

High-Dimensional Statistics

Vitaliy Pozdnyakov

March 2020

Solutions to exercises from the book "High-Dimensional Probability", Roman Vershynin, 2020
<https://www.math.uci.edu/~rvershyn/>

Notation:

- cc – coffee cup
- $\sum := \sum_{i=1}^N$
- $\prod := \prod_{i=1}^N$

1 Exercise 2.1.4 (1 cc)

Let $g \sim \mathcal{N}(0, 1)$. Show that for all $t \geq 1$, we have

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

1.1 Solution

$$\mathbb{E} g^2 \mathbf{1}\{g > t\} = \int_{-\infty}^{+\infty} x^2 \mathbf{1}\{x > t\} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} x^2 e^{-x^2/2} dx$$

Make the substitution $u = x$ and $dv = x e^{-x^2/2} dx$, then $du = dx$ and

$$v = \int dv = \int x e^{-x^2/2} dx = \left[z = -\frac{x^2}{2}; dz = -x dx \right] = - \int e^z dz = -e^z = -e^{-x^2/2}$$

Next, integrate by parts

$$\int_a^b u dv = uv \Big|_a^b - \int_a^b v du$$

$$\frac{1}{\sqrt{2\pi}} \int_t^{+\infty} x^2 e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \left(-x e^{-x^2/2} \Big|_t^{+\infty} \right) + \int_t^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} + \mathbb{P}\{g > t\}$$

Next, apply the theorem "Tails of the normal distribution":

$$\mathbb{P}\{g > t\} \leq \frac{1}{t} \cdot \frac{1}{\sqrt{2\pi}} e^{t^2/2}$$

Thereby,

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

2 Exercise 2.2.3 (1 cc)

(Bounding the hyperbolic cosine)

Show that

$$\cosh(x) \leq \exp(x^2/2) \text{ for all } x \in \mathbb{R}.$$

2.1 Solution

Apply Taylor's expansion for the left side of inequation

$$\cosh(x)' = \sinh(x); \cosh(x)'' = \cosh(x); \cosh(x)''' = \sinh(x); \dots$$

$$\cosh(0) = 1; \sinh(0) = 0$$

Then

$$\begin{aligned} \cosh(x) &= \cosh(0) + \frac{\sinh(0)}{1!}x^1 + \frac{\cosh(0)}{2!}x^2 + \frac{\sinh(0)}{3!}x^3 + \dots \\ &= 1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + O(x^8) \end{aligned}$$

For the right side

$$\exp(x^2/2)' = \exp(x^2/2)x; \exp(x^2/2)'' = \exp(x^2/2)(x^2 + 1); \exp(x^2/2)''' = \exp(x^2/2)x(x^2 + 3); \dots$$

$$\exp(0^2/2) = 1; \exp(0^2/2)' = 0; \exp(0^2/2)'' = 1; \exp(0^2/2)''' = 0; \dots$$

Then

$$\begin{aligned} \exp(x^2/2) &= \exp(0^2/2) + \frac{\exp(0^2/2)'0}{1!}x^1 + \frac{\exp(0^2/2)''(0^2 + 1)}{2!}x^2 + \frac{\exp(0^2/2)'''(0^2 + 3)}{3!}x^3 + \dots \\ &= 1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{384} + O(x^8) \end{aligned}$$

Obviously,

$$1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + O(x^8) \leq 1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{384} + O(x^8)$$

Thus,

$$\cosh(x) \leq \exp(x^2/2) \text{ for all } x \in \mathbb{R}.$$

3 Exercise 2.2.7. (2 cc)

Let X_1, \dots, X_N be independent random variables. Assume that $X_i \in [m_i, M_i]$ for every i . Prove that for any $t > 0$

$$\mathbb{P} \left\{ \sum (X_i - \mathbb{E}X_i) \geq t \right\} \leq \exp \left(- \frac{2t^2}{\sum (M_i - m_i)^2} \right)$$

3.1 Solution

Let $\xi_i = X_i - \mathbb{E}X_i$. Apply Markov's inequality and a property of the exp function

$$\mathbb{P}\left\{\sum \xi_i \geq t\right\} = \mathbb{P}\left\{e^{\lambda \sum \xi_i} \geq e^{\lambda t}\right\} \leq e^{-\lambda t} \mathbb{E}\left[e^{\lambda \sum \xi_i}\right] \text{ for any } \lambda > 0$$

Apply independence assumption

$$\mathbb{E}\left[e^{\lambda \sum \xi_i}\right] = \mathbb{E}\left[\prod e^{\lambda \xi_i}\right] = \prod \mathbb{E}\left[e^{\lambda \xi_i}\right]$$

Apply Hoeffding's Lemma

$$\mathbb{E}\left[e^{\lambda \xi_i}\right] \leq e^{\frac{\lambda^2 (M_i - m_i)^2}{8}}$$

Thus,

$$\mathbb{P}\left\{\sum \xi_i \geq t\right\} \leq e^{-\lambda t} \prod e^{\frac{\lambda^2 (M_i - m_i)^2}{8}} = e^{-\lambda t + \frac{\lambda^2}{8} \sum (M_i - m_i)^2}$$

Let $\lambda = 4t / \sum (M_i - m_i)^2$, then

$$\mathbb{P}\left\{\sum (X_i - \mathbb{E}X_i) \geq t\right\} \leq \exp\left(-\frac{2t^2}{\sum (M_i - m_i)^2}\right)$$

4 Exercise 2.2.8. (2 cc)

(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 $1/2 + \delta$ 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 $\epsilon \in (0, 1)$, the answer is correct with probability at least $1 - \epsilon$, as long as

$$N \leq \frac{1}{2\delta^2} \ln(1/\epsilon)$$

4.1 Solution

Let X_1, X_2, \dots, X_N be a random sequence with probabilities

$$\mathbb{P}\{X_i = 0\} = 1/2 - \delta$$

$$\mathbb{P}\{X_i = 1\} = 1/2 + \delta$$

Then the event {The answer is incorrect} can be written as

$$\begin{aligned} \left\{\sum X_i \geq N/2\right\} &= \left\{\sum (X_i - \mathbb{E}X_i) \geq \sum (1/2 - \mathbb{E}X_i)\right\} \\ &= \left\{\sum (X_i - \mathbb{E}X_i) \geq \sum (1/2 - 1/2 + \delta)\right\} \\ &= \left\{\sum (X_i - \mathbb{E}X_i) \geq \delta N\right\} \end{aligned}$$

Apply Hoeffding inequality

$$\mathbb{P}\left\{\sum (X_i - \mathbb{E}X_i) \geq t\right\} \leq \exp\left(-\frac{2t^2}{\sum (M_i - m_i)^2}\right)$$

Make substitutions $t = \delta N$, $M_i = 1$, $m_i = 0$

$$\mathbb{P}\left\{\sum(X_i - \mathbb{E}X_i) \geq \delta N\right\} \leq \exp(-2\delta^2 N)$$

We can observe that

$$\left\{N \leq \frac{1}{2\delta^2} \ln(1/\epsilon)\right\} \Rightarrow \{\exp(-2\delta^2 N) < \epsilon\}$$

Thereby, $\mathbb{P}\{\text{The answer is incorrect}\} \leq \epsilon$ and $\mathbb{P}\{\text{The answer is correct}\} \geq 1 - \epsilon$ under that constraint.

5 Exercise 2.2.10 (2 cc)

Let X_1, \dots, 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 $\mathbb{E}e^{-tX_i} \leq 1/t$ for all $t > 0$.
(b) Deduce that, for any $\epsilon > 0$, we have $\mathbb{P}\{\sum X_i \leq \epsilon N\} \leq (e\epsilon)^N$.

5.1 Solution (a)

Let $\phi_i(x)$ be a PDF of X_i

$$\mathbb{E} \exp(-tX_i) = \int_0^\infty e^{-tx} \phi_i(x) dx$$

Since $\phi_i(x) \leq 1$ for all i and x , then

$$\int_0^\infty e^{-tx} \phi_i(x) dx \leq \int_0^\infty e^{-tx} dx = \frac{1}{t} \text{ for all } t > 0$$

5.2 Solution (b)

Rewrite the probability as

$$\mathbb{P}\left\{\sum X_i \leq \epsilon N\right\} = \mathbb{P}\left\{\sum -X_i/\epsilon \geq -N\right\}$$

Apply Markov's inequality and a property of the exp function

$$\mathbb{P}\left\{\sum -X_i/\epsilon \geq -N\right\} = \mathbb{P}\left\{\exp\left(\sum -X_i/\epsilon\right) \geq \exp(-N)\right\} \leq e^N \prod \mathbb{E}e^{-X_i/\epsilon}$$

Apply the result from the section (a) and then

$$\mathbb{P}\left\{\sum X_i \leq \epsilon N\right\} \leq e^N \prod \epsilon = (e\epsilon)^N$$

6 Exercise 2.3.2 (2 cc)

(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

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

6.1 Solution

Since $t < \mu$, $t > 0$ and $S_N > 0$ we can rewrite the probability as

$$\mathbb{P}\{S_N \leq t\} = \mathbb{P}\{(t/\mu)^{S_N} \geq (t/\mu)^t\}$$

Apply Markov's inequality

$$\mathbb{P}\{(t/\mu)^{S_N} \geq (t/\mu)^t\} \leq (\mu/t)^t \mathbb{E}(t/\mu)^{S_N}$$

Consider the expectation

$$\mathbb{E}(t/\mu)^{S_N} = \prod \mathbb{E}(t/\mu)^{X_i}$$

$$\mathbb{E}(t/\mu)^{X_i} = (t/\mu)^1 p_i + (t/\mu)^0 (1 - p_i) = 1 + [(t/\mu) - 1] p_i \leq \exp\left(\left[\frac{t}{\mu} - 1\right] p_i\right)$$

$$\prod \mathbb{E}(t/\mu)^{X_i} \leq \exp\left(\left[\frac{t}{\mu} - 1\right] \sum p_i\right) = \exp\left(\left[\frac{t}{\mu} - 1\right] \mu\right) = e^t e^{-\mu}$$

Thereby

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

7 Exercise 2.3.3 (2 cc)

(Poisson tails)

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

$$\mathbb{P}\{X \geq t\} \leq e^{-\lambda} \left(\frac{e\lambda}{t}\right)^t$$

7.1 Solution

Chernoff's inequality

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

Assume that $N \rightarrow \infty$, $\max p_i \rightarrow 0$ and $\mathbb{E}S_N \rightarrow \lambda < \infty$. Then, using Poisson limit theorem we have

$$S_N \rightarrow \text{Pois}(\lambda)$$

Thereby

$$\mathbb{P}\{X \geq t\} \leq e^{-\lambda} \left(\frac{e\lambda}{t}\right)^t$$

8 Exercise 2.3.5 (3 cc)

(Chernoff's inequality: small deviations)

Show that, in the setting of Theorem 2.3.1, for $\delta \in (0, 1]$ we have

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq 2e^{-c\mu\delta^2}$$

where $c > 0$ is an absolute constant.

8.1 Solution

Consider Chernoff's inequality with $t > \mu$ (upper tails). Let $t = \mu + \delta\mu$, so

$$\mathbb{P}\{S_N - \mu \geq \delta\mu\} \leq e^{-\mu} \left(\frac{e}{1+\delta} \right)^{\mu+\delta\mu}$$

In opposite, if $t < \mu$ and $t = \mu - \delta\mu$ (lower tails) we obtain

$$\mathbb{P}\{S_N - \mu \leq -\delta\mu\} \leq e^{-\mu} \left(\frac{e}{1-\delta} \right)^{\mu-\delta\mu}$$

Apply addition theorem of probability for lower and upper tails

$$\mathbb{P}(\{S_N - \mu \geq \delta\mu\} \cup \{S_N - \mu \leq -\delta\mu\}) = \mathbb{P}\{|S_N - \mu| \geq \delta\mu\}$$

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq e^{-\mu} \left(\frac{e}{1-\delta} \right)^{\mu-\delta\mu} + e^{-\mu} \left(\frac{e}{1+\delta} \right)^{\mu+\delta\mu} \leq 2e^{-\mu} \left(\frac{e}{1+\delta} \right)^{\mu+\delta\mu}$$

Consider the inequality

$$\begin{aligned} 2e^{-\mu} \left(\frac{e}{1+\delta} \right)^{\mu+\delta\mu} &\leq 2e^{-c\mu\delta^2} \\ \frac{e^{\delta\mu}}{(1+\delta)^{\mu+\delta\mu}} &\leq e^{-c\mu\delta^2} \\ e^{\delta+c\delta^2} &\leq (1+\delta)^{1+\delta} \end{aligned}$$

Take log

$$\begin{aligned} \delta + c\delta^2 &\leq (1+\delta) \ln(1+\delta) \\ c &\leq \frac{(1+\delta) \ln(1+\delta) - \delta}{\delta^2} \end{aligned}$$

Thereby,

$$\mathbb{P}\{|S_N - \mu| \geq \delta\mu\} \leq 2e^{-c\mu\delta^2}$$

for all $c \leq [(1+\delta) \ln(1+\delta) - \delta]/\delta^2$.

9 Exercise 2.3.8 (2 cc)

(Normal approximation to Poisson)

Let $X \sim Pois(\lambda)$. Show that, as $\lambda \rightarrow \infty$, we have

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

9.1 Solution

Rewrite the RV X as a sum $Z_1 + \dots + Z_N$ where $Z_i \sim Pois(\lambda/N)$ and apply Lindeberg-Levy central limit theorem

$$\frac{X - \mathbb{E}X}{\sqrt{\text{Var}(X)}} = \frac{X - N\mathbb{E}Z_i}{\sqrt{N\text{Var}(Z_i)}} = \frac{X - \lambda}{\sqrt{\lambda}} \rightarrow \mathcal{N}(0, 1) \text{ as } N \rightarrow \infty$$

10 Exercise 2.4.2 (1 cc)

(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 $O(\log n)$.

10.1 Solution

Since $d_i \geq 0$ and $n > 0$, then for any i we have $d_i = O(\log n) \equiv \exists C_i : d_i \leq C_i \log n$ and we need to prove the following statement

$$\mathbb{P}\{\forall i : d_i \leq C_i \log n\} \geq 0.9$$

where C_i is a constant. We can rewrite probability of the opposite case as

$$\mathbb{P}\{\exists i : d_i \geq C_i \log n\} \leq \sum \mathbb{P}\{d_i \geq C_i \log n\} \leq 0.1$$

Apply Chernoff's inequality for upper bound (since $d \leq C_i \log n$)

$$\sum \mathbb{P}\{d_i \geq C_i \log n\} \leq ne^{-d} \left(\frac{ed}{C_i \log n} \right)^{C_i \log n} \leq 0.1$$

Note that there exists C_d such that $d \leq C_d \log n$ and then

$$ne^{-C_d \log n} \left(\frac{eC_d \log n}{C_i \log n} \right)^{C_i \log n} = n^{C_i - C_d + 1} \left(\frac{C_d}{C_i} \right)^{C_i \log n} \leq 0.1$$

Hence, the statement holds for any C_i such that

$$C_i \geq \frac{-C_d \log n + \log n + \log 10}{\log n} \left[W \left(\frac{-C_d \log n + \log n + \log 10}{C_d e \log n} \right) \right]^{-1}$$

where $W(\cdot)$ is the product log function.

11 Exercise 2.4.3 (2 cc)

(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

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

11.1 Solution

Apply the same approach as in the previous exercise and obtain

$$\sum \mathbb{P}\left\{d_i \geq C_i \frac{\log n}{\log \log n}\right\} \leq ne^{-C_d} \left(\frac{eC_d \log \log n}{C_i \log n} \right)^{(C_i \log n)/(\log \log n)} \leq 0.1$$

Hence, the statement holds for any C_i such that

$$C_i \geq \frac{(\log n - C_d + \log 10) \log \log n}{\log n} \left[W \left(\frac{\log n - C_d + \log 10}{C_d e} \right) \right]^{-1}$$

where $W(\cdot)$ is the product log function.

12 Exercise 3.2.6 (1 cc)

(Distance between independent isotropic vectors)

Let X and Y be independent, mean zero, isotropic random vectors in \mathbb{R}^n . Check that

$$\mathbb{E}\|X - Y\|_2^2 = 2n$$

12.1 Solution

$$\mathbb{E}\|X - Y\|_2^2 = \mathbb{E}\left[\sum (X_i - Y_i)^2\right] = \sum (\mathbb{E}X_i^2 - \mathbb{E}X_iY_i + \mathbb{E}Y_i^2)$$

$\mathbb{E}X_i^2 = \mathbb{E}Y_i^2 = 1$ by isotropy, $\mathbb{E}X_iY_i = \mathbb{E}X_i\mathbb{E}Y_i$ by independence and $\mathbb{E}X_i = \mathbb{E}Y_i = 0$ by task description. Thereby,

$$\mathbb{E}\|X - Y\|_2^2 = \sum (1 - 0 + 1) = 2n$$

13 Exercise 3.3.1 (1 cc)

Show that the spherically distributed random vector X is isotropic. Argue that the coordinates of X are not independent.

13.1 Solution

14 Exercise 3.3.3 (2 cc)

(Rotation invariance)

Deduce the following properties from the rotation invariance of the normal distribution.

- (a) Consider a random vector $g \sim \mathcal{N}(0, I_n)$ and a fixed vector $u \in \mathbb{R}^n$. Then $\langle g, u \rangle \sim \mathcal{N}(0, \|u\|_2^2)$
- (b) Consider independent random variables $X_i \sim \mathcal{N}(0, \sigma_i^2)$. Then $\sum X_i \sim \mathcal{N}(0, \sigma^2)$ where $\sigma^2 = \sum \sigma_i^2$
- (c) Let G be an $m \times n$ Gaussian random matrix, i.e. the entries of G are independent $\mathcal{N}(0, 1)$ random variables. Let $u \in \mathbb{R}^n$ be a fixed unit vector. Then $Gu \sim \mathcal{N}(0, I_m)$.

14.1 Solution (a)

Note that $\langle g, u \rangle = \sum g_i u_i$ where $g_i u_i \sim \mathcal{N}(0, u_i^2)$ by summation property of Gaussian distribution and then

$$\mathbb{E}\langle g, u \rangle = \sum \mathbb{E}g_i u_i = 0$$

$$\text{Var}\langle g, u \rangle = \sum \text{Var} g_i u_i = \sum u_i^2 = \|u\|_2^2$$

Thereby,

$$\langle g, u \rangle \sim \mathcal{N}\left(\sum \mathbb{E}g_i u_i, \sum \text{Var} g_i u_i\right) = \mathcal{N}(0, \|u\|_2^2)$$

14.2 Solution (b)

Consider the PDF $f_Z(z)$ where $Z = X + Y$ with independent $X \sim \mathcal{N}(0, \sigma_X^2)$ and $Y \sim \mathcal{N}(0, \sigma_Y^2)$

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f(z-x)f(x)dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sigma_Y \sqrt{2\pi}} \exp\left[-\frac{(z-x)^2}{2\sigma_Y^2}\right] \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left[-\frac{x^2}{2\sigma_X^2}\right] dx \\ &= [\dots] = \frac{1}{\sqrt{\sigma_X^2 + \sigma_Y^2} \sqrt{2\pi}} \exp\left[-\frac{z^2}{2(\sigma_X^2 + \sigma_Y^2)}\right] \end{aligned}$$

Thereby $Z \sim \mathcal{N}(0, \sigma_X^2 + \sigma_Y^2)$ and consequently $\sum X_i \sim \mathcal{N}(0, \sigma^2)$ where $\sigma^2 = \sum \sigma_i^2$

14.3 Solution (c)

Denote g_{ij} as an entry of the matrix G and v_i as an entry of the vector Gu

$$\begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix} := Gu = \begin{pmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \dots & \dots & \dots & \dots \\ g_{m1} & g_{m2} & \dots & g_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

Then $v_i = g_{i1}u_1 + g_{i2}u_2 + \dots + g_{in}u_n \sim \mathcal{N}(0, \|u\|_2^2)$ by solution (a) and (b). Since u is a unit vector we obtain that $v_i \sim \mathcal{N}(0, 1)$. Entries g_{ij} are independent, then entries v_i are independent too and consequently $\forall i \neq j : \text{cov}(v_i, v_j) = 0$. Thereby $Gu \sim \mathcal{N}(0, I_n)$.

15 Exercise 4.4.2 (1 cc)

Let $x \in \mathbb{R}^n$ and \mathcal{N} be an ε -net of the sphere S^{n-1} . Show that

$$\sup_{y \in \mathcal{N}} \langle x, y \rangle \leq \|x\|_2 \leq \frac{1}{1-\varepsilon} \sup_{y \in \mathcal{N}} \langle x, y \rangle$$

15.1 Solution

First, consider the lower bound. By the definition of the dot product we have

$$\langle x, y \rangle = \|x\|_2 \|y\|_2 \cos \theta$$

Use the fact that $y \in \mathcal{N} \subset S^{n-1}$, so $\|y\|_2 = 1$. Also $\cos \theta \leq 1$, consequently

$$\sup_{y \in \mathcal{N}} \langle x, y \rangle \leq \sup_{y \in S^{n-1}} \langle x, y \rangle = \|x\|_2$$

Next, use the similar approach as in Lemma 4.4.1 for the upper bound. Fix a vector $y \in S^{n-1}$ for which

$$\langle x, y \rangle = \|x\|_2$$

and choose $y_0 \in \mathcal{N}$ that approximates y so that

$$\|y - y_0\|_2 \leq \varepsilon$$

By the definition of the dot product, this implies

$$\langle x, y - y_0 \rangle = \|x\|_2 \|y - y_0\|_2 \cos \theta \leq \|x\|_2 \varepsilon \cos \theta \leq \|x\|_2 \varepsilon$$

Using a distributive law, we find that

$$\langle x, y_0 \rangle = \langle x, y \rangle - \langle x, y - y_0 \rangle \geq \|x\|_2 - \|x\|_2 \varepsilon = (1 - \varepsilon) \|x\|_2$$

Dividing both sides of this inequality by $1 - \varepsilon$, we complete the proof.

16 Exercise 4.4.3 (2 cc)

(Quadratic form on a net)

Let A be an $m \times n$ matrix and $\varepsilon \in [0, 1/2)$.

(a) Show that for any ε -net \mathcal{N} of the sphere S^{n-1} and any ε -net \mathcal{M} of the sphere S^{m-1} , we have

$$\sup_{x \in \mathcal{N}, y \in \mathcal{M}} \langle Ax, y \rangle \leq \|A\| \leq \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}, y \in \mathcal{M}} \langle Ax, y \rangle$$

(b) Moreover, if $m = n$ and A is symmetric, show that

$$\sup_{x \in \mathcal{N}} |\langle Ax, x \rangle| \leq \|A\| \leq \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}} |\langle Ax, x \rangle|$$

16.1 Solution (a)

First, consider the lower bound. Using the same explanations as in Exercise 4.4.2 and the lower bound from Lemma 4.4.1 we obtain

$$\|A\| \geq \sup_{x \in \mathcal{N}} \|Ax\|_2 \geq \sup_{x \in \mathcal{N}, y \in \mathcal{M}} \langle Ax, y \rangle$$

Next, again apply Lemma 4.4.1 and Exercise 4.4.2 for the upper bound

$$\begin{aligned} \|A\| &\leq \frac{1}{1 - \varepsilon} \sup_{x \in \mathcal{N}} \|Ax\|_2 \\ &\leq \frac{1}{1 - \varepsilon} \sup_{x \in \mathcal{N}} \left(\frac{1}{1 - \varepsilon} \sup_{y \in \mathcal{M}} \langle Ax, y \rangle \right) = \frac{1}{1 - 2\varepsilon + \varepsilon^2} \sup_{x \in \mathcal{N}, y \in \mathcal{M}} \langle Ax, y \rangle \\ &\leq \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}, y \in \mathcal{M}} \langle Ax, y \rangle \quad (\text{since } \varepsilon \in [0, 1/2)) \end{aligned}$$

16.2 Solution (b)

First, use Lemma 4.4.1 and the Cauchy–Bunyakovsky–Schwarz inequality for the lower bound

$$\begin{aligned} \|A\| &\geq \sup_{x \in \mathcal{N}} \|Ax\|_2 = \sup_{x \in \mathcal{N}} \|Ax\|_2 \|x\|_2 && (\text{by the fact that } \|x\|_2 = 1) \\ &\geq \sup_{x \in \mathcal{N}} |\langle Ax, x \rangle| && (\text{by Cauchy–Bunyakovsky–Schwarz inequality}) \end{aligned}$$

Next, use the similar approach as in Lemma 4.4.1. Fix a vector $x \in S^{n-1}$ for which

$$\|Ax\|_2 = \|Ax\|_2 \|x\|_2 = |\langle Ax, x \rangle|$$

and choose $x_0 \in \mathcal{N}$ that approximates x so that

$$\|x - x_0\|_2 \leq \varepsilon$$

By the Cauchy–Bunyakovsky–Schwarz inequality, this implies

$$|\langle Ax, x - x_0 \rangle| \leq \|Ax\|_2 \|x - x_0\|_2 \leq \|Ax\|_2 \varepsilon$$

Using the triangle inequality of absolute value, we find that

$$|\langle Ax, x_0 \rangle| \geq |\langle Ax, x \rangle| - |\langle Ax, x - x_0 \rangle| \geq \|Ax\|_2 - \|Ax\|_2 \varepsilon = (1 - \varepsilon) \|Ax\|_2$$

Dividing both sides of this inequality by $1 - \varepsilon$, we prove that

$$\|Ax\|_2 \leq \frac{1}{1 - \varepsilon} \sup_{x \in \mathcal{N}} |\langle Ax, x \rangle|$$

Finally, apply the upper bound of Lemma 4.4.1 and complete the proof

$$\begin{aligned} \|A\| &\leq \frac{1}{1 - \varepsilon} \sup_{x \in \mathcal{N}} \|Ax\|_2 \\ &\leq \frac{1}{1 - \varepsilon} \sup_{x \in \mathcal{N}} \left(\frac{1}{1 - \varepsilon} \sup_{x \in \mathcal{N}} |\langle Ax, x \rangle| \right) = \frac{1}{1 - 2\varepsilon + \varepsilon^2} \sup_{x \in \mathcal{N}} |\langle Ax, x \rangle| \\ &\leq \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}} |\langle Ax, x \rangle| \quad (\text{since } \varepsilon \in [0, 1/2)) \end{aligned}$$

17 Exercise 4.4.4 (3 cc)

(Deviation of the norm on a net)

Let A be an $m \times n$ matrix, $\mu \in \mathbb{R}$ and $\varepsilon \in [0, 1/2)$. Show that for any ε -net \mathcal{N} of the sphere S^{n-1} , we have

$$\sup_{x \in S^{n-1}} |\|Ax\|_2 - \mu| \leq \frac{C}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}} |\|Ax\|_2 - \mu|$$

17.1 Solution

Assume without loss of generality that $\mu = 1$. Rewrite $\|Ax\|_2^2 - 1$ in the quadratic form

$$\|Ax\|_2^2 - 1 = \langle Rx, x \rangle$$

where $R = A^T A - I_n$. Next, use the definition of the operator norm and the result from Exercise 4.4.3 (b)

$$\sup_{x \in S^{n-1}} |\langle Rx, x \rangle| = \|R\| \leq \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}} |\langle Rx, x \rangle|$$

Applying the elementary inequality

$$\max(|z - 1|, |z - 1|^2) \leq |z^2 - 1|, z > 0$$

for $z = \|Ax\|_2$, we conclude that

$$|\langle Rx, x \rangle| \leq |\|Ax\|_2 - 1|$$

And thereby

$$\sup_{x \in S^{n-1}} |\|Ax\|_2 - 1| \leq \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}} |\|Ax\|_2 - 1|$$

Reject the assumption $\mu = 1$ and complete the proof.