit{for all } \lambda \in \mathbb{R}.$$

If $1 < \alpha < 2$, then properties a-d are also equivalent to the following one.

(f) The MGF of X satisfies

$$\mathbb{E}\exp(\lambda X) \le \exp\left(K_6^{\alpha/(\alpha-1)}|\lambda|^{\alpha/(\alpha-1)}\right) \quad \textit{for all } |\lambda| \ge 1/2e.$$

Proof.

a $\Rightarrow$ b. Assume $K_1 = 1$. Note that p is not necessarily an integer. For all $p \ge 1$,

$$\mathbb{E}|X|^p = \int_0^\infty pt^{p-1}\mathbb{P}(|X| \ge t)dt = 2p\int_0^\infty t^{p-1}e^{-t^\alpha}dt = \frac{2p}{\alpha}\int_0^\infty u^{p/\alpha-1}e^{-u}du = 2\Gamma\left(\frac{p}{\alpha}+1\right).$$

There exists $K_2 > 0$ such that $\mathbb{E}|X|^p = 2\Gamma\left(\frac{p}{\alpha}+1\right) \le K_2\left(\frac{p}{\alpha}\right)^{p/\alpha}$.

b $\Rightarrow$ c. Assume $K_2 = 1$. For $0 < \lambda < 1/2e^{1/\alpha}$,

$$\mathbb{E}\exp(\lambda^\alpha|X|^\alpha) = \sum_{p=0}^\infty \frac{\lambda^{\alpha p}\mathbb{E}|X|^{\alpha p}}{p!} \le \sum_{p=0}^\infty \frac{\lambda^{\alpha p}p^p e^p}{p^p} \le \sum_{p=0}^\infty \lambda^{\alpha p}e^p = \frac{1}{1-2e\lambda^\alpha} \le \exp(2e\lambda^\alpha).$$

c $\Rightarrow$ d is trivial.

d $\Rightarrow$ a. Assume $K_4 = 1$. $\mathbb{P}(|X| \ge t) \le e^{-t^\alpha}\mathbb{E}\,e^{|X|^\alpha} \le 2e^{-t^\alpha}$.

b c $\Rightarrow$ e. Assume $K_2, K_3 \le 1$. Let $\beta = \alpha/(\alpha-1)$. We have $1 < \beta < \alpha$. We have the following inequality

$$\exp(x) \le x + \exp\left(|x|^\beta\right) \quad \text{for all } x \in [-1,1].$$

It follows that

$$
\begin{aligned}
\mathbb{E}\,\exp(\lambda X) &\le \mathbb{E}\,\exp\left(|\lambda|^{\beta}|x|^{\beta}\right) \\
&= \sum_{p=0}^{\infty} \frac{|\lambda|^{\beta p}\mathbb{E}\,|X|^{\beta p}}{p!} \\
&\le \sum_{p=0}^{\infty} |\lambda|^{\beta p}\left(\frac{\beta p}{\alpha}\right)^{\beta p/\alpha}\frac{e^{p}}{p^{p}} \\
&= \sum_{p=0}^{\infty} |\lambda|^{\beta p} p^{p(\beta/\alpha-1)}(\alpha-1)^{-p/(\alpha-1)}e^{p} \\
&\le \sum_{p=0}^{\infty} |\lambda|^{\beta p}(\alpha-1)^{-p/(\alpha-1)}e^{p} \\
&= \frac{1}{1-|\lambda|^{\beta}(\alpha-1)^{-1/(\alpha-1)}e} \\
&\le \exp\left(2(\alpha-1)^{-1/(\alpha-1)}e|\lambda|^{\beta}\right) \\
&\le \exp\left(2e|\lambda|^{\beta}\right) \\
&\le \exp\left((2e)^{\beta}|\lambda|^{\beta}\right)
\end{aligned}
$$

for all λ satisfying $|\lambda|^{\beta} \le (\alpha-1)^{1/(\alpha-1)}/2e$. For λ with large absolute value, by Young's inequality,

$$
\lambda x \le |\lambda||x| \le \frac{|\lambda|^{\beta}}{\beta} + \frac{|x|^{\alpha}}{\alpha}.
$$

It follows that

$$
\mathbb{E}\,e^{\lambda X} \le e^{|\lambda|^{\beta}/\beta}\mathbb{E}\,e^{|X|^{\alpha}/\alpha} \le e^{|\lambda|^{\beta}/\beta}e^{1/\alpha} \le e^{C|\lambda|^{\beta}}
$$

for $|\lambda|^{\beta} \ge (\alpha-1)^{1/(\alpha-1)}/2e$ and $C = \frac{\alpha-1}{\alpha} + \frac{2e(\alpha-1)^{-1/(\alpha-1)}}{\alpha} \le 2e+1$. Thus $\mathbb{E}\,e^{\lambda X} \le \exp\left[(2e+1)^{\beta}|\lambda|^{\beta}\right]$.

e $\Rightarrow$ a. Assume $K_5 = 1$. Let $\lambda > 0$.

$$
\mathbb{P}\left(X \ge t\right) = \mathbb{P}\left(e^{\lambda X} \ge e^{\lambda t}\right) \le e^{-\lambda t}\mathbb{E}\,e^{\lambda X} \le \exp\left(-\lambda t + \lambda^{\alpha/(\alpha-1)}\right).
$$

Substituting $\lambda = t^{\alpha-1}(\alpha-1)^{\alpha-1}/\alpha^{\alpha-1}$ gives

$$
\mathbb{P}\left(X \ge t\right) \le \exp\left(-\frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha}}t^{\alpha}\right) \le \exp\left(-t^{\alpha}/2^{\alpha}\right).
$$

Repeating this argument for $-X$, we obtain the same bound for $\mathbb{P}(X \le -t)$. We conclude that

$$\mathbb{P}(|X| \ge t) \le 2\exp\left(-t^\alpha/2^\alpha\right).$$

c $\Rightarrow$ f. Assume $K_3 = 1$. Let $\beta = \alpha/(\alpha-1)$. By Young's inequality,

$$\mathbb{E}\, e^{\lambda X} \le e^{|\lambda|^\beta/\beta}\mathbb{E}\, e^{|X|^\alpha/\alpha} \le e^{|\lambda|^\beta/\beta}e^{1/\alpha} \le \exp\left(C|\lambda|^\beta\right)$$

for $|\lambda| \ge 1/2e$ and $C = \frac{1}{\beta} + \frac{(2e)^\beta}{\alpha} \le (2e+1)^\beta$.

f $\Rightarrow$ a. Assume $K_6 = 1$. If $t \le 2(\log 2)^{1/\alpha}$, we have $2\exp(-t^\alpha/2^\alpha) \ge 1$. Otherwise,

$$\lambda := t^{\alpha-1}(\alpha-1)^{\alpha-1}/\alpha^{\alpha-1} \ge 2^{\alpha-1}(\log 2)^{(\alpha-1)/\alpha}\frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha-1}} \ge 1/2e.$$

From the proof of e $\Rightarrow$ a, we have

$$\mathbb{P}(|X| \ge t) \le 2\exp\left(-t^\alpha/2^\alpha\right).$$

$\square$

Exercise 2.7.4. Let $X \sim Exp(1)$. For $|\lambda| < 1$, we have $\mathbb{E}\exp(\lambda X) = \frac{1}{1-\lambda}$. Suppose $f(x) := e^{kx} - \frac{1}{1-x} \ge 0$ for $x \in [-\delta, \delta]$ with $\delta > 0$. $f(\delta) \ge 0$ gives $k \ge -\frac{\log(1-\delta)}{\delta} > 1$. Then we have $f'(0) > 0$. Note that $f(0) = 0$. This induces a contradiction.

Exercise 2.7.11. The proof is similar to Exercise 2.5.7.

Exercise 2.8.5. For $|z| < 3$, we have

$$e^z = \sum_{k=0}^\infty \frac{z^k}{k|} \le 1 + z + \frac{z^2}{2}\sum_{k=0}^\infty \frac{|z|^k}{(k+2)!/2} \le 1 + z + \frac{z^2}{2}\sum_{k=0}^\infty \frac{|z|^k}{3^k} = 1 + z + \frac{z^2/2}{1-|z|/3}.$$

Then

$$\mathbb{E}\exp(\lambda X) \le 1 + \mathbb{E}\,\frac{\lambda^2X^2/2}{1-\lambda|X|/3} \le \exp(\mathbb{E}\, X^2 g(\lambda)).$$

Exercise 2.8.6.
$$\mathbb{P}\left\{\sum_{i=1}^{N} X_i \geq t\right\} \leq e^{-\lambda t}\prod_{i=1}^{N} e^{\lambda X_i} \leq \exp\left(-\lambda t + g(\lambda)\sigma^2\right).$$

Substituting $\lambda = \frac{t}{\sigma^2+Kt/3}$ gives

$$\mathbb{P}\left\{\sum_{i=1}^{N} X_i \geq t\right\} \leq \exp\left(-\frac{t^2/2}{\sigma^2 Kt/3}\right).$$

Repeating the argument for $-X_i$, we obtain the same bound for $\mathbb{P}\{\sum_{i=1}^{N} \leq -t\}$. A combination of the two bounds completes the proof.

3 Random vectors in high dimensions

Exercise 3.1.4

(a) By Theorem 3.1.1 we have $\|\|X\|_2 - \sqrt{n}\|_{\psi_2} \leq C_0 K^2$. Recall that for sub-gaussian r.v. X, $\|X\|_{L^p} \leq C_1\|X\|_{\psi_2}\sqrt{p}$, for all $p \geq 1$. Hence we have:

$$|\mathbb{E}\|X\|_2 - \sqrt{n}| \leq \mathbb{E}\,|\|X\|_2 - \sqrt{n}| \leq C_1\|X\|_{\psi_2} \leq C_0C_1K^2$$

The proof is completed.

(b) By the same argument used in **(a)** we can get $\mathbb{E}\,(\|X\|_2 - \sqrt{n})^2 \leq (CK^2\sqrt{2})^2$, which means $2\sqrt{n}\mathbb{E}\,\|X\|_2 \geq n + \mathbb{E}\,\|X\|_2^2 - 2C^2K^4$. Note that $\mathbb{E}\,\|X\|_2^2 = n$, thus $\mathbb{E}\,\|X\|_2 \geq \sqrt{n} - 2C^2K^4/\sqrt{n} = \sqrt{n} + o(1)$. And it is trivial that $E\|X\|_2 \leq (E\|X\|_2^2)^{1/2} = \sqrt{n}$. So the answer is yes.

Exercise 3.1.5 Use Exercise 3.1.4 we have $\mathbb{E}\,\|X\|_2 \geq \sqrt{n} - 2C^2K^4/\sqrt{n}$. Therefore,

$$\mathrm{Var}\,(\|X\|_2) = \mathbb{E}\,\|X\|_2^2 - (\mathbb{E}\,\|X\|_2)^2 \leq n - \left(\sqrt{n} - \frac{2C^2K^4}{\sqrt{n}}\right)^2 = 4C^2K^4 - \frac{4C^4K^8}{n}$$

The proof is completed.

Exercise 3.1.6 We have

$$\mathbb{E}\,(\|X\|_2 - \sqrt{n})^2 \le \frac{\mathbb{E}\,(\|X\| - \sqrt{n})^2(\|X\| + \sqrt{n})^2}{n} = \frac{\mathbb{E}\,(\|X\|_2^2 - n)^2}{n}.$$

Note that

$$\begin{aligned}
\mathbb{E}\,(\|X\|_2^2 - n)^2 &= \mathbb{E}\,\|X\|_2^4 + n^2 - 2n\mathbb{E}\,\|X\|_2^2 \\
&= \sum_{i\neq j} \mathbb{E}\,X_i^2 X_j^2 + \sum_i \mathbb{E}\,X_i^4 = nK^4 + n(n-1) + n^2 - 2n^2 \\
&\le nK^4.
\end{aligned}$$

Therefore we can get $\mathrm{Var}\,(\|X\|_2) \le \mathbb{E}\,(\|X\|_2 - \sqrt{n})^2 \le \mathbb{E}\,(\|X\|_2^2 - n)^2/n = K^4$.

Exercise 3.1.7 For any $\lambda > 0$, we have

$$\begin{aligned}
\mathbb{P}\,(\|X\|_2 \le \epsilon\sqrt{n}) &= \mathbb{P}\,(-\|X\|_2^2 \ge -\epsilon^2 n) \\
&= \mathbb{P}\,(e^{-\lambda\|X\|_2^2} \ge e^{-\lambda\epsilon^2 n}) \\
&\le \frac{\mathbb{E}\,e^{-\lambda\|X\|_2^2}}{e^{-\lambda\epsilon^2 n}} \qquad\qquad , \\
&= \left(\frac{\mathbb{E}\,e^{-\lambda X_1^2}}{e^{-\lambda\epsilon^2}}\right)^n
\end{aligned}$$

and

$$\mathbb{E}\,e^{-\lambda x_1^2} = \int e^{-\lambda t^2} f(t)dt \le \int e^{-\lambda t^2} dt = \sqrt{\frac{\pi}{\lambda}}.$$

Setting $\lambda = \epsilon^{-2}$ yields

$$\begin{aligned}
\mathbb{P}\,(\|X\|_2 \le \epsilon\sqrt{n}) &\le \left(\frac{\mathbb{E}\,e^{-\lambda X_1^2}}{e^{-\lambda\epsilon^2}}\right)^n \\
&\le (e^{\lambda\epsilon^2}\sqrt{\pi/\lambda})^n \\
&= (e\sqrt{\pi}\epsilon)^n.
\end{aligned}$$

Exercise 3.3.1 By rotational invariance property of X, it suffices to prove $\mathbb{E}\,X_1^2 = 1$ and $\mathbb{E}\,X_1X_2 = 0$. Since we always have $X_1^2 + \cdots + X_n^2 = n$, $\mathbb{E}\,X_1^2 = 1$ holds by the rotational invariance, too. Note that $X_1X_2 \overset{d}{=} -X_1X_2$, so we have $\mathbb{E}\,X_1X_2 = 0$. It is obvious the coordinates of

X are not independent, as $X_1^2 + \cdots + X_n^2 = n$ always holds.

Exercise 3.3.7 Note that $dg = r^{n-1} dr d\sigma(\theta)$, where $\sigma(\theta)$ denotes the area element of $\mathbb{S}^{n-1}$. Thus we have $\mathrm{e}^{-\|g\|_2^2/2}\mathrm{d}g = C\mathrm{e}^{-r^2/2} r^{n-1}\mathrm{d}r \cdot \mathrm{d}\sigma(\theta)$, which completes the proof.

Exercise 3.4.3

1. By triangular inequality we have

$$\begin{aligned}
\|\langle X, x\rangle\|_{\psi_2} &= \left\|\sum_{i=1}^n X_i x_i\right\|_{\psi_2} \\
&\leq \sum_{i=1}^n \|X_i x_i\|_{\psi_2} \\
&= \sum_{i=1}^n |x_i| \|X_i\|_{\psi_2} \\
&\leq \left(\sum_{i=1}^n \|X_i\|_{\psi_2}^2\right)^{1/2}, \quad \forall x \in \mathbb{S}^{n-1}.
\end{aligned}$$

The last inequality holds by Cauchy-Schwartz's inequality.

2. Let $X_1 \sim N(0, 1)$, and we simply define $X = (X_1, \ldots, X_1)$. Then

$$\begin{aligned}
\|X\|_{\psi_2} &\geq \|\langle X, 1_n/\sqrt{n}\rangle\|_{\psi_2} \\
&= \sqrt{n}\|X_1\|_{\psi_2} \\
&\gg \|X_1\|_{\psi_2}.
\end{aligned}$$

Exercise 3.4.4 First we'll verify $\|X_i\|_{\psi_2} = \sqrt{n/\log(n+1)}$, $\forall i \leq n$. Recall that:

$$\|X_i\|_{\psi_2} = \inf\{K > 0 : \mathbb{E}\exp\{X_i^2/K^2\} \leq 2\}.$$

Let $K = \sqrt{n/\log(n+1)}$ and plug it into $\mathbb{E}\exp\{X^2/K^2\}$ yields:

$$\begin{aligned}\mathbb{E}\exp\{X_i^2/K^2\} &= \frac{1}{n}\exp\{n/K^2\} + \frac{n-1}{n} \\ &= \frac{n+1}{n} + \frac{n-1}{n} \\ &= 2.\end{aligned}$$

So $\|X_i\|_{\psi_2} = \sqrt{n/\log(n+1)},\ i \le n$, and $\|X\|_{\psi_2} \ge \|\langle X, e_i\rangle\|_{\psi_2} = \|X_i\|_{\psi_2} = \sqrt{n/\log(n+1)}$. In addition, $\forall x \in \mathbb{S}^{n-1}$, we have

$$\begin{aligned}\mathbb{E}\exp\{\langle X, x\rangle^2/K^2\} &= \frac{1}{n}\sum_{i=1}^{n}\exp\{nx_i^2/K^2\} \\ &\le \frac{1}{n}\exp\{n/K^2\} + \frac{n-1}{n} \quad \text{(by Karamata's inequality)} \\ &= \mathbb{E}\exp\{X_i^2/K^2\}\end{aligned}$$

Therefore, $\|X\|_{\psi_2} \le \|X_i\|_{\psi_2} = \sqrt{n/\log(n+1)}$. And we may conclude that $\|X\|_{\psi_2} = \sqrt{n/\log(n+1)}$.

Exercise 3.4.5 Recall that:

$$\|X\|_{\psi_2} = \sup_{\|\mathbf{y}\|=1} \inf\{K > 0 : \mathbb{E}\exp\{(X^T\mathbf{y})^2/K^2\} \le 2\}.$$

By the assumption $\|X\|_{\psi_2} \le C$,

$$\exists C_0 > 0, \text{ s.t. } \mathbb{E}\exp\{(X^T\mathbf{y})^2/C_0^2\} \le 3,\ \forall\, \mathbf{y} \in \mathbb{S}^{d-1}.$$

Let $T := \{\mathbf{x}_1, \dots,\ \mathbf{x}_{|T|}\}, P(X = \mathbf{x}_i) := p_i$.

$$p_j\exp\{(\mathbf{x}_j^\top\mathbf{y})^2/C_0^2\} \le \sum_{i=1}^{|T|} p_i\exp\{(\mathbf{x}_i^\top\mathbf{y})^2/C_0^2\} = \mathbb{E}\exp\{(X^T\mathbf{y})^2/C_0^2\} \le 3$$

Thus

$$\left(\mathbf{x}_j^\top\mathbf{y}\right)^2 \le C_0^2 \ln\frac{3}{p_j},\ \forall\, \mathbf{y} \in \mathbb{S}^{d-1}.$$

Therefore

$$\|\mathbf{x}_j\|^2 \le C_0^2 \ln \frac{3}{p_j},\ \forall\, \mathbf{y} \in \mathbb{S}^{d-1}.$$

Notice that $\mathbb{E}XX^\top = I_n$ (X is isotropic),

$$n = \operatorname{tr}\left(\mathbb{E}XX^\top\right) = \sum_{i=1}^{|T|} p_i \|\mathbf{x}_i\|_2^2 \le \sum_{i=1}^{|T|} C_0^2 p_i \ln \frac{3}{p_i} \le t_0^2 \ln(3|T|).$$

The last inequality follows from $x \ln x$ is convex and Jensen inequality. Hence

$$|T| \ge \frac{1}{3} \exp\left\{\frac{n}{C_0^2}\right\}.$$

Exercise 3.4.7 Since X enjoys the rotational invariance property, it suffices to prove $\|X_1\|_{\psi_2} \le C$. Define $Y = \sqrt{n}\frac{X}{\|X\|_2}$, it follows that $Y \sim \text{Unif}(\sqrt{n}\mathbb{S}^{n-1})$. From Theorem 3.4.6 we know $\|Y_1\|_{\psi_2} \le C$, and $|Y_1| \ge |X_1|$ always holds. So we can get $\|X_1\|_{\psi_2} \le C$.

Exercise 3.4.9

(a) By rotational invariance we have $\mathbb{E}\, X_i X_j = 0,\ i \ne j$ and $\mathbb{E}\, X_1^2 = \cdots \mathbb{E}\, X_n^2$. Then it is clear that one can make the distribution isotropic by scaling the ball. We denote a n-dimensional l_1-ball of radius r by $B_n(r)$. Suppose $X \sim \text{Unif}(B_n(1))$, then

$$\begin{aligned}
\mathbb{P}\left(|X_1| \le t\right) &= \int_{-t}^{t} \int_{-1+|x_1|}^{1-|x_1|} \cdots \int_{-1+|x_1|+\cdots+|x_{n-1}|}^{1-|x_1|-\cdots-|x_{n-1}|} \frac{1}{\text{Vol}(B_n(1))} dx_n \ldots dx_2 dx_1 \\
&= \int_{-t}^{t} \frac{\text{Vol}(B_{n-1}(1-|x_1|))}{\text{Vol}(B_n(1))} dx_1 \\
&= \int_{-t}^{t} (1-|x_1|)^{n-1} dx_1 \frac{\text{Vol}(B_{n-1}(1))}{\text{Vol}(B_n(1))}
\end{aligned}$$

Setting $t = 1$ yields $\mathbb{P}\left(|X_1| \le 1\right) = \frac{2}{n}\frac{\text{Vol}(B_{n-1}(1))}{\text{Vol}(B_n(1))} = 1$, so

$$\mathbb{P}\left(|X_1| > t\right) = 1 - \frac{n}{2}\int_{-t}^{t} (1-|x_1|)^{n-1} dx_1 = (1-t)^n.$$

Now using results from Exercise 1.2.3 we have:

$$\mathbb{E}\, X_1^2 = \int_0^1 2t\mathbb{P}\left(|X_1| > t\right)dt = \int_0^1 2t(1-t)^n dt = \frac{2}{(n+1)(n+2)}.$$

Therefore the scale factor should be $r = \sqrt{(n+1)(n+2)/2}$.

(b) Using similar argument in part (a), we may get for $X \sim \mathrm{Unif}(B_n(r))$,

$$\mathbb{P}\left(|X_1| > t\right) = \left(1 - \frac{t}{r}\right)^n \to e^{-t},\ n \to \infty,$$

since $r \asymp n$. So obviously $\|X_1\|_{\psi_2}$ is not bounded by any an absolute constant as n grows, and neither is $\|X\|_{\psi_2}$.

Exercise 3.4.10 (This example is copied from *https://mathoverflow.net/questions/326183/*) Define $\mu_X = \frac{1}{2}\mu_{aZ} + \frac{1}{2}\mu_{bZ}$, where μ_U denotes the probability distribution of a random vector U, $Z \sim N(0, I_n)$, and $a,\ b$ are constants such that $0 < a < 1 < b,\ a^2 + b^2 = 2$. Then one can verify that X is isotropic. For any unit vector u and a real number $s > 0$,

$$\mathbb{E}\, \exp\left\{\langle X, u\rangle^2/s^2\right\} = \frac{1}{2\sqrt{1 - 2a^2/s^2}} + \frac{1}{2\sqrt{1 - 2b^2/s^2}} < 2,$$

if s is large enough (depending only on $a,\ b$), so that by the definition of sub-gaussian norm we have $\|X\|_{\psi_2} \le s$.
On the other hand, let $t := (b-1)\sqrt{n}/2$, we have

$$\begin{aligned}
2\mathbb{E}\, e^{(\|X\|-\sqrt{n})^2/t^2} &> \mathbb{E}\, e^{(\|bZ\|-\sqrt{n})^2/t^2} \\
&> \mathbb{E}\, e^{(\|bZ\|-\sqrt{n})^2/t^2} 1_{\|Z\|^2 > n} \\
&> e^{(b\sqrt{n}-\sqrt{n})^2/t^2}\mathbb{P}\left(\|Z\|^2 > n\right) \\
&= e^4\mathbb{P}\left(\|Z\|^2 > n\right) \to e^4/2 \\
&> 4,
\end{aligned}$$

since $\mathbb{P}\left(\|Z\|^2 > n\right) \to 1/2$ by CLT. Therefore, $\|\|X\| - \sqrt{n}\|_{\psi_2} \ge t = (b-1)\sqrt{n}/2 \to \infty$, as desired.

Exercise 3.5.3 First we'll prove the following lemma:

Lemma. *Suppose A is either positive-semidefinite or has zero diagonal, we have*

$$\max_{x\in\{-1,1\}^n} |\langle Ax, x\rangle| = \max_{x\in\{-1,0,1\}^n} |\langle Ax, x\rangle|$$

Proof. Obviously we have $\max_{x\in\{-1,1\}^n} |\langle Ax, x\rangle| \le \max_{x\in\{-1,0,1\}^n} |\langle Ax, x\rangle|$. And next we will prove $\max_{x\in\{-1,1\}^n} |\langle Ax, x\rangle| \ge \max_{x\in\{-1,0,1\}^n} |\langle Ax, x\rangle|$.

When A is positive-semidefinite, we always have $|\langle Ax, x\rangle| = \langle Ax, x\rangle$, and $\langle Ax, x\rangle$ is convex in x. Also note that $[-1,1]^n = \mathrm{conv}\{u : u \in \{-1,1\}^n\}$ is a polyhedron with vertices $\{u : u \in \{-1,1\}^n\}$, so $\langle Ax, x\rangle$ must attain its maximum at one of the vertices. Then we may get

$$\max_{x\in\{-1,0,1\}^n} \langle Ax, x\rangle \le \max_{x\in[-1,1]^n} \langle Ax, x\rangle = \max_{x\in\{-1,1\}^n} \langle Ax, x\rangle$$

When A has zero diagonal, suppose

$$\max_{x\in\{-1,1\}^n} |\langle Ax, x\rangle| < \max_{x\in\{-1,0,1\}^n} |\langle Ax, x\rangle|,$$

define $x_0 = \mathrm{argmax}_{x\in\{-1,0,1\}^n} \langle Ax, x\rangle$, then x_0 has at least one zero coordinate, thus we can assume $x_0(i_0) = 0$. Consider $x \in \{-1,0,1\}^n$ such that $x(j) = x_0(j)$, $\forall j \ne i_0$, we have

$$\langle Ax, x\rangle - \langle Ax_0, x_0\rangle = 2\sum_{j\ne i_0} a_{i_0 j} x(i_0) x(j) = 2x(i_0)\sum_{j\ne i_0} a_{i_0 j} x(j).$$

Then setting $x_1 = x_0 + \mathrm{sgn}(\langle Ax_0, x_0\rangle) \cdot \mathrm{sgn}(\sum_{j\ne i_0} a_{i_0 j} x(j)) \cdot e_{i_0}$ yields $|\langle Ax_1, x_1\rangle| \ge |\langle Ax_0, x_0\rangle|$. Repeat this procedure for at most n times we will get a $x_* \in \{-1,1\}^n$ such that $|\langle Ax_*, x_*\rangle| \ge |\langle Ax_0, x_0\rangle|$, which is a contradiction. $\square$

Suppose $x, y \in \{-1,1\}^n$, by polarization identity we'll get $\langle Ax, y\rangle = \langle Au, u\rangle - \langle Av, v\rangle$, where $u = (x+y)/2 \in \{-1,0,1\}$ and $v = (x-y)/2 \in \{-1,0,1\}$. Since A is positive-semidefinite or has zero diagonal and $|\langle Ax, x\rangle| \le 1$, $\forall x \in \{-1,1\}^n$, from the lemma we know $|\langle Au, u\rangle| \le 1$, $|\langle Av, v\rangle| \le 1$. Thus $|\langle Ax, y\rangle| \le |\langle Au, u\rangle| + |\langle Av, v\rangle| = 2$, $\forall x, y \in \{-1,1\}$. Then the conclusion holds via Grothendieck's inequality.

Exercise 3.6.4 The algorithm is simple. We may just repeat the procedure in Proposition 3.6.3, and terminate as long as the number of cuts exceeds $(0.5-\epsilon)|E|$. Define $p_0 = \mathbb{P}\,(\text{a single try fails})$, then we may find

$$\begin{aligned}
p_0 &= \mathbb{P}\left(\frac{1}{4}\sum_{i,j=1}^{n} A_{ij}(1-x_i x_j) \le \left(\frac{1}{2}-\epsilon\right)|E|\right) \\
&= \mathbb{P}\left(\frac{1}{4}\sum_{i,j=1}^{n} A_{ij}x_i x_j \ge \epsilon|E|\right) \\
&= \mathbb{P}\left(\frac{1}{4}\sum_{i,j=1}^{n} A_{ij}\,\mathrm{sign}(x_i)\,\mathrm{sign}(x_j) \ge \epsilon|E|\right)
\end{aligned}$$

Let $f(x) = \frac{1}{4}\sum_{i,j=1}^{n} A_{ij}\,\mathrm{sign}(x_i)\,\mathrm{sign}(x_j)$. By Theorem 2.9.1, we have

$$p_0 \le \exp\left(-\frac{8\epsilon^2|E|^2}{\sum_{i=1}^{n} d_i^2}\right),$$

where d_i is the degree of the i-th vertex. It remains to bound $\sum_{i=1}^{n} d_i^2$. D.de Caen [2] has proved

$$\sum_{i=1}^{n} d_i^2 \le |E|\left(\frac{2|E|}{n-1}+n-2\right) \quad \text{when } n \ge 2.$$

Then we obtain

$$p_0 \le \exp\left(-\frac{8\epsilon^2|E|}{2|E|/(n-1)+n-2}\right) \quad \text{when } n \ge 2.$$

Futhermore, the number of runs needed N obeys the geometric distribution with parameter $1-p_0$. And $\mathbb{P}\,(N<\infty)=1$, $\mathbb{E}\,N = 1/(1-p_0)$.

Exercise 3.6.7 First we set $(X_1, X_2) := (\langle g,u\rangle, \langle g,v\rangle) \sim N(0, [\begin{smallmatrix} 1 & \cos\alpha \\ \cos\alpha & 1 \end{smallmatrix}])$, then we have

$$\begin{aligned}
&\mathbb{E}\,\mathrm{sign}\langle g,u\rangle \mathrm{sign}\langle g,v\rangle \\
&= \mathbb{P}\,(X_1>0, X_2>0) + \mathbb{P}\,(X_1 \le 0, X_2 \le 0) - \mathbb{P}\,(X_1 \le 0, X_2 > 0) - \mathbb{P}\,(X_1 > 0, X_2 \le 0) \\
&= 2\mathbb{P}\,(X_1>0, X_2>0) - (1 - 2\mathbb{P}\,(X_1>0, X_2>0)) \qquad \text{(by symmetry)} \\
&= 4\mathbb{P}\,(X_1>0, X_2>0) - 1
\end{aligned}$$

Note that $(X_1, X_2) \overset{d}{=} (Z_1, \sin\alpha Z_1 + \cos\alpha Z_2)$, where $(Z_1, Z_2) \sim N(0, I_2)$. So

$$\mathbb{P}(X_1 > 0, X_2 > 0) = \mathbb{P}(Z_1 > 0, \sin\alpha Z_1 + \cos\alpha Z_2 > 0) = \frac{\pi - \alpha}{2\pi}.$$

Thus $\mathbb{E}\,\mathrm{sign}\langle g, u\rangle \mathrm{sign}\langle g, v\rangle = 4\mathbb{P}(X_1 > 0, X_2 > 0) - 1 = (\pi - 2\alpha)/\pi = 2\arcsin\langle u, v\rangle/\pi$.

References

[1] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013.

[2] D. de Caen. An upper bound on the sum of squares of degrees in a graph. *Discrete Mathematics*, 185(1):245–248, 1998.

[3] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[4] Gautam Kamath. Bounds on the expectation of the maximum of samples from a gaussian. 2015.

[5] S. Y. Novak. Poisson approximation. *arXiv: Probability*, 2018.

[6] Francesco Orabona and David Pal. Optimal non-asymptotic lower bound on the minimax regret of learning with expert advice. *arXiv preprint arXiv:1511.02176*, 2015.