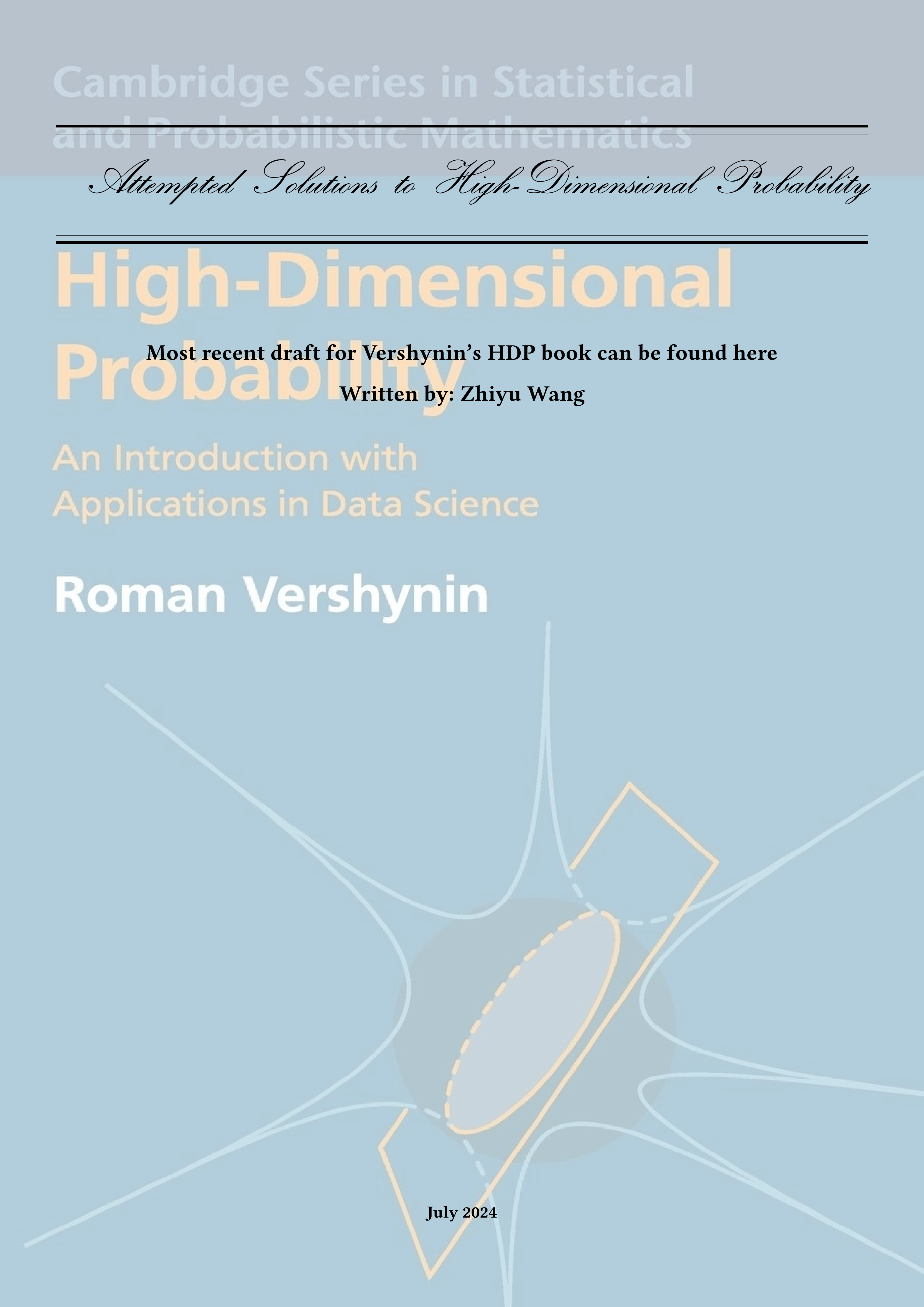
Attempted Solutions to High-Dimensional Probability

Most recent draft for Vershynin's HDP book can be found here

Written by: Zhiyu Wang



July 2024

Contents

1 Preliminaries on random variables

1.1 Basic quantities associated with random variables

1.2 Some classical inequalities

> ***Exercise 1.2.2*** *(Generalization of integral identity).* Prove the following extension of Lemma 1.2.1, which is valid for any random variable X (not necessarily non-negative):
>
> $$\mathbb{E}X = \int_0^\infty \mathbb{P}\{X > t\}dt - \int_{-\infty}^0 \mathbb{P}\{X < t\}dt$$

Attempted Solution 1.2.2. Write $X = X\mathbf{1}_{\{X\geq 0\}} - (-X\mathbf{1}_{\{X<0\}})$ where $X\mathbf{1}_{\{X\geq 0\}}$ and $(-X\mathbf{1}_{\{X<0\}})$ are non-negative random variables. Observe that

$$\begin{aligned}
\mathbb{E}X &\overset{(1)}{=} \mathbb{E}X\mathbf{1}_{\{X\geq 0\}} - \mathbb{E}(-X\mathbf{1}_{\{X<0\}}) \\
&\overset{(2)}{=} \int_0^\infty \mathbb{P}\{X\mathbf{1}_{\{X\geq 0\}} > t\}dt - \int_0^\infty \mathbb{P}\{-X\mathbf{1}_{\{X<0\}} > u\}du \\
&\overset{(3)}{=} \int_0^\infty \mathbb{P}\{X > t\}dt - \int_0^\infty \mathbb{P}\{X < -u\}du \\
&\overset{(4)}{=} \int_0^\infty \mathbb{P}\{X > t\}dt - \int_{-\infty}^0 \mathbb{P}\{X < t\}dt.
\end{aligned}$$

(1) linearity of expectation;

(2) Lemma 1.2.1;

(3) $\{X\mathbf{1}_{\{X\geq 0\}} > t\} = \{X > t\}$ and $\{-X\mathbf{1}_{\{X<0\}} > u\} = \{X < -u\}$ as they are the same events respectively;

(4) change of variable $t = -u$ in the second term.

> ***Exercise 1.2.3*** *(p-moments via tails).* Let X be a random variable and $p \in (0, \infty)$. Show that
>
> $$\mathbb{E}|X|^p = \int_0^\infty pt^{p-1}\mathbb{P}\{|X| > t\}dt$$
>
> whenever the right hand side is finite.
> **Hint:** Use the integral identity for $|X|^p$ and change variables.

Attempted Solution 1.2.3. Observe that

$$\begin{aligned}
\mathbb{E}|X|^p &\overset{(1)}{=} \int_0^\infty \mathbb{P}\left\{|X|^p \geq u\right\} du \\
&\overset{(2)}{=} \int_0^\infty \mathbb{P}\{|X| \geq t\}pt^{p-1}dt.
\end{aligned}$$

(1) Lemma 1.2.1;

(2) change of variables $u = t^p$.

> ***Exercise 1.2.6.*** Deduce Chebyshev's inequality by squaring both sides of the bound $|X - \mu| \geq t$ and applying Markov's inequality.

Attempted Solution 1.2.6. Observe that

$$\mathbb{P}\{|X-\mu| \geq t\} \overset{(1)}{=} \mathbb{P}\{(X-\mu)^2 \geq t^2\} \overset{(2)}{\leq} \frac{\mathbb{E}(X-\mu)^2}{t^2} = \frac{\sigma^2}{t^2}.$$

(1) squaring both sides of the bound $|X - \mu| \geq t$;

(2) Markov's inequality.

1.3 Limit theorems

Exercise 1.3.3. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with mean μ and finite variance. Show that

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right| = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \qquad \text{as } N \to \infty.$$

Attempted Solution 1.3.3. Observe that

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right| \overset{(1)}{=} \left\|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right\|_{L^1} \overset{(2)}{\leq} \left\|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right\|_{L^2} \overset{(3)}{=} \sqrt{\mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right)} \overset{(4)}{=} \frac{\sqrt{\mathrm{Var}(X_1)}}{\sqrt{N}}.$$

(1) definition of L^1 norm;

(2) $\|X\|_{L^p} \leq \|X\|_{L^q}$ for any $0 \leq p \leq q \leq \infty$;

(3) $\mathbb{E}\frac{1}{N}\sum_{i=1}^{N} X_i = \frac{1}{N} N\mu = \mu$;

(4) finite variance

2 Concentration of sums of independent random variables

2.1 Why concentration inequalities?

Exercise 2.1.4 *(Truncated normal distribution).* Let $g \sim \mathcal{N}(0,1)$. Show that for all $t \geq 1$, we have

$$\mathbb{E} g^2 \mathbf{1}_{\{g>t\}} = t \cdot \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \mathbb{P}\{g > t\} \leq \left(t + \frac{1}{t}\right) \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Hint: Integrate by parts.

Attempted Solution 2.1.4. Recall that the density of $Y \sim \mathcal{N}(0,1)$ is

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, y \in \mathbb{R}.$$

Note that $f'(y) = -yf(y)$. We have

$$\begin{aligned}
\mathbb{E} g^2 \mathbf{1}_{\{g>t\}} &\overset{(1)}{=} \int_t^\infty x^2 f(x) dx \\
&\overset{(2)}{=} \int_t^\infty -x f'(x) dx \\
&\overset{(3)}{=} -\left([x f(x)]_t^\infty - \int_t^\infty f(x) dx \right) \\
&\overset{(4)}{=} \int_t^\infty f(x) dx + t f(x) \\
&\overset{(5)}{=} \mathbb{P}\{g > t\} + t \cdot \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \\
&\overset{(6)}{\leq} \left(t + \frac{1}{t}\right) \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.
\end{aligned}$$

(1) can be viewed as the definition of expectation of functions of a continuous random variable with density f. But rigorously, it is really the corollary of the following proposition (Check the details if you care!):
Proposition. If $h : \mathbb{R} \to \mathbb{R}$ is Lebesgue integrable on $(\mathbb{R}, \mathcal{B}, \mu)$, then

$$Eh(X) = \int_{\mathbb{R}} h(t) d\mu(t).$$

Here $\mathcal{B}$ is the Borel σ-algbera on $\mathbb{R}$ and μ is the law of random variable X.
Corollary. If X has density f, then $Eh(X) = \int_{\mathbb{R}} h(t) f(t) dt$ and in particular $EX = \int_{\mathbb{R}} t f(t) dt$.

(2) already checked;

(3) integration by parts with $u = x$ and $dv = f'(x)dx$;

(4) $\lim_{x\to\infty} x f(x) = 0$;

(5) clear;

(6) upper bound in Proposition 2.1.2.

2.2 Hoeffding's inequality

Exercise 2.2.3 *(Bounding the hyperbolic cosine).* Show that $\cosh(x) \leq \exp(x^2/2)$ for all $x \in \mathbb{R}$.
Hint: Compare the Taylor's expansions of both sides.

Attempted Solution 2.2.3. By the definition of the exponential function and hyperbolic cosine, we have

$$\exp\left(\frac{x^2}{2}\right) := \sum_{n=0}^{\infty} \frac{x^{2n}}{2^n n!},$$
$$\cosh(x) := \frac{e^x + e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{x^n + (-x)^n}{2n!} \overset{\text{odd terms vanish}}{=} \sum_{n=0}^{\infty} \frac{2x^{2n}}{2(2n)!}.$$

It remains to check that $(2n)! \geq 2^n n!$ which follows easily from

$$\underbrace{1 \ldots n}_{\text{same}} \underbrace{(n+1) \cdots (2n)}_{\text{bigger}} \geq \underbrace{1 \ldots n}_{\text{same}} \underbrace{2 \ldots 2}_{\text{smaller}}.$$

Exercise 2.2.7. Prove Theorem 2.2.6, possibly with some absolute constant instead of 2 in the tail.

Attempted Solution 2.2.7.

Exercise 2.2.8 *(Boosting randomized algorithms).* Imagine we have an algorithm for solving some decision problem (e.g. is a given number p a prime?). Suppose the algorithm makes a decision at random and returns the correct answer with probability $\frac{1}{2} + \delta$ with some $\delta > 0$, which is just a bit better than a random guess. To improve the performance, we run the algorithm N times and take the majority vote. Show that, for any $\varepsilon \in (0, 1)$, the answer is correct with probability at least $1 - \varepsilon$, as long as

$$N \geq \frac{1}{2\delta^2} \ln\left(\frac{1}{\varepsilon}\right).$$

Hint: Apply Hoeffding's inequality for X_i being the indicators of the wrong answers.

Attempted Solution 2.2.8.

Exercise 2.2.9 *(Robust estimation of the mean).* Suppose we want to estimate the mean μ of a random variable X from a sample $X_1, \ldots, X_N$ drawn independently from the distribution of X. We want an ε-accurate estimate, i.e. one that falls in the interval $(\mu - \varepsilon, \mu + \varepsilon)$. (a) Show that a sample [2] of size $N = O\left(\sigma^2/\varepsilon^2\right)$ is sufficient to compute an ε-accurate estimate with probability at least 3/4, where $\sigma^2 = \operatorname{Var} X$. Hint: Use the sample mean $\hat{\mu} := \frac{1}{N}\sum_{i=1}^{N} X_i$. (b) Show that a sample of size $N = O\left(\log\left(\delta^{-1}\right)\sigma^2/\varepsilon^2\right)$ is sufficient to compute an ε-accurate estimate with probability at least $1 - \delta$. **Hint:** Use the median of $O\left(\log\left(\delta^{-1}\right)\right)$ weak estimates from part 1.

Attempted Solution 2.2.9.

Exercise 2.2.10 *(Small ball probabilities).* Let $X_1, \ldots, X_N$ be non-negative independent random variables with continuous distributions. Assume that the densities of X_i are uniformly bounded by 1 . (a) Show that the MGF of X_i satisfies

$$\mathbb{E}\exp\left(-tX_i\right) \leq \frac{1}{t} \quad \text{for all } t > 0.$$

(b) Deduce that, for any $\varepsilon > 0$, we have

$$\mathbb{P}\left\{\sum_{i=1}^{N} X_i \leq \varepsilon N\right\} \leq (e\varepsilon)^N.$$

Hint: Rewrite the inequality $\sum X_i \leq \varepsilon N$ as $\sum\left(-X_i/\varepsilon\right) \geq -N$ and proceed like in the proof of Hoeffding's inequality. Use part 1 to bound the MGF.

Attempted Solution 2.2.10.

2.3 Chernoff's inequality

Exercise 2.3.2 *(Chernoff's inequality: lower tails).* Modify the proof of Theorem 2.3.1 to obtain the following bound on the lower tail. For any $t < \mu$, we have

$$\mathbb{P}\{S_N \le t\} \le e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

Attempted Solution 2.3.2.

***Exercise 2.3.3** (Poisson tails).* Let $X \sim \text{Pois}(\lambda)$. Show that for any $t > \lambda$, we have

$$\mathbb{P}\{X \ge t\} \le e^{-\lambda}\left(\frac{e\lambda}{t}\right)^t.$$

Hint: Combine Chernoff's inequality with Poisson limit theorem (Theorem 1.3.4).

Attempted Solution 2.3.3.

***Exercise 2.3.5** (Chernoff's inequality: small deviations).* Show that, in the setting of Theorem 2.3.1, for $\delta \in (0, 1]$ we have

$$\mathbb{P}\{|S_N - \mu| \ge \delta\mu\} \le 2e^{-c\mu\delta^2}$$

where $c > 0$ is an absolute constant. **Hint:** Apply Theorem 2.3.1 and Exercise 2.3.2 $t = (1 \pm \delta)\mu$ and analyze the bounds for small δ.

Attempted Solution 2.3.5.

***Exercise 2.3.6** (Poisson distribution near the mean).* Let $X \sim \text{Pois}(\lambda)$. Show that for $t \in (0, \lambda]$, we have

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

Hint: Combine Exercise 2.3.5 with the Poisson limit theorem (Theorem 1.3.4).

Attempted Solution 2.3.6.

***Exercise 2.3.8** (Normal approximation to Poisson).* Let $X \sim \text{Pois}(\lambda)$. Show that, as $\lambda \to \infty$, we have

$$\frac{X - \lambda}{\sqrt{\lambda}} \to N(0, 1) \quad \text{in distribution.}$$

Hint: Derive this from the central limit theorem. Use the fact that the sum of independent Poisson distributions is a Poisson distribution.

Attempted Solution 2.3.8.

2.4 Application: degrees of random graphs

***Exercise 2.4.2** (Bounding the degrees of sparse graphs).* Consider a random graph $G \sim G(n, p)$ with expected degrees $d = O(\log n)$. Show that with high probability (say, 0.9), all vertices of G have degrees $O(\log n)$. **Hint:** Modify the proof of Proposition 2.4.1.

Attempted Solution 2.4.2.

***Exercise 2.4.3** (Bounding the degrees of very sparse graphs).* Consider a random graph $G \sim G(n, p)$ with expected degrees $d = O(1)$. Show that with high probability (say, 0.9), all vertices of G have degrees

$$O\left(\frac{\log n}{\log\log n}\right).$$

Attempted Solution 2.4.3.

***Exercise 2.4.4** (Sparse graphs are not almost regular).* Consider a random graph $G \sim G(n, p)$ with expected degrees $d = o(\log n)$. Show that with high probability, (say, 0.9), G has a vertex with degree $^{3}10d$. **Hint:** The principal difficulty is that the degrees d_i are not independent. To fix this, try to replace d_i by some d_i' that are independent. (Try to include not all vertices in the counting.) Then use Poisson approximation (2.9).

Attempted Solution 2.4.4.

> ***Exercise 2.4.5*** *(Very sparse graphs are far from being regular).* Consider a random graph $G \sim G(n, p)$ with expected degrees $d = O(1)$. Show that with high probability, (say, 0.9), G has a vertex with degree
>
> $$\Omega\left(\frac{\log n}{\log \log n}\right).$$

Attempted Solution 2.4.5.

2.5 Sub-gaussian distributions

> ***Exercise 2.5.1*** *(Moments of the normal distribution).* Show that for each $p \geq 1$, the random variable $X \sim \mathcal{N}(0,1)$ satisfies
>
> $$\|X\|_{L^p} = \left(\mathbb{E}|X|^p\right)^{1/p} = \sqrt{2}\left[\frac{\Gamma((1+p)/2)}{\Gamma(1/2)}\right]^{1/p}.$$
>
> Deduce that
>
> $$\|X\|_{L^p} = \mathcal{O}(\sqrt{p}) \text{ as } p \to \infty.$$
>
> Finally, a classical formula gives the moment generating function of $X \sim \mathcal{N}(0,1)$:
>
> $$\mathbb{E}\exp(\lambda X) = e^{\lambda^2/2} \quad \text{for all } \lambda \in \mathbb{R}.$$

Attempted Solution 2.5.1. Observe that

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

Taking pth roots of both sides gives the result. Now let us focus on the asymptotics whose details can be found here. X has mgf

$$\begin{aligned}
\mathbb{E}\exp(\lambda X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{tx} e^{-\frac{1}{2}x^2}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(tx - \frac{1}{2}x^2\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(\underbrace{-\frac{1}{2}\left(x^2 - 2tx + t^2\right) + \frac{1}{2}t^2}_{\text{Completing the square}}\right) dz \\
&= \exp\left(\frac{1}{2}t^2\right) \int_{-\infty}^\infty \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-t)^2\right)}_{\text{Pdf of } \mathcal{N}(t,1)} dx = e^{t^2/2}.
\end{aligned}$$

> ***Exercise 2.5.4***. Show that the condition $\mathbb{E}X = 0$ is necessary for property (v) to hold.

Attempted Solution 2.5.4. Recall that *Jensen's inequality* states that for any random variable X and a convex function $\varphi : \mathbb{R} \to \mathbb{R}$, we have

$$\varphi(\mathbb{E}X) \leq \mathbb{E}\varphi(X).$$

Observe that $\exp(\lambda\mathbb{E}X) \overset{Jensen}{\leq} \mathbb{E}\exp(\lambda X) \overset{(v)}{\leq} \exp\left(K_5^2\lambda^2\right)$. This implies that $\lambda\mathbb{E}X \leq K_5^2\lambda^2$. Note that $\forall\lambda > 0$, we have

$$-K_5^2\lambda \leq \mathbb{E}X \leq K_5^2\lambda.$$

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

***Exercise 2.5.5** (On property iii in Proposition 2.5.2).* (a) Show that if $X \sim \mathcal{N}(0,1)$, the function $\lambda \mapsto \mathbb{E}\exp\left(\lambda^2X^2\right)$ is only finite in some bounded neighborhood of zero.

(b) Suppose that some random variable X satisfies $\mathbb{E}\exp\left(\lambda^2X^2\right) \leq \exp\left(K\lambda^2\right)$ for all $\lambda \in \mathbb{R}$ and some constant K. Show that X is a bounded random variable, i.e. $\|X\|_\infty < \infty$.

Attempted Solution 2.5.5.
(a) Observe that $\mathbb{E}\left[\exp\left(\lambda^2X^2\right)\right] = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{\lambda^2x^2}e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\exp\left(\left(\lambda^2-\frac{1}{2}\right)x^2\right)dx$. Now note that this integral is finite if and only if $\lambda^2 - \frac{1}{2} < 0$ if and only if $\lambda \in (-\frac{1}{2}, \frac{1}{2})$. It follows from the fact that $\int_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi}$.
(b) We want to show that X is almost surely bounded. Fix $\epsilon > 0$. Consider the probability that $|X| \geq \sqrt{K+\epsilon}$. We have

$$\begin{aligned}
\mathbb{P}\left\{|X| \geq \sqrt{K+\epsilon}\right\} &\overset{(1)}{=} \mathbb{P}\left\{e^{\lambda^2X^2} \leq e^{\lambda^2(K+\epsilon)}\right\}\\
&\overset{(2)}{\leq} \inf_{\lambda\in\mathbb{R}}\left(e^{\lambda^2(K+\epsilon)}\mathbb{E}e^{\lambda^2X^2}\right)\\
&\overset{(3)}{\leq} \inf_{\lambda\in\mathbb{R}}\left(e^{\lambda^2(K+\epsilon)}e^{K\lambda^2}\right)\\
&= \inf_{\lambda\in\mathbb{R}}(e^{-\lambda^2\epsilon})\\
&= 0.
\end{aligned}$$

(1) multiply both sides by λ^2 and take the exponential;

(2) Markov's inequality and optimize over $\lambda \in \mathbb{R}$;

(3) assumption.

Thus, $\mathbb{P}\left\{|X| \geq \sqrt{K+\epsilon}\right\} = 0$. Hence $\mathbb{P}\left\{|X| < \sqrt{K+\epsilon}\right\} = 1 - \mathbb{P}\left\{|X| \leq \sqrt{K+\epsilon}\right\} = 1$.

***Exercise 2.5.7**.* Check that $\|\cdot\|_{\psi_2}$ is indeed a norm on the space of subgaussian random variables.

Attempted Solution 2.5.7.

***Exercise 2.5.9**.* Check that Poisson, exponential, Pareto and Cauchy distributions are not sub-gaussian.

Attempted Solution 2.5.9.

- If $X \sim \text{Posi}(\lambda)$, then
$$\mathbb{P}\{X \geq t\} \geq \mathbb{P}\{X = \lceil t\rceil\} = e^{-\lambda}\lambda^{\lceil t\rceil}/\lceil t\rceil! = e^{-\lambda}\lambda^s/s!$$
where $s := \lceil t\rceil$ for simplicity. Stirling's approximation gives $s! \sim \sqrt{2\pi s}\left(\frac{s}{e}\right)^s$. Hence
$$\frac{e^{-\lambda}\lambda^s}{s!} \sim \frac{e^{-\lambda}\lambda^s}{\sqrt{2\pi s}\left(\frac{s}{e}\right)^s} \sim \frac{e^{-\lambda}\lambda^s}{s^{1/2}\left(\frac{s}{e}\right)^s} = \frac{e^{-\lambda}}{s^{1/2}}\left(\frac{\lambda e}{s}\right)^s.$$
Taking the logarithm of the above RHS, we have
$$\log\left(\frac{e^{-\lambda}}{s^{1/2}}\left(\frac{\lambda e}{s}\right)^s\right) = -\lambda + s\log\lambda + s - s\log s - \frac{1}{2}\log s + cs^2 = cs^2 - s\log s + \mathcal{O}(s) \overset{s\to\infty}{\longrightarrow} \infty$$
since the term cs^2 will dominate $-s\log s$ for any positive c as $s \to \infty$. Hence the tail decay of Poisson distribution is strictly slower than $\exp(-ct^2)$ meaning that it is not sub-gaussian.

- If $X \sim \text{Exp}(\lambda)$, then
$$\mathbb{P}\{X \geq t\} = \exp(-\lambda t)$$
which decays strictly slower than $\exp(-ct^2)$ meaning that it is not sub-gaussian.

- If $X \sim \text{Pareto}(a, \theta)$, then
$$\mathbb{P}\{X \geq t\} = \left(\frac{a}{t}\right)^{\theta} = \exp(\theta \log a - \theta \log t)$$
which decays strictly slower than $\exp(-ct^2)$ meaning that it is not sub-gaussian.

- If $X \sim$ Cauchy (μ, σ), then $\mathbb{E}|X| = \infty$. Since the first moment of X is not bounded by any constant, (ii) implies that X is not sub-gaussian.

Exercise 2.5.10 *(Maximum of sub-gaussians).* Let $X_1, X_2, \ldots$, be a sequence of sub-gaussian random variables, which are not necessarily independent. Show that

$$\mathbb{E} \max_i \frac{|X_i|}{\sqrt{1 + \log i}} \leq CK,$$

where $K = \max_i \|X_i\|_{\psi_2}$. Deduce that for every $N \geq 2$ we have

$$\mathbb{E} \max_{i \leq N} |X_i| \leq CK\sqrt{\log N}$$

Hint: Denote $Y_i := X_i/(CK\sqrt{1 + \log i})$ with absolute constant C chosen sufficiently large. Use subgaussian tail bound (2.14)and then a union bound to conclude that $\mathbb{P}\{\exists i : |Y_i| \geq t\} \lesssim e^{-t^2}$ for any $t \geq 1$. Use the integrated tail formula (Lemma 1.2.1), breaking the integral into two integrals: one over $[0, 1]$ (whose value should be trivial to bound) and the other over $[1, \infty)$ (where you can use the tail bound obtained before).

Attempted Solution 2.5.10.

Exercise 2.5.11 *(Lower bound).* Show that the bound in Exercise 2.5.10 is sharp. Let $X_1, X_2, \ldots, X_N$ be independent $N(0, 1)$ random variables. Prove that

$$\mathbb{E} \max_{i \leq N} X_i \geq c\sqrt{\log N}.$$

Attempted Solution 2.5.11.

2.6 General Hoeffding's and Khintchine's inequalities

2.7 Sub-exponential distributions

2.8 Bernstein's inequality

3 Random vectors in high dimensions

3.1 Concentration of the norm

3.2 Covariance matrices and principal component analysis

3.3 Examples of high-dimensional distributions

3.4 Sub-gaussian distributions in higher dimensions

3.5 Application: Grothendieck's inequality and semidefinite programming

3.6 Application: Maximum cut for graphs

3.7 Kernel trick, and tightening of Grothendieck's inequality

4 Random matrices

4.1 Preliminaries on matrices

4.2 Nets, covering numbers and packing numbers

4.3 Application: error correcting codes

4.4 Upper bounds on random sub-gaussian matrices

4.5 Application: community detection in networks

4.6 Two-sided bounds on sub-gaussian matrices

4.7 Application: covariance estimation and clustering

5 Concentration without independence

5.1 Concentration of Lipschitz functions on the sphere

5.2 Concentration on other metric measure spaces

5.3 Application: Johnson-Lindenstrauss Lemma

5.4 Matrix Bernstein's inequality

5.5 Application: community detection in sparse networks

5.6 Application: covariance estimation for general distributions

6 Quadratic forms, symmetrization and contraction

6.1 Decoupling

6.2 Hanson-Wright Inequality

6.3 Concentration of anisotropic random vectors

6.4 Symmetrization

6.5 Random matrices with non-i.i.d. entries

6.6 Application: matrix completion

6.7 Contraction Principle

7 Random processes

7.1 Basic concepts and examples

7.2 Slepian's inequality

7.3 Sharp bounds on Gaussian matrices

7.4 Sudakov's minoration inequality

7.5 Gaussian width

7.6 Stable dimension, stable rank, and Gaussian complexity

7.7 Random projections of sets

8 Chaining

8.1 Dudley's inequality

8.2 Application: empirical processes

8.3 VC dimension

8.4 Application: statistical learning theory

8.5 Generic chaining

8.6 Talagrand's majorizing measure and comparison theorems

8.7 Chevet's inequality

9 Deviations of random matrices and geometric consequences

9.1 Matrix deviation inequality

9.2 Random matrices, random projections and covariance estimation

9.3 Johnson-Lindenstrauss Lemma for infinite sets

9.4 Random sections: M^* bound and Escape Theorem

10 Sparse Recovery

10.1 High-dimensional signal recovery problems

10.2 Signal recovery based on M^* bound

10.3 Recovery of sparse signals

10.4 Low-rank matrix recovery

10.5 Exact recovery and the restricted isometry property

10.6 Lasso algorithm for sparse regression

11 Dvoretzky-Milman's Theorem

11.1 Deviations of random matrices with respect to general norms

11.2 Johnson-Lindenstrauss embeddings and sharper Chevet inequality

11.3 Dvoretzky-Milman's Theorem