

# 1 Entropy method beyond uniform gradient bounds

Throughout, let  $X \in \mathbb{R}^n$  be a random vector satisfying the MLSI

$$\text{Ent}(e^{u(X)}) \leq C \mathbb{E}[\|\nabla u(X)\|_2^2 e^{u(X)}] \quad \text{for all smooth } u : \mathbb{R}^n \rightarrow \mathbb{R}, \tag{1}$$

where  $C > 0$  is a fixed constant. For a nonnegative random variable  $Y$ , recall the concentration entropy

$$\text{Ent}(Y) := \mathbb{E}[Y \log Y] - (\mathbb{E}Y) \log(\mathbb{E}Y).$$

For a real random variable  $W$ , define its log-MGF

$$\kappa_W(\eta) := \log \mathbb{E}e^{\eta W},$$

whenever this is finite. For centered quantities, we follow Lecture 7 and write

$$\psi_Z(\theta) := \log \mathbb{E}e^{\theta Z} \quad \text{when } \mathbb{E}Z = 0.$$

You may use the following fact from Lecture 7 without proof: if  $Z$  is centered, then

$$\frac{\text{Ent}(e^{\theta Z})}{\mathbb{E}e^{\theta Z}} = \theta \psi'_Z(\theta) - \psi_Z(\theta) = \theta^2 \left( \frac{\psi_Z(\theta)}{\theta} \right)'. \tag{2}$$

In particular, an upper bound on

$$\frac{\text{Ent}(e^{\theta Z})}{\theta^2 \mathbb{E}e^{\theta Z}}$$

can be integrated using Herbst's argument to bound  $\psi_Z(\theta)$ .

- (a) **Young's inequality for entropy.** Let  $Y \geq 0$  satisfy  $\mathbb{E}Y = 1$ , and let  $W$  be any real-valued random variable. Show that

$$\mathbb{E}[WY] \leq \log \mathbb{E}e^W + \text{Ent}(Y). \tag{3}$$

Equivalently, in our notation:

$$\mathbb{E}[WY] \leq \kappa_W(1) + \text{Ent}(Y).$$

Identify the equality condition.

**Solution:**

Let  $(\Omega, \mathcal{F}, \phi)$  be a probability space, and both  $Y$  and  $Z := \frac{e^W}{\mathbb{E}e^W}$  are non-negative random variables on that space. This gives,  $\mathbb{E}_\phi[Y] = 1$  and  $\mathbb{E}_\phi[Z] = 1$ , so we define two new probability measures

$$d\mu = Y d\phi \quad d\nu = Z d\phi.$$

Now, the KL divergence between the two measures is given by,

$$D_{\text{KL}}(\mu \parallel \nu) = \mathbb{E}_\phi \left[ Y \log \left( \frac{Y}{Z} \right) \right] = \mathbb{E}_\phi [Y (\log Y - W + \log \mathbb{E}_\phi e^W)].$$

Now, using the non-negative property of the KL divergence, we get

$$\log \mathbb{E}_\phi e^W + \mathbb{E}_\phi [Y \log Y] \geq \mathbb{E}_\phi [WY].$$

Here,  $\text{Ent}(Y) = \mathbb{E}_\phi [Y \log Y] - (\mathbb{E}_\phi Y) \log(\mathbb{E}_\phi Y) = \mathbb{E}_\phi [Y \log Y]$  because  $\mathbb{E}_\phi Y = 1$ . Hence,

$$\log \mathbb{E}_\phi e^W + \text{Ent}(Y) \geq \mathbb{E}_\phi [WY]. \quad \square$$

Here, the equality holds iff the KL is zero, i.e., iff  $Y = \frac{e^W}{\mathbb{E}e^W}$  almost surely.  $\square$

- (b) **A nonlinear Bernstein inequality from MLSI.** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be smooth, and define the centered random variable

$$Z := f(X) - \mathbb{E}f(X), \quad \psi_f(\theta) := \log \mathbb{E}e^{\theta Z}, \quad G := \|\nabla f(X)\|_2^2.$$

- (b1) Show that for every  $\theta \geq 0$ ,

$$\text{Ent}(e^{\theta Z}) \leq C\theta^2 \mathbb{E}[G e^{\theta Z}].$$

**Solution:**

Let  $u(X) = \theta Z$ , then  $\|\nabla u(X)\|_2^2 = \theta^2 \|\nabla f(X)\|_2^2$ . As  $X$  satisfies MLSI, using (1), we get

$$\text{Ent}(e^{\theta Z}) \leq C\mathbb{E}\left[\|\nabla u(X)\|_2^2 e^{u(X)}\right] = C\theta^2 \mathbb{E}[G e^{\theta Z}]. \quad \square$$

- (b2) Fix  $\eta > 0$ . Apply part (a) with

$$Y = \frac{e^{\theta Z}}{\mathbb{E}e^{\theta Z}}, \quad W = \eta G,$$

and show that

$$\frac{\mathbb{E}[G e^{\theta Z}]}{\mathbb{E}e^{\theta Z}} \leq \frac{\kappa_G(\eta)}{\eta} + \frac{\text{Ent}(e^{\theta Z})}{\eta \mathbb{E}e^{\theta Z}}.$$

**Solution:**

Here,  $\mathbb{E}[WY] = \frac{\eta \mathbb{E}[G e^{\theta Z}]}{\mathbb{E}e^{\theta Z}}$  and  $\log \mathbb{E}e^W = \log \mathbb{E}e^{\eta G} = \kappa_G(\eta)$ . Similarly, using homogeneity  $\text{Ent}(Y) = \frac{\text{Ent}(e^{\theta Z})}{\mathbb{E}e^{\theta Z}}$ . Thus, using (3) from part (a),

$$\frac{\eta \mathbb{E}[G e^{\theta Z}]}{\mathbb{E}e^{\theta Z}} \leq \kappa_G(\eta) + \frac{\text{Ent}(e^{\theta Z})}{\mathbb{E}e^{\theta Z}}.$$

Rearranging, we get,

$$\frac{\mathbb{E}[G e^{\theta Z}]}{\mathbb{E}e^{\theta Z}} \leq \frac{\kappa_G(\eta)}{\eta} + \frac{\text{Ent}(e^{\theta Z})}{\eta \mathbb{E}e^{\theta Z}}. \quad \square$$

- (b3) Combining parts (b1) and (b2), deduce that whenever  $\eta > C\theta^2$ ,

$$\frac{\text{Ent}(e^{\theta Z})}{\theta^2 \mathbb{E}e^{\theta Z}} \leq \frac{C \kappa_G(\eta)}{\eta - C\theta^2}.$$

**Solution:**

Combining parts (b1) and (b2), we get,

$$\frac{\mathbb{E}[G e^{\theta Z}]}{\mathbb{E}e^{\theta Z}} \leq \frac{\kappa_G(\eta)}{\eta} + \frac{C\theta^2 \mathbb{E}[G e^{\theta Z}]}{\eta \mathbb{E}e^{\theta Z}} \implies (\eta - C\theta^2) \mathbb{E}[G e^{\theta Z}] \leq \kappa_G(\eta) \mathbb{E}e^{\theta Z}.$$

Rearranging, we get,  $\frac{\mathbb{E}[G e^{\theta Z}]}{\mathbb{E}e^{\theta Z}} \leq \frac{\kappa_G(\eta)}{\eta - C\theta^2}$ , which is only valid if  $\eta > C\theta^2$ . Using part (b1) again, to lower bound  $\mathbb{E}[G e^{\theta Z}]$  gives,

$$\frac{\text{Ent}(e^{\theta Z})}{\theta^2 \mathbb{E}e^{\theta Z}} \leq \frac{C \kappa_G(\eta)}{\eta - C\theta^2}. \quad \square$$

(b4) Use the Herbst identity (2) to show that whenever  $\eta > C\theta^2$ ,

$$\psi_f(\theta) \leq \frac{C\theta^2 \kappa_G(\eta)}{\eta - C\theta^2}.$$

**Solution:**

Using Herbst's identity (2) and part (b3),

$$\left(\frac{\psi_f(\theta)}{\theta}\right)' \leq \frac{C \kappa_G(\eta)}{\eta - C\theta^2}.$$

Integrating from 0 to  $\theta$  gives,

$$\frac{\psi_f(\theta)}{\theta} - \lim_{t \rightarrow 0} \frac{\psi_f(t)}{t} \leq \int_0^\theta \frac{C \kappa_G(\eta)}{\eta - Ct^2} dt \stackrel{(i)}{\leq} \int_0^\theta \frac{C \kappa_G(\eta)}{\eta - C\theta^2} dt = \frac{C\theta \kappa_G(\eta)}{\eta - C\theta^2},$$

where the limit is zero because  $\psi_f(0) = 0$  and  $\psi_f'(0) = \mathbb{E}Z = 0$ , and (i) follows from the fact that for all  $t \in [0, \theta]$ ,  $\eta - C\theta^2 \leq \eta - Ct^2$ . Hence, we get,

$$\psi_f(\theta) \leq \frac{C\theta^2 \kappa_G(\eta)}{\eta - C\theta^2}. \quad \square$$

(b5) Deduce the one-sided tail bound

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \inf_{\eta > 0} \inf_{0 \leq \theta < \sqrt{\eta/C}} \exp\left(-\theta t + \frac{C\theta^2 \kappa_G(\eta)}{\eta - C\theta^2}\right).$$

This is a nonlinear Bernstein inequality: the tail of  $f(X) - \mathbb{E}f(X)$  is controlled by the exponential moments of the random energy  $G = \|\nabla f(X)\|_2^2$ .

**Solution:**

Fix  $\eta > 0$  and  $0 \leq \theta < \sqrt{\eta/C}$ , so that  $\eta > C\theta^2$ . Using the Chernoff bound:

$$\mathbb{P}\{Z \geq t\} = \mathbb{P}\{e^{\theta Z} \geq e^{\theta t}\} \leq e^{-\theta t} \mathbb{E}e^{\theta Z} = \exp(-\theta t + \psi_f(\theta)).$$

Using the bound for  $\psi_f(\theta)$  from part (b4), we get

$$\mathbb{P}\{Z \geq t\} \leq \exp\left(-\theta t + \frac{C\theta^2 \kappa_G(\eta)}{\eta - C\theta^2}\right).$$

Since this holds for every  $\eta > 0$  and every  $0 \leq \theta < \sqrt{\eta/C}$ , taking the infimum over both parameters yields the one-sided nonlinear Bernstein inequality:

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \inf_{\eta > 0} \inf_{0 \leq \theta < \sqrt{\eta/C}} \exp\left(-\theta t + \frac{C\theta^2 \kappa_G(\eta)}{\eta - C\theta^2}\right). \quad \square$$

(c) **[Bonus] Application: self-bounded functions.** Assume now that the centered random variable  $Z = f(X) - \mathbb{E}f(X)$  is *self-bounded* in the sense that

$$G = \|\nabla f(X)\|_2^2 \leq aZ + b \quad \text{almost surely,}$$

for some constants  $a, b \geq 0$ .

(c1) Show that for every  $\eta \geq 0$ ,

$$\kappa_G(\eta) \leq b\eta + \psi_f(a\eta).$$

**[Bonus] Solution:**

Using the definition of log-MGF:

$$\kappa_G(\eta) = \log \mathbb{E} e^{\eta G} \stackrel{(i)}{\leq} \log \mathbb{E} e^{\eta(aZ+b)} = \log \mathbb{E} e^{a\eta Z} + \log(e^{b\eta}) = b\eta + \psi_f(a\eta),$$

where (i) uses the fact that  $G$  is upper bounded by  $aZ + b$  almost surely and  $\eta \geq 0$ .  $\square$

(c2) Assume  $a > 0$ . In part (b4), choose  $\eta = \theta/a$  and prove that for every

$$0 \leq \theta < \frac{1}{2Ca},$$

we have

$$\psi_f(\theta) \leq \frac{Cb\theta^2}{1-2Ca\theta}.$$

**[Bonus] Solution:**

Plugging  $\eta = \theta/a$  in the result from part (b4), we get

$$\psi_f(\theta) \leq \frac{C\theta^2 \kappa_G(\eta)}{\theta/a - C\theta^2} = \frac{aC\theta^2 \kappa_G(\eta)}{\theta - aC\theta^2}.$$

Using the bound for  $\kappa_G(\eta)$  from part (c1) gives,

$$\psi_f(\theta) \leq \frac{aC\theta^2 (b\eta + \psi_f(a\eta))}{\theta - aC\theta^2} \implies \psi_f(\theta) [\theta - 2aC\theta^2] \leq abC\theta^2\eta.$$

Now, we can divide on both sides by  $[\theta - 2aC\theta^2]$  without flipping the inequality only if  $[\theta - 2aC\theta^2] > 0 \implies 0 \leq \theta < \frac{1}{2Ca}$ . This gives,

$$\psi_f(\theta) \leq \frac{abC\theta^2\eta}{\theta - 2aC\theta^2} = \frac{bC\theta(a\eta)}{1 - 2aC\theta} = \frac{Cb\theta^2}{1 - 2Ca\theta}. \quad \square$$

(c3) Deduce that there exists a constant  $c = c(C) > 0$  such that

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \exp\left(-c \min\left\{\frac{t^2}{b}, \frac{t}{a}\right\}\right) \quad \text{for all } t \geq 0.$$

Briefly explain what happens in the simpler case  $a = 0$ .

**[Bonus] Solution:**

Using the bound for  $\psi_f(\theta)$  from part (c2) and Chernoff gives,

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \inf_{0 \leq \theta < 1/(2Ca)} \exp\left(-\theta t + \frac{Cb\theta^2}{1 - 2Ca\theta}\right).$$

Now, restrict  $\theta \leq \frac{1}{4Ca}$  such that  $\frac{Cb\theta^2}{1 - 2Ca\theta} \leq 2Cb\theta^2$ . Hence, for all  $0 \leq \theta \leq \frac{1}{4Ca}$ ,

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \exp(-\theta t + 2Cb\theta^2).$$

Define  $\theta = \min \left\{ \frac{t}{4Cb}, \frac{1}{4Ca} \right\} \leq \frac{1}{4Ca}$ , then we will have two cases:

1. If  $\frac{t}{4Cb} \leq \frac{1}{4Ca}$  then,  $-\theta t + 2Cb\theta^2 = -\frac{t^2}{8Cb}$ , which gives

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \exp\left(-\frac{t^2}{8Cb}\right).$$

2. If  $\frac{1}{4Ca} < \frac{t}{4Cb}$  equivalently  $\frac{b}{a} < t$  then,  $-\theta t + 2Cb\theta^2 = -\frac{t}{4Ca} + \frac{b}{8Ca^2} < -\frac{t}{4Ca} + \frac{t}{8Ca} = -\frac{t}{8Ca}$ , which gives

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \exp\left(-\frac{t}{8Ca}\right).$$

Hence, combining both cases, we get,

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \exp\left(-\frac{1}{8C} \min\left\{\frac{t^2}{b}, \frac{t}{a}\right\}\right) \quad \text{for all } t \geq 0. \quad \square$$

If  $a = 0$ , then the self-bounded assumption becomes  $G \leq b$  a.s. So, the bound in (c1) becomes,

$$\kappa_G(\eta) \leq b\eta.$$

Following a similar derivation above removes the linear-tail regime, and we obtain:

$$\mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} \leq \exp\left(-c\frac{t^2}{b}\right) \quad \text{for all } t \geq 0. \quad \square$$

(d) **[Bonus] Application: positive semidefinite quadratic forms.** This part gives a general-distribution analogue of the Gaussian quadratic-form concentration bounds you proved in Homework 3, Exercise 5 (and Mid Exam, Exercise 2).

Let  $A \in \mathbb{R}^{n \times n}$  be positive semidefinite, and let  $X \in \mathbb{R}^n$  be isotropic:

$$\mathbb{E}[XX^\top] = I_n.$$

Define

$$f(X) := X^\top AX - \text{tr}(A).$$

(d1) Show that  $\mathbb{E}f(X) = 0$ , and prove that

$$\|\nabla f(X)\|_2^2 = 4X^\top A^2 X \leq 4\|A\|_{\text{op}} X^\top AX = 4\|A\|_{\text{op}} (f(X) + \text{tr}(A)).$$

Thus  $f(X)$  is self-bounded with

$$a = 4\|A\|_{\text{op}}, \quad b = 4\|A\|_{\text{op}} \text{tr}(A).$$

**[Bonus] Solution:**

$$\mathbb{E}[X^\top AX] \stackrel{(i)}{=} \mathbb{E}[\text{tr}(X^\top AX)] \stackrel{(ii)}{=} \mathbb{E}[\text{tr}(AXX^\top)] \stackrel{(iii)}{=} \text{tr}(A\mathbb{E}[XX^\top]) \stackrel{(iv)}{=} \text{tr}(A),$$

where (i) uses the trace trick for scalar, (ii) uses cyclic property of trace, (iii) uses the linearity of trace operator, and (iv) uses the given fact that  $\mathbb{E}[XX^\top] = I_n$ . Hence,  $\mathbb{E}f(X) = 0$ .  $\square$

Now,

$$\nabla f(X) = (A + A^\top)X \stackrel{(i)}{=} 2AX \implies \|\nabla f(X)\|_2^2 = 4X^\top A^\top AX = 4X^\top A^2 X,$$

where (i) is true because  $A$  is symmetric (being positive semidefinite).  $A$  being positive semidefinite, its eigenvalues are non-negative, so on the spectral law  $A^2 \leq \|A\|_{\text{op}} A$ . Hence,

$$\|\nabla f(X)\|_2^2 \leq 4 \|A\|_{\text{op}} X^\top AX = 4 \|A\|_{\text{op}} (f(X) + \text{tr}(A)). \quad \square$$

(d2) Conclude from part (c) that

$$\mathbb{P}\{X^\top AX - \text{tr}(A) \geq t\} \leq \exp\left(-c \min\left\{\frac{t^2}{\|A\|_{\text{op}} \text{tr}(A)}, \frac{t}{\|A\|_{\text{op}}}\right\}\right) \quad \text{for all } t \geq 0,$$

for a constant  $c > 0$  depending only on the MLSI constant  $C$ .

**[Bonus] Solution:**

From part (d1), we have  $f(X)$  is self-bounded with

$$a = 4 \|A\|_{\text{op}}, \quad \text{and} \quad b = 4 \|A\|_{\text{op}} \text{tr}(A).$$

Hence, using the tail bound proved in part (c3), we get

$$\mathbb{P}\{X^\top AX - \text{tr}(A) \geq t\} \leq \exp\left(-c \min\left\{\frac{t^2}{\|A\|_{\text{op}} \text{tr}(A)}, \frac{t}{\|A\|_{\text{op}}}\right\}\right) \quad \text{for all } t \geq 0. \quad \square$$

(d3) Compare this bound with the Gaussian quadratic-form bound from Homework 3, Exercise 5 (or Mid Exam, Exercise 2), where for Gaussian  $X$  one gets a sharper Frobenius-scale variance term involving  $\|A\|_{\text{F}}^2$  instead of  $\|A\|_{\text{op}} \text{tr}(A)$ .

**[Bonus] Solution:**

From Exercise 2 in Mid Exam, when  $X \sim \mathcal{N}(0, I_n)$ ,

$$\mathbb{P}\{X^\top AX - \text{tr}(A) \geq t\} \leq \exp\left(-c \min\left\{\frac{t^2}{\|A\|_{\text{F}}^2}, \frac{t}{\|A\|_{\text{op}}}\right\}\right) \quad \text{for all } t \geq 0.$$

Hence, for Gaussian  $X$ , the tail bound is stronger in the quadratic regime (i.e., in small-to-moderate deviation scale) because  $A \geq 0 \implies \lambda_i(A) \geq 0$  for all  $i \in \{1, \dots, n\}$  gives

$$\|A\|_{\text{F}}^2 = \sum_{i=1}^n \{\lambda_i(A)\}^2 \leq \lambda_{\max}(A) \sum_{i=1}^n \lambda_i(A) = \|A\|_{\text{op}} \text{tr}(A). \quad \square$$

## 2 Practice with covering and packing

Throughout, let  $(T, d)$  be a metric space, let  $K \subset T$ , and let  $\varepsilon > 0$ . Recall the following definitions:

- A set  $\mathcal{M} \subset K$  is an  $\varepsilon$ -net of  $K$  if for every  $x \in K$  there exists  $y \in \mathcal{M}$  such that

$$d(x, y) \leq \varepsilon.$$

Equivalently, the balls of radius  $\varepsilon$  centered at points of  $\mathcal{M}$  cover  $K$ .

- The *covering number* of  $K$  is

$$\mathcal{N}(K, d, \varepsilon) := \min\{|\mathcal{M}| : \mathcal{M} \subset K \text{ is an } \varepsilon\text{-net of } K\}.$$

- A set  $\mathcal{M} \subset K$  is  $\varepsilon$ -separated if

$$d(x, y) > \varepsilon \quad \text{for all distinct } x, y \in \mathcal{M}.$$

The *packing number* of  $K$  is

$$\mathcal{P}(K, d, \varepsilon) := \max\{|\mathcal{M}| : \mathcal{M} \subset K \text{ is } \varepsilon\text{-separated}\}.$$

- The *metric entropy* is the logarithm of the covering number:

$$H(K, d, \varepsilon) := \log \mathcal{N}(K, d, \varepsilon).$$

When the metric is clear from context, we write  $\mathcal{N}(K, \varepsilon)$ ,  $\mathcal{P}(K, \varepsilon)$ , and  $H(K, \varepsilon)$ .

- (a) **Monotonicity properties.** The first goal is to get comfortable with the definitions and with a subtle point: covering numbers are monotone in the scale  $\varepsilon$ , but not monotone in the set  $K$ .

- (a1) Show that the functions

$$\varepsilon \mapsto \mathcal{N}(K, d, \varepsilon) \quad \text{and} \quad \varepsilon \mapsto \mathcal{P}(K, d, \varepsilon)$$

are (weakly) decreasing in  $\varepsilon$ .

**Solution:**

Take  $0 < \varepsilon_1 \leq \varepsilon_2$ . If  $\mathcal{M} \subset K$  is an  $\varepsilon_1$ -net, then  $\forall x \in K, \exists y \in \mathcal{M}$  such that  $d(x, y) \leq \varepsilon_1 \leq \varepsilon_2$ . So, every  $\varepsilon_1$ -net is also an  $\varepsilon_2$ -net, i.e.,

$$\mathcal{P}(K, d, \varepsilon_2) \leq \mathcal{P}(K, d, \varepsilon_1).$$

Similarly, if  $\mathcal{M} \subset K$  is  $\varepsilon_2$ -separated, then  $\forall$  distinct  $x, y \in \mathcal{M}, d(x, y) > \varepsilon_2 \geq \varepsilon_1$ . So, every  $\varepsilon_2$ -separated set is also an  $\varepsilon_1$ -separated set, i.e.,

$$\mathcal{N}(K, d, \varepsilon_2) \leq \mathcal{N}(K, d, \varepsilon_1).$$

Hence, both the functions are weakly decreasing in  $\varepsilon$ .  $\square$

- (a2) Show by example that, in general,

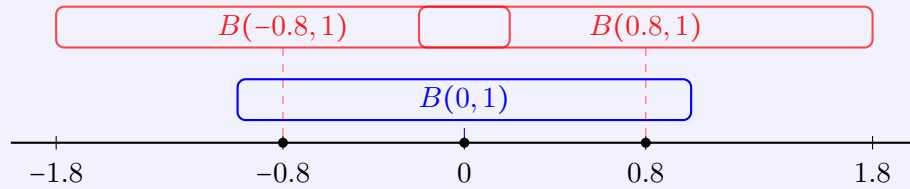
$$L \subset K \quad \not\Rightarrow \quad \mathcal{N}(L, d, \varepsilon) \leq \mathcal{N}(K, d, \varepsilon).$$

In other words, the covering number is not monotone in the underlying set.

**Solution:**

Take  $K = \{-0.8, 0, 0.8\}$  and  $L = \{-0.8, 0.8\} \subset K$  with metric  $d(x, y) = |x - y|$  and  $\varepsilon = 1$ .

One ball centered at  $0 \in K$  covers all of  $K$ ,  
so  $\mathcal{N}(K, d, 1) = 1$ .



But no single point of  $L$  covers both endpoints, since  
 $d(-0.8, 0.8) = 1.6 > 1$ , so  $\mathcal{N}(L, d, 1) = 2$ .

Figure 1: Covering number is not monotone in the set.

(a3) Prove the following approximate monotonicity statement: if  $L \subset K$ , then

$$\mathcal{N}(L, d, \varepsilon) \leq \mathcal{N}(K, d, \varepsilon/2).$$

**Solution:**

Let  $\mathcal{M} \subset K$  be an  $\varepsilon/2$ -net of  $K$  with  $|\mathcal{M}| = \mathcal{N}(K, d, \varepsilon/2)$ . For each  $x_i \in \mathcal{M}$  if  $B(x_i, \varepsilon/2) \cap L \neq \emptyset$ , choose a point

$$y_i \in B(x_i, \varepsilon/2).$$

Let  $\mathcal{M}'$  be the set of all such valid  $y_i$ . Then,

$$\mathcal{M}' \subset L, \quad \text{and} \quad |\mathcal{M}'| \leq \mathcal{N}(K, d, \varepsilon/2).$$

Now, fix any  $x \in L$ . Since  $L \subset K$  and  $\mathcal{M}$  is an  $\varepsilon/2$ -net of  $K$ ,  $\exists x_i \in \mathcal{M}$  such that

$$d(x, x_i) \leq \varepsilon/2.$$

Hence  $B(x_i, \varepsilon/2) \cap L \neq \emptyset$ , so there exists a corresponding  $y_i \in \mathcal{M}'$ . Using the triangle inequality,

$$d(x, y_i) \leq d(x, x_i) + d(x_i, y_i) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

because  $d(x_i, y_i) \leq \varepsilon/2$  by construction. Thus  $\mathcal{M}'$  is an  $\varepsilon$ -net of  $L$ . Hence,

$$\mathcal{N}(L, d, \varepsilon) \leq |\mathcal{M}'| \leq \mathcal{N}(K, d, \varepsilon/2). \quad \square$$

(b) **Packing versus covering.** This part develops the basic equivalence between covering and packing.

(b1) Let  $\mathcal{M} \subset K$  be a maximal  $\varepsilon$ -separated set, meaning that  $\mathcal{M}$  is  $\varepsilon$ -separated and one cannot add any new point of  $K$  to  $\mathcal{M}$  while preserving  $\varepsilon$ -separation. Show that  $\mathcal{M}$  is an  $\varepsilon$ -net of  $K$ .

**Solution:**

Consider an arbitrary point  $x \in K$ . If  $x \in \mathcal{M}$ , then the distance between  $x$  and  $\mathcal{M}$  is trivially 0. If instead  $x \in K \setminus \mathcal{M}$ , then  $\mathcal{M}' = \mathcal{M} \cup \{x\}$  is not  $\varepsilon$ -separated by the maximality of  $\mathcal{M}$ . This implies  $\exists y \in \mathcal{M}$  such that  $d(x, y) \leq \varepsilon$ . Hence, for all  $x \in K$ , there exists some  $y \in \mathcal{M}$  such that  $d(x, y) \leq \varepsilon$ , i.e.,  $\mathcal{M}$  is an  $\varepsilon$ -net of  $K$ .  $\square$

(b2) Deduce that for every  $K \subset T$  and every  $\varepsilon > 0$ ,

$$\mathcal{P}(K, d, 2\varepsilon) \leq \mathcal{N}(K, d, \varepsilon) \leq \mathcal{P}(K, d, \varepsilon).$$

Briefly explain why packing numbers and covering numbers are therefore equivalent up to universal changes of scale, and why the same is true for metric entropy.

**Solution:**

From part (b1), every maximal  $\varepsilon$ -separated set of  $K$  is an  $\varepsilon$ -net of  $K$ . We know,  $\mathcal{N}(K, d, \varepsilon)$  is the smallest cardinality of an  $\varepsilon$ -net of  $K$ , while  $\mathcal{P}(K, d, \varepsilon)$  is the maximum cardinality of an  $\varepsilon$ -separated subset of  $K$ . This gives,  $\mathcal{N}(K, d, \varepsilon) \leq \mathcal{P}(K, d, \varepsilon)$ . Now, let  $\mathcal{M}$  be a maximal  $2\varepsilon$ -separated set of  $K$ , so  $|\mathcal{M}| = \mathcal{P}(K, d, 2\varepsilon)$ . Then, any  $\varepsilon$ -ball can contain at most one point from the  $2\varepsilon$ -separated set  $\mathcal{M}$ . Thus, any  $\varepsilon$ -cover must have at least  $\mathcal{P}(K, d, 2\varepsilon)$  balls, i.e.,  $\mathcal{P}(K, d, 2\varepsilon) \leq \mathcal{N}(K, d, \varepsilon)$ . Hence, for every  $\varepsilon > 0$ ,

$$\mathcal{P}(K, d, 2\varepsilon) \leq \mathcal{N}(K, d, \varepsilon) \leq \mathcal{P}(K, d, \varepsilon). \quad \square$$

(c) **[Bonus] Volumetric bounds in Euclidean space.** From now on, let  $T = \mathbb{R}^n$  equipped with the Euclidean metric  $d(x, y) = \|x - y\|_2$ . Write

$$B_2^n := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}.$$

For sets  $A, B \subset \mathbb{R}^n$ , recall the Minkowski sum

$$A + B := \{a + b : a \in A, b \in B\}.$$

This part asks you to prove the standard volumetric bounds from the lecture note in full detail.

(c1) Let  $K \subset \mathbb{R}^n$  be measurable. Prove the volumetric bounds

$$\frac{\text{vol}(K)}{\text{vol}(\varepsilon B_2^n)} \leq \mathcal{N}(K, \varepsilon) \leq \mathcal{P}(K, \varepsilon) \leq \frac{\text{vol}(K + (\varepsilon/2)B_2^n)}{\text{vol}((\varepsilon/2)B_2^n)}.$$

**[Bonus] Solution:**

Let  $\mathcal{M}$  be an  $\varepsilon$ -net of  $K$  with  $|\mathcal{M}| = \mathcal{N}(K, \varepsilon)$ . By the definition of  $\varepsilon$ -nets,

$$K \subseteq \bigcup_{x \in \mathcal{M}} B(x, \varepsilon) = \bigcup_{x \in \mathcal{M}} (x + \varepsilon B_2^n).$$

Taking the volume on both sides and using the monotonicity and subadditivity of volume gives,

$$\text{vol}(K) \leq \sum_{x \in \mathcal{M}} \text{vol}(x + \varepsilon B_2^n) = \sum_{x \in \mathcal{M}} \text{vol}(\varepsilon B_2^n) = \mathcal{N}(K, \varepsilon) \cdot \text{vol}(\varepsilon B_2^n),$$

where we used the fact that volume is translation invariant. This proves the lower bound. Now, let  $\mathcal{M}'$  be a maximal  $\varepsilon$ -packing of  $K$  with  $|\mathcal{M}'| = \mathcal{P}(K, \varepsilon)$ . Because the points in  $\mathcal{M}'$  are  $\varepsilon$ -separated, the  $\varepsilon/2$ -balls centered at the distinct points in  $\mathcal{M}'$  are pairwise disjoint. Combining this fact with the additivity of volume for a disjoint union and translation-invariance of volume gives,

$$|\mathcal{M}'| \cdot \text{vol}((\varepsilon/2)B_2^n) = \text{vol}\left(\bigcup_{x \in \mathcal{M}'} (x + (\varepsilon/2)B_2^n)\right) = \text{vol}(\mathcal{M}' + (\varepsilon/2)B_2^n).$$

Since  $\mathcal{M}' \subset K$ , monotonicity of the volume gives,

$$\mathcal{P}(K, \varepsilon) \cdot \text{vol}((\varepsilon/2)B_2^n) \leq \text{vol}(K + (\varepsilon/2)B_2^n).$$

Finally, combining with the results from part (b2), we get,

$$\frac{\text{vol}(K)}{\text{vol}(\varepsilon B_2^n)} \leq \mathcal{N}(K, \varepsilon) \leq \mathcal{P}(K, \varepsilon) \leq \frac{\text{vol}(K + (\varepsilon/2)B_2^n)}{\text{vol}((\varepsilon/2)B_2^n)}. \quad \square$$

(c2) Deduce that for every  $\varepsilon > 0$ ,

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(B_2^n, \varepsilon) \leq \left(1 + \frac{2}{\varepsilon}\right)^n.$$

In particular, for  $0 < \varepsilon \leq 1$ ,

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(B_2^n, \varepsilon) \leq \left(\frac{3}{\varepsilon}\right)^n.$$

**[Bonus] Solution:**

Substituting  $K = B_2^n$  in part (c1), we get,

$$\frac{\text{vol}(B_2^n)}{\text{vol}(\varepsilon B_2^n)} \leq \mathcal{N}(B_2^n, \varepsilon) \leq \frac{\text{vol}(B_2^n + (\varepsilon/2)B_2^n)}{\text{vol}((\varepsilon/2)B_2^n)}.$$

The Lebesgue volume in  $\mathbb{R}^n$  is homogeneous of degree  $n$ , so

$$\text{vol}(\varepsilon B_2^n) = \varepsilon^n \cdot \text{vol}(B_2^n).$$

And, the unit ball in  $\mathbb{R}^n$ , i.e.,  $B_2^n$ , is convex, so

$$B_2^n + (\varepsilon/2)B_2^n = (1 + \varepsilon/2)B_2^n \implies \text{vol}(B_2^n + (\varepsilon/2)B_2^n) = (1 + \varepsilon/2)^n \cdot \text{vol}(B_2^n).$$

Hence, for every  $\varepsilon > 0$ ,

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(B_2^n, \varepsilon) \leq \left(\frac{1 + \varepsilon/2}{\varepsilon/2}\right)^n = \left(1 + \frac{2}{\varepsilon}\right)^n. \quad \square$$

For  $0 < \varepsilon \leq 1$ , since  $1 \leq (1/\varepsilon)$ ,  $(1 + 2/\varepsilon) \leq 3/\varepsilon$ , so the bound becomes,

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(B_2^n, \varepsilon) \leq \left(\frac{3}{\varepsilon}\right)^n. \quad \square$$

(c3) Show that the same upper bound holds for the Euclidean sphere  $S^{n-1}$ :

$$\mathcal{N}(S^{n-1}, \varepsilon) \leq \left(1 + \frac{2}{\varepsilon}\right)^n.$$

**[Bonus] Solution:**

Substituting  $K = S^{n-1}$  in part (c1) and taking the upper bound gives,

$$\mathcal{N}(S^{n-1}, \varepsilon) \leq \frac{\text{vol}(S^{n-1} + (\varepsilon/2)B_2^n)}{\text{vol}((\varepsilon/2)B_2^n)}.$$

Since  $S^{n-1} \subset B_2^n$ , monotonicity of the volume gives,

$$\mathcal{N}(S^{n-1}, \varepsilon) \leq \frac{\text{vol}(B_2^n + (\varepsilon/2)B_2^n)}{\text{vol}((\varepsilon/2)B_2^n)} = \left(1 + \frac{2}{\varepsilon}\right)^n. \quad \square$$

(c4) Conclude that for  $0 < \varepsilon \leq 1/2$ ,

$$H(B_2^n, \varepsilon) = \log \mathcal{N}(B_2^n, \varepsilon) \asymp n \log\left(\frac{e}{\varepsilon}\right),$$

up to absolute constants.

**[Bonus] Solution:**

Taking logarithms in the bounds from part (c2), we obtain

$$n \log\left(\frac{1}{\varepsilon}\right) \leq H(B_2^n, \varepsilon) = \log \mathcal{N}(B_2^n, \varepsilon) \leq n \log\left(1 + \frac{2}{\varepsilon}\right).$$

For  $0 < \varepsilon \leq 1/2$ , we have  $1 \leq \frac{1}{2\varepsilon}$ , which gives  $1 + \frac{2}{\varepsilon} \leq \frac{e}{\varepsilon}$ , so

$$H(B_2^n, \varepsilon) \leq n \log\left(\frac{e}{\varepsilon}\right).$$

Also,

$$\log\left(\frac{e}{\varepsilon}\right) = 1 + \log\left(\frac{1}{\varepsilon}\right) \asymp \log\left(\frac{1}{\varepsilon}\right) \quad \text{for } 0 < \varepsilon \leq \frac{1}{2},$$

because  $\log(1/\varepsilon) \geq \log 2$  gives  $\log\left(\frac{e}{\varepsilon}\right) \leq \left(1 + \frac{1}{\log 2}\right) \log\left(\frac{1}{\varepsilon}\right)$ . Therefore

$$H(B_2^n, \varepsilon) = \log \mathcal{N}(B_2^n, \varepsilon) \asymp n \log\left(\frac{e}{\varepsilon}\right),$$

up to absolute constants.  $\square$

### 3 Covariance estimation with sub-Gaussian random vectors

Throughout, let  $X$  denote a random vector in  $\mathbb{R}^d$ . We say that  $X$  is a *sub-Gaussian random vector* if every one-dimensional marginal  $\langle X, u \rangle$  is a sub-Gaussian random variable. Its sub-Gaussian norm is defined by

$$\|X\|_{\psi_2} := \sup_{u \in S^{d-1}} \|\langle X, u \rangle\|_{\psi_2}.$$

We also recall that  $X$  is *isotropic* if

$$\mathbb{E}X = 0, \quad \mathbb{E}[XX^\top] = I_d.$$

As in the previous homework sets,  $C, c > 0$  denote positive absolute constants whose values may change from line to line.

(a) **Coordinates and dependence.** This part introduces the basic geometry of the sub-Gaussian norm for random vectors.

(a1) Let  $X_1, \dots, X_d$  be independent, mean-zero, sub-Gaussian random variables, and let

$$X = (X_1, \dots, X_d) \in \mathbb{R}^d.$$

Show that  $X$  is a sub-Gaussian random vector and that

$$\max_{1 \leq i \leq d} \|X_i\|_{\psi_2} \leq \|X\|_{\psi_2} \leq C \max_{1 \leq i \leq d} \|X_i\|_{\psi_2}.$$

**Solution:**

Let  $e_i$  be the  $i$ -th standard basis vector of  $\mathbb{R}^d$ , then  $e_i \in S^{d-1}$  and  $\langle X, e_i \rangle = X_i$ . So,

$$\|X_i\|_{\psi_2} = \|\langle X, e_i \rangle\|_{\psi_2} \leq \|X\|_{\psi_2}.$$

Taking the maximum over  $i$  gives,

$$\max_{1 \leq i \leq d} \|X_i\|_{\psi_2} \leq \|X\|_{\psi_2}.$$

Now, fix  $u = (u_1, \dots, u_d)^\top \in S^{d-1}$ , then  $\sum_{i=1}^d u_i^2 = 1$  and  $\langle X, u \rangle = \sum_{i=1}^d u_i X_i$ . Because  $X_i$  are independent and mean-zero,  $u_i X_i$  will also be independent and mean-zero sub-Gaussian random variables. Then, using sub-Gaussian Hoeffding's inequality gives,

$$\|\langle X, u \rangle\|_{\psi_2}^2 \leq C \sum_{i=1}^d u_i^2 \|X_i\|_{\psi_2}^2 \leq C \max_{1 \leq i \leq d} \|X_i\|_{\psi_2}^2 \sum_{i=1}^d u_i^2 = C \max_{1 \leq i \leq d} \|X_i\|_{\psi_2}^2.$$

Taking supremum over  $u \in S^{d-1}$  and combining with the lower bound above, we get

$$\max_{1 \leq i \leq d} \|X_i\|_{\psi_2} \leq \|X\|_{\psi_2} \leq C \max_{1 \leq i \leq d} \|X_i\|_{\psi_2},$$

and hence,  $X$  is a sub-Gaussian random vector.  $\square$

(a2) Show by example that the independence assumption in part (a1) is essential: construct a random vector  $X \in \mathbb{R}^d$  with dependent coordinates such that

$$\|X\|_{\psi_2} \geq c\sqrt{d} \max_{1 \leq i \leq d} \|X_i\|_{\psi_2}.$$

**Solution:**

Let  $Z \sim \mathcal{N}(0, 1)$  be the standard normal and set  $X = (Z, \dots, Z)^\top \in \mathbb{R}^d$ , so  $\max_{1 \leq i \leq d} \|X_i\|_{\psi_2} = \|Z\|_{\psi_2}$ .

Now, choose  $u = \left(\frac{1}{\sqrt{d}}, \dots, \frac{1}{\sqrt{d}}\right)^\top \in S^{d-1}$ , then

$$\langle X, u \rangle = \sum_{i=1}^d \frac{1}{\sqrt{d}} Z = \sqrt{d} Z \implies \|\langle X, u \rangle\|_{\psi_2} = \sqrt{d} \|Z\|_{\psi_2}.$$

Hence, for  $X = (Z, \dots, Z)^\top$ ,

$$\|X\|_{\psi_2} \geq \sqrt{d} \|Z\|_{\psi_2}. \quad \square$$

(a3) [Bonus] Show that the uniform distribution on the cube  $[-1, 1]^d$  is sub-Gaussian with

$$\|X\|_{\psi_2} \leq C.$$

Show the same for the Boolean cube  $\{-1, +1\}^d$ .

**[Bonus] Solution:**

For the uniform distribution on  $[-1, 1]^d$ , the coordinates  $X_i$  are independent and each  $X_i \sim \text{Uniform}[-1, 1]$ . So,  $\mathbb{E}X_i = 0$  and  $|X_i| \leq 1$  almost surely. Using Hoeffding's lemma, a bounded random variable is sub-Gaussian with absolute-constant  $\psi_2$ -norm, so

$$\|X_i\|_{\psi_2} \leq C, \quad \text{for all } i.$$

Now, applying part (a1),  $X$  is sub-Gaussian with

$$\|X\|_{\psi_2} \leq C \max_{1 \leq i \leq d} \|X_i\|_{\psi_2} \leq C. \quad \square$$

Similarly, if  $X$  is uniform in Boolean cube  $\{-1, +1\}^d$ , then the coordinates  $X_i$  are independent Rademacher variables with  $|X_i| \leq 1$  almost surely. Hence, using similar argument each  $X_i$  is sub-Gaussian with  $\|X_i\|_{\psi_2} \leq C$ , and therefore by part (a1),

$$\|X\|_{\psi_2} \leq C. \quad \square$$

(a4) [Bonus] Let  $X \sim \mathcal{N}(0, \Sigma)$  in  $\mathbb{R}^d$ , where  $\Sigma \geq 0$ . Show that  $X$  is sub-Gaussian and

$$\|X\|_{\psi_2} \leq C \sqrt{\|\Sigma\|_{\text{op}}}.$$

**[Bonus] Solution:**

Take any  $u \in S^{d-1}$ , then  $\langle X, u \rangle$  will be a Gaussian with  $\mathbb{E}[\langle X, u \rangle] = 0$  and  $\text{Var}(\langle X, u \rangle) = u^\top \Sigma u$ . We know, if  $Y \sim \mathcal{N}(0, \sigma^2)$ , then  $\|Y\|_{\psi_2} \leq C\sigma$ , for some absolute constant  $C$ . Similarly,

$$\|\langle X, u \rangle\|_{\psi_2} \leq C \sqrt{u^\top \Sigma u}.$$

Here  $\Sigma \geq 0$ , so  $u^\top \Sigma u \leq \|\Sigma\|_{\text{op}} \|u\|_2^2 = \|\Sigma\|_{\text{op}}$  because  $\|u\|_2^2 = 1$ . So,  $\|\langle X, u \rangle\|_{\psi_2} \leq C \sqrt{\|\Sigma\|_{\text{op}}}$  and taking the supremum over all  $u \in S^{d-1}$  gives

$$\|X\|_{\psi_2} \leq \sup_{u \in S^{d-1}} \|\langle X, u \rangle\|_{\psi_2} \leq C \sqrt{\|\Sigma\|_{\text{op}}} \implies X \text{ is sub-Gaussian.} \quad \square$$

- (b) **Covariance estimation for sub-Gaussian data.** We now prove the operator-norm covariance estimation bound from Lecture 12 for sub-Gaussian data.

Let  $X_1, \dots, X_N$  be i.i.d. copies of a random vector  $X \in \mathbb{R}^d$ , and define the sample covariance matrix

$$\widehat{\Sigma} := \frac{1}{N} \sum_{i=1}^N X_i X_i^\top.$$

In this part, you may use the following two facts from earlier lectures/homework:

- If  $W$  is sub-Gaussian, then  $W^2 - \mathbb{E}W^2$  is sub-exponential and

$$\|W^2 - \mathbb{E}W^2\|_{\psi_1} \leq C \|W\|_{\psi_2}^2. \quad (4)$$

- If  $\mathcal{N}$  is a  $1/4$ -net of  $S^{d-1}$ , then  $|\mathcal{N}| \leq 12^d$ , and for every symmetric matrix  $A$ ,

$$\|A\|_{\text{op}} \leq 2 \max_{u \in \mathcal{N}} |u^\top A u|. \quad (5)$$

(b1) Assume first that  $X$  is isotropic and  $\|X\|_{\psi_2} \leq K$ . Fix  $u \in S^{d-1}$  and define

$$Z_i(u) := \langle X_i, u \rangle^2 - 1.$$

Show that  $Z_i(u)$  are independent, mean-zero, sub-exponential random variables satisfying

$$\|Z_i(u)\|_{\psi_1} \leq CK^2.$$

Deduce from Bernstein's inequality that for every  $t \geq 0$ ,

$$\mathbb{P} \left\{ |u^\top (\widehat{\Sigma} - I_d) u| \geq t \right\} \leq 2 \exp \left[ -cN \min \left( \frac{t^2}{K^4}, \frac{t}{K^2} \right) \right].$$

**Solution:**

Since the  $X_i$  are i.i.d. and  $Z_i(u)$  is function only of  $X_i$ , so  $Z_i(u)$  are independent. Here,  $X$  is isotropic (i.e.  $\Sigma = I_d$ ), so  $u^\top \Sigma u = 1$ , and as shown in part (a4),  $\langle X_i, u \rangle \sim \mathcal{N}(0, 1)$ . So,

$$\mathbb{E}[Z_i(u)] = \mathbb{E}[\langle X_i, u \rangle^2 - 1] = \text{Var}(\langle X_i, u \rangle) - 1 = 0.$$

Since  $\|X\|_{\psi_2} \leq K$ , by definition of the vector sub-Gaussian norm,  $\|\langle X_i, u \rangle\|_{\psi_2} \leq K$ . Now, using the given fact (4) with  $W = \langle X_i, u \rangle$  gives,  $\langle X_i, u \rangle^2 - \mathbb{E}[\langle X_i, u \rangle^2] = Z_i(u)$  is sub-exponential with

$$\|Z_i(u)\|_{\psi_1} \leq C \|\langle X_i, u \rangle\|_{\psi_2}^2 \leq CK^2. \quad \square$$

Now,

$$u^\top (\widehat{\Sigma} - I_d) u = \left[ u^\top \left( \frac{1}{N} \sum_{i=1}^N X_i X_i^\top \right) u \right] - 1 = \frac{1}{N} \sum_{i=1}^N (\langle X_i, u \rangle^2 - 1) = \frac{1}{N} \sum_{i=1}^N Z_i(u).$$

Since the  $Z_i(u)$  are independent and sub-exponential, using the Bernstein inequality for the sum of independent sub-exponential random variables gives

$$\mathbb{P} \left\{ |u^\top (\widehat{\Sigma} - I_d) u| \geq t \right\} \leq 2 \exp \left[ -c \min \left( \frac{t^2 N^2}{\sum_{i=1}^N \|Z_i(u)\|_{\psi_1}^2}, \frac{tN}{\max_i \|Z_i(u)\|_{\psi_1}} \right) \right].$$

Simplifying, we get, for every  $t \geq 0$ ,

$$\mathbb{P} \left\{ |u^\top (\widehat{\Sigma} - I_d) u| \geq t \right\} \leq 2 \exp \left[ -cN \min \left( \frac{t^2}{K^4}, \frac{t}{K^2} \right) \right]. \quad \square$$

- (b2) Continue in the isotropic case. Use a  $1/4$ -net of  $S^{d-1}$ , a union bound, and the deterministic net reduction above to show that

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - I_d\|_{\text{op}} \geq CK^2 \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d} \quad \text{provided } N \geq d.$$

**Solution:**

Let  $\mathcal{N}$  be a  $1/4$ -net of  $S^{d-1}$ , then using part (b1) to each  $u \in \mathcal{N}$  and union bounding gives,

$$\mathbb{P} \left\{ \max_{u \in \mathcal{N}} |u^\top (\widehat{\Sigma} - I_d)u| \geq t \right\} \leq 2|\mathcal{N}| \exp \left[ -cN \min \left( \frac{t^2}{K^4}, \frac{t}{K^2} \right) \right].$$

Here,  $(\widehat{\Sigma} - I_d)$  is symmetric, so using (5),  $|\mathcal{N}| \leq 12^d$  and  $\|\widehat{\Sigma} - I_d\|_{\text{op}} \leq 2 \max_{u \in \mathcal{N}} |u^\top (\widehat{\Sigma} - I_d)u|$ . Then,

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - I_d\|_{\text{op}} \geq 2t \right\} \leq 2 \exp \left[ d \log 12 - cN \min \left( \frac{t^2}{K^4}, \frac{t}{K^2} \right) \right].$$

Now, we choose  $t$  so that the exponent is at most  $-d$ , up to constants. That is, we want,

$$cN \min \left( \frac{t^2}{K^4}, \frac{t}{K^2} \right) \geq d.$$

So, in the quadratic regime:  $cN \frac{t^2}{K^4} \geq d \implies t \geq K^2 \sqrt{\frac{d}{N}}$ .

Similarly, in the exponential regime:  $cN \frac{t}{K^2} \geq d \implies t \geq K^2 \frac{d}{N}$ .

So, if we choose  $t = CK^2 \left( \sqrt{\frac{d}{N}} + \frac{d}{N} \right)$  with  $C$  large enough so that we can absorb the constant,  $\log 12$ , from the exponent, we will have

$$d \log 12 - cN \min \left( \frac{t^2}{K^4}, \frac{t}{K^2} \right) \leq -d.$$

Hence,

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - I_d\|_{\text{op}} \geq CK^2 \left( \sqrt{\frac{d}{N}} + \frac{d}{N} \right) \right\} \leq 2e^{-d}. \quad \square$$

When  $N \geq d$ ,  $\left( \sqrt{\frac{d}{N}} + \frac{d}{N} \right) \leq 2\sqrt{\frac{d}{N}}$ , and hence after adjusting the constants, we get

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - I_d\|_{\text{op}} \geq CK^2 \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d}. \quad \square$$

- (b3) [Bonus] Now assume  $X$  is mean-zero with covariance matrix

$$\Sigma = \mathbb{E}[XX^\top],$$

and assume for simplicity that  $\Sigma > 0$ . Suppose moreover that there exists  $K \geq 1$  such that

$$\|\langle X, u \rangle\|_{\psi_2} \leq K(\mathbb{E}\langle X, u \rangle^2)^{1/2} \quad \text{for all } u \in S^{d-1}.$$

Show that

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - \Sigma\|_{\text{op}} \geq CK^2 \|\Sigma\|_{\text{op}} \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d} \quad \text{provided } N \geq d.$$

In particular, if  $X \sim \mathcal{N}(0, \Sigma)$ , conclude that

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - \Sigma\|_{\text{op}} \geq C \|\Sigma\|_{\text{op}} \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d} \quad \text{provided } N \geq d.$$

**[Bonus] Solution:**

Define the whitened vector  $Y = \Sigma^{-1/2}X$ , then

$$\mathbb{E}[Y] = 0 \quad \text{and} \quad \mathbb{E}[YY^\top] = \mathbb{E}[\Sigma^{-1/2}XX^\top\Sigma^{-1/2}] = I_d,$$

so  $Y$  is isotropic. Now, for any  $u \in S^{d-1}$ ,

$$\langle Y, u \rangle = \langle \Sigma^{-1/2}X, u \rangle = \langle X, \Sigma^{-1/2}u \rangle.$$

Let  $v = \frac{\Sigma^{-1/2}u}{\|\Sigma^{-1/2}u\|_2} \in S^{d-1}$ , provided  $\Sigma^{-1/2}u \neq 0$ . Then,

$$\|\langle X, \Sigma^{-1/2}u \rangle\|_{\psi_2} = \|\Sigma^{-1/2}u\|_2 \|\langle X, v \rangle\|_{\psi_2} \leq K \|\Sigma^{-1/2}u\|_2 (\mathbb{E}[\langle X, v \rangle^2])^{1/2}.$$

But  $\mathbb{E}[\langle X, v \rangle^2] = v^\top \Sigma v$ , so with  $v = \frac{\Sigma^{-1/2}u}{\|\Sigma^{-1/2}u\|_2}$ ,

$$v^\top \Sigma v = \frac{u^\top \Sigma^{-1/2} \Sigma \Sigma^{-1/2} u}{\|\Sigma^{-1/2}u\|_2^2} = \frac{\|u\|_2^2}{\|\Sigma^{-1/2}u\|_2^2} = \frac{1}{\|\Sigma^{-1/2}u\|_2^2}.$$

Therefore,

$$\|\langle Y, u \rangle\|_{\psi_2} = \|\langle X, \Sigma^{-1/2}u \rangle\|_{\psi_2} \leq K.$$

Taking supremum over all  $u \in S^{d-1}$ ,

$$\|Y\|_{\psi_2} \leq K \implies Y \text{ is a sub-Gaussian vector.} \quad \square$$

Now, define the sample covariance of the whitened samples:

$$\widehat{\Sigma}_Y := \frac{1}{N} \sum_{i=1}^N Y_i Y_i^\top = \Sigma^{-1/2} \left( \frac{1}{N} \sum_{i=1}^N X_i X_i^\top \right) \Sigma^{-1/2} = \Sigma^{-1/2} \widehat{\Sigma} \Sigma^{-1/2}.$$

Since  $Y$  is isotropic and  $\|Y\|_{\psi_2} \leq K$ , part (b2) gives,

$$\mathbb{P} \left\{ \|\widehat{\Sigma}_Y - I_d\|_{\text{op}} \geq CK^2 \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d} \quad \text{provided } N \geq d. \quad \square$$

We know,

$$\widehat{\Sigma} - \Sigma = \Sigma^{1/2} (\widehat{\Sigma}_Y - I_d) \Sigma^{1/2} \implies \|\widehat{\Sigma} - \Sigma\|_{\text{op}} \leq \|\Sigma\|_{\text{op}} \|\widehat{\Sigma}_Y - I_d\|_{\text{op}}.$$

Hence,

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - \Sigma\|_{\text{op}} \geq CK^2 \|\Sigma\|_{\text{op}} \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d} \quad \text{provided } N \geq d. \quad \square$$

Now, if  $X \sim \mathcal{N}(0, \Sigma)$ , then  $\langle X, u \rangle$  will be Gaussian with

$$\mathbb{E}[\langle X, u \rangle] = 0 \quad \text{and} \quad \text{Var}(\langle X, u \rangle) = u^\top \text{Var}(X)u = u^\top \Sigma u.$$

Thus,  $\langle X, u \rangle \sim \mathcal{N}(0, \mathbb{E}[\langle X, u \rangle^2])$  implies

$$\|\langle X, u \rangle\|_{\psi_2} \leq C (\mathbb{E}[\langle X, u \rangle^2])^{1/2} \quad \text{for all } u \in S^{d-1}.$$

Thus, the assumption holds with an absolute constant  $K$ , and the bound becomes

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - \Sigma\|_{\text{op}} \geq C \|\Sigma\|_{\text{op}} \sqrt{\frac{d}{N}} \right\} \leq 2e^{-d} \quad \text{provided } N \geq d. \quad \square$$

## 4 Learning a spike model

This exercise studies the simplest structured covariance model, known as a spike model:

$$\Sigma = I_d + \beta uu^\top,$$

where  $u \in S^{d-1}$  is an unknown signal direction and  $\beta > 0$  is the signal-to-noise ratio (SNR). The leading eigenvector of  $\Sigma$  is exactly  $u$ , so if we can show that the sample covariance  $\widehat{\Sigma}$  is close to  $\Sigma$  in operator norm, then matrix perturbation theory should imply that the top eigenvector of  $\widehat{\Sigma}$  is close to  $u$  (up to sign).

Throughout,  $C, c > 0$  denote absolute constants whose values may change from line to line. Also, you may use the conclusion of part (a2) in part (b), even if you do not prove it.

(a) **[Bonus] Projection matrices and perturbation of top eigenvectors.** For a unit vector  $u \in \mathbb{R}^d$ , write

$$P_u := uu^\top,$$

the orthogonal projection onto the line spanned by  $u$ .

(a1) Let  $u, v \in S^{d-1}$ . Show that there exists a sign  $s \in \{-1, 1\}$  such that

$$\frac{1}{2} \|u - sv\|_2 \leq \|P_u - P_v\|_{\text{op}} \leq 2 \|u - sv\|_2.$$

### [Bonus] Solution:

Choose the sign  $s \in \{-1, +1\}$  so that  $\langle u, sv \rangle \geq 0$ . Then  $u$  and  $sv$  lie in a two-dimensional subspace and there is  $\theta \in [0, \pi/2]$  and some unit vector  $w \perp u$  such that

$$sv = \cos \theta u + \sin \theta w.$$

Since  $P_v = P_{sv}$ , we compare  $P_u$  and  $P_{sv}$ . In the orthonormal basis  $(u, w)$  of  $\text{span}\{u, v\}$ ,

$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad sv = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

So the corresponding projection matrices are:

$$P_u = uu^\top = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad P_{sv} = (sv)(sv)^\top = \begin{pmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix}.$$

This gives,

$$P_u - P_v = P_u - P_{sv} = \begin{pmatrix} \sin^2 \theta & -\sin \theta \cos \theta \\ -\sin \theta \cos \theta & -\sin^2 \theta \end{pmatrix},$$

where  $\text{tr}(P_u - P_v) = 0$  and  $\det(P_u - P_v) = -\sin^2 \theta$ , so using the characteristic polynomial

$$\lambda^2 - \text{tr}(P_u - P_v) + \det(P_u - P_v) = 0,$$

the eigen values of  $P_u - P_v$  are  $\pm \sin \theta$ , so  $\|P_u - P_v\|_{\text{op}} = \sin \theta$ . Now,

$$\|u - sv\|_2^2 = \|u - (\cos \theta u + \sin \theta w)\|_2^2 = (1 - \cos \theta)^2 + \sin^2 \theta = 4 \sin^2(\theta/2).$$

This gives,  $\|u - sv\|_2 = 2 \sin(\theta/2)$ , so we now compare  $\sin \theta$  and  $2 \sin(\theta/2)$ . We know,  $\sin \theta = 2 \sin(\theta/2) \cos(\theta/2)$  and  $\cos(\theta/2) \in [1/\sqrt{2}, 1]$  for  $\theta \in [0, \pi/2]$ , and hence

$$\frac{1}{\sqrt{2}} \|u - sv\|_2 \leq \|P_u - P_v\|_{\text{op}} \leq \|u - sv\|_2. \quad \square$$

(a2) Let  $A, B \in \mathbb{R}^{d \times d}$  be symmetric matrices with

$$\delta := \lambda_1(A) - \lambda_2(A) > 0.$$

Assume that  $v_1(A)$  and  $v_1(B)$  are unit top eigenvectors of  $A$  and  $B$ , respectively. Using the Davis-Kahan theorem from Lecture 12, deduce that there exists a sign  $s \in \{-1, 1\}$  such that

$$\|v_1(A) - s v_1(B)\|_2 \leq C \frac{\|A - B\|_{\text{op}}}{\delta}.$$

**[Bonus] Solution:**

We apply Davis-Kahan with  $k = 1$  because the eigengap is at  $k = 1$  and is given by

$$\delta = \lambda_1(A) - \lambda_2(A) > 0,$$

the projections will be rank-one projector onto  $\text{span}\{v_1(A)\}$  and  $\text{span}\{v_1(B)\}$ . Using the Davis-Kahan theorem, provided that  $\|A - B\|_{\text{op}} \leq \delta/2$ ,

$$\|P_{v_1(A)} - P_{v_1(B)}\|_{\text{op}} \leq \frac{2\|A - B\|_{\text{op}}}{\delta}.$$

Using part (a1) with  $u = v_1(A)$  and  $v = v_1(A)$ , there is some  $s \in \{-1, +1\}$  such that

$$\frac{1}{2} \|v_1(A) - s v_1(B)\|_2 \leq \|P_{v_1(A)} - P_{v_1(B)}\|_{\text{op}}.$$

Combining with Davis-Kahan, we get  $\|v_1(A) - s v_1(B)\|_2 \leq \frac{4\|A - B\|_{\text{op}}}{\delta}$ .

Now, if instead  $\|A - B\|_{\text{op}} > \delta/2$ , then trivially

$$\|v_1(A) - s v_1(B)\|_2 \leq 2 \leq \frac{4\|A - B\|_{\text{op}}}{\delta}.$$

Hence, in all cases, there exists a sign  $s \in \{-1, 1\}$  such that

$$\|v_1(A) - s v_1(B)\|_2 \leq C \frac{\|A - B\|_{\text{op}}}{\delta}. \quad \square$$

(b) **Learning a rank-one spike model.** Let  $u \in S^{d-1}$  and  $\beta > 0$ , and consider the covariance matrix

$$\Sigma := I_d + \beta u u^\top.$$

Let  $X_1, \dots, X_n$  be i.i.d. mean-zero random vectors in  $\mathbb{R}^d$  with covariance matrix  $\Sigma$ , and define the sample covariance

$$\widehat{\Sigma} := \frac{1}{n} \sum_{i=1}^n X_i X_i^\top.$$

In this part, you may use the following covariance-estimation fact from the previous exercise: if  $\|X_i\|_{\psi_2} \leq K$ , then

$$\mathbb{P} \left\{ \|\widehat{\Sigma} - \Sigma\|_{\text{op}} \geq CK^2 \left( \sqrt{\frac{d}{n}} + \frac{d}{n} \right) \right\} \leq 2e^{-d}.$$

(b1) Show that the two largest eigenvalues of  $\Sigma$  are

$$\lambda_1(\Sigma) = 1 + \beta, \quad \lambda_2(\Sigma) = 1,$$

and that the top eigenvector is  $v_1(\Sigma) = u$ .

**Solution:**

For the vector  $u$ ,

$$\Sigma u = (I_d + \beta uu^\top)u = u + \beta u(u^\top u) = (1 + \beta)u.$$

So,  $u$  is an eigenvector of  $\Sigma$  with eigenvalue  $1 + \beta$ . Now, take any  $v \perp u$ . Then  $u^\top v = 0$ , so

$$\Sigma v = (I_d + \beta uu^\top)v = v + \beta u(u^\top v) = v.$$

So, every vector orthogonal to  $u$  is an eigenvector of  $\Sigma$  with eigenvalue 1. This gives the eigenspace decomposition of  $\Sigma$ :

$$\mathbb{R}^d = \text{span}\{u\} \oplus u^\perp.$$

Hence,

$$\lambda_1(\Sigma) = 1 + \beta, \quad \text{and} \quad \lambda_2(\Sigma) = 1,$$

only when  $\beta > 0$ , and so the top eigenvector is  $v_1(\Sigma) = u$  up to sign.  $\square$

(b2) Assume in addition that the  $X_i$  are sub-Gaussian and satisfy

$$\|X_i\|_{\psi_2} \leq 10.$$

Let  $v = v_1(\widehat{\Sigma})$  be a unit top eigenvector of the sample covariance matrix. Show that if

$$n \geq C \frac{d}{\beta^2},$$

then

$$\min_{s \in \{-1, 1\}} \|u - sv\|_2 \leq 0.1$$

with probability at least  $1 - 2e^{-d}$ .

**Solution:**

$\Sigma$  has an eigengap of  $\beta > 0$  at  $k = 1$ , using part (a2) with  $A = \Sigma$  and  $B = \widehat{\Sigma}$ ,  $\exists s \in \{-1, 1\}$  s.t.

$$\|u - sv\|_2 \leq \frac{C \|\Sigma - \widehat{\Sigma}\|_{\text{op}}}{\beta}.$$

We have  $\|X_i\|_{\psi_2} \leq 10$ , so put  $K = 10$  in the covariance estimation fact given in the problem,

$$\|\widehat{\Sigma} - \Sigma\|_{\text{op}} \leq C \left( \sqrt{\frac{d}{n}} + \frac{d}{n} \right), \quad \text{with probability at least } 1 - 2e^{-d}.$$

Now, if  $n \geq \frac{Cd}{\beta^2}$ , then  $\sqrt{\frac{d}{n}} \leq \frac{\beta}{\sqrt{C}}$  and  $\frac{d}{n} \leq \frac{\beta^2}{C}$ . Since  $\beta \leq C'$  under  $\|X_i\|_{\psi_2} \leq 10$ , the second term is also  $\leq C'\beta/C$ . Thus,  $\sqrt{\frac{d}{n}} + \frac{d}{n} \leq C\beta$ , which implies  $\|\widehat{\Sigma} - \Sigma\|_{\text{op}} \leq C\beta$  with probability at least  $1 - 2e^{-d}$ . Finally, plugging this back into Davis–Kahan from part (a2) above,

$$\min_{s \in \{-1, 1\}} \|u - sv\|_2 \leq \|u - sv\|_2 \leq 0.1, \quad \text{with probability at least } 1 - 2e^{-d},$$

provided the constant in  $n \geq \frac{Cd}{\beta^2}$  is chosen large enough.  $\square$

## 5 Learning a Gaussian mixture model

One of the simplest models of structured high-dimensional data is a *Gaussian mixture model*. In the two-cluster version, one observes points drawn from one of two Gaussian distributions with different means. A basic example is

$$X = G + \theta t u,$$

where  $u \in S^{d-1}$  is a fixed unit vector,  $t > 0$  controls the separation between the two clusters,  $G \sim \mathcal{N}(0, I_d)$ , and  $\theta \in \{-1, +1\}$  is a Rademacher random variable independent of  $G$ . Equivalently,  $X$  is drawn from  $N(tu, I_d)$  or from  $N(-tu, I_d)$  with probability  $1/2$  each.

Thus the clusters are centered at  $\pm tu$ , and the direction  $u$  is the signal we would like to learn from data. Since the model is symmetric under  $u \mapsto -u$ , the best we can hope for is recovery of  $u$  up to sign.

Let  $X_1, \dots, X_n$  be i.i.d. copies of  $X$ , and define the sample covariance matrix

$$\widehat{\Sigma} := \frac{1}{n} \sum_{i=1}^n X_i X_i^\top.$$

In this problem, you may use the following two results proved earlier in Problems 3 and 4:

- Covariance estimation for sub-Gaussian data: If  $Y \in \mathbb{R}^d$  is mean zero with covariance matrix  $\Sigma_Y$ , and if

$$\|\langle Y, v \rangle\|_{\psi_2} \leq K \| \langle Y, v \rangle \|_{L_2} \quad \text{for all } v \in S^{d-1},$$

then

$$\mathbb{P} \left\{ \|\widehat{\Sigma}_Y - \Sigma_Y\|_{\text{op}} \geq CK^2 \left( \sqrt{\frac{d}{n}} + \frac{d}{n} \right) \|\Sigma_Y\|_{\text{op}} \right\} \leq 2e^{-d}.$$

- Davis–Kahan for top eigenvectors: If  $A, B$  are symmetric and  $\lambda_1(A) - \lambda_2(A) = \delta > 0$ , then

$$\min_{s \in \{-1, 1\}} \|v_1(A) - s v_1(B)\|_2 \leq C \frac{\|A - B\|_{\text{op}}}{\delta}.$$

Assume throughout that

$$\|u\|_2 = 1 \quad \text{and} \quad t \geq 0.1.$$

- (a) **Covariance and spike structure.** This part shows that the Gaussian mixture model has a rank-one spiked covariance structure.

- (a1) Show that  $\mathbb{E}X = 0$  and compute the covariance matrix of  $X$ , proving that

$$\Sigma := \mathbb{E}[X X^\top] = I_d + t^2 u u^\top.$$

**Solution:**

$$\mathbb{E}X = \mathbb{E}[G + \theta t u] = \mathbb{E}[G] + \mathbb{E}[\theta] t u = 0 + 0 = 0. \quad \square$$

Similarly, as  $\mathbb{E}X = 0$ ,

$$\Sigma = \text{Var}(X) = \mathbb{E}[X X^\top] = \text{Var}(G + \theta t u) \stackrel{(i)}{=} \text{Var}(G) + t^2 u \text{Var}(\theta) u^\top \stackrel{(ii)}{=} I_d + t^2 u u^\top,$$

where (i) uses the fact that  $G$  and  $\theta$  are independent, and (ii) uses the variance of a Rademacher random variable is 1.  $\square$

(a2) Deduce that the two largest eigenvalues of  $\Sigma$  are

$$\lambda_1(\Sigma) = 1 + t^2, \quad \lambda_2(\Sigma) = 1,$$

and that the top eigenspace is  $\text{span}(u)$ , so one may choose

$$v_1(\Sigma) = u.$$

**Solution:**

Using the conclusion for  $\Sigma = I_d + \beta uu^\top$  in Exercise 4 (b1) with  $\beta = t^2$ , the result follows.  $\square$

(b) **Learning the signal direction from data.** We now show that the top eigenvector of the sample covariance recovers the cluster-separation direction.

(b1) Show that there exists an absolute constant  $K$  such that for every  $v \in S^{d-1}$ ,

$$\|\langle X, v \rangle\|_{\psi_2} \leq K \|\langle X, v \rangle\|_{L_2}.$$

**Solution:**

Choose any  $v \in S^{d-1}$ , then

$$\langle X, v \rangle = X^\top v = G^\top v + \theta t u^\top v = \langle G, v \rangle + \theta t \langle u, v \rangle.$$

First, since  $G$  is isotropic, for all  $v \in S^{d-1}$ ,  $\langle G, v \rangle \sim \mathcal{N}(0, 1)$  (as proved in Exercise 3b1). So,

$$\|\langle G, v \rangle\|_{\psi_2} \leq C, \quad \text{for some absolute constant } C > 0.$$

Second,  $\theta$  is Rademacher, so  $|\theta t \langle u, v \rangle| \leq t |\langle u, v \rangle|$ . We know, a bounded random variable is sub-Gaussian. This gives, for some  $C > 0$ ,

$$\|\theta t \langle u, v \rangle\|_{\psi_2} \leq Ct |\langle u, v \rangle|.$$

Therefore, using triangle inequality for  $\psi_2$ -norm,

$$\|\langle X, v \rangle\|_{\psi_2} \leq \|\langle G, v \rangle\|_{\psi_2} + \|\theta t \langle u, v \rangle\|_{\psi_2} \leq C(1 + t |\langle u, v \rangle|).$$

Now, using  $\mathbb{E}[\theta^2] = 1$ ,

$$\mathbb{E}[\langle X, v \rangle^2] = \mathbb{E}[\langle G, v \rangle^2] + t^2 \langle u, v \rangle^2 = 1 + t^2 \langle u, v \rangle^2,$$

because  $\mathbb{E}[\theta \langle G, v \rangle] = (\mathbb{E}\theta)(\mathbb{E}[\langle G, v \rangle]) = 0$  by independence and centering. Thus, we have

$$\|\langle X, v \rangle\|_{L_2}^2 = \mathbb{E}[\langle X, v \rangle^2] = 1 + t^2 \langle u, v \rangle^2.$$

Using the elementary bound  $1 + |a| \leq 2\sqrt{1 + a^2}$  for all  $a \in \mathbb{R}$  with  $a = t \langle u, v \rangle$  gives,

$$1 + t |\langle u, v \rangle| \leq 2\sqrt{1 + t^2 \langle u, v \rangle^2} = 2 \|\langle X, v \rangle\|_{L_2}.$$

Hence, there exists an absolute constant  $K$  such that for every  $v \in S^{d-1}$ ,

$$\|\langle X, v \rangle\|_{\psi_2} \leq K \|\langle X, v \rangle\|_{L_2}. \quad \square$$

(b2) Let  $v := v_1(\widehat{\Sigma})$  be a unit top eigenvector of the sample covariance matrix. Show that if

$$n \geq C d$$

for a sufficiently large absolute constant  $C$ , then with probability at least  $1 - 2e^{-d}$ ,

$$\min_{s \in \{-1, 1\}} \|u - sv\|_2 \leq 0.1.$$

Thus, with  $n = O(d)$  unlabeled samples, one can recover the direction of separation in the Gaussian mixture model up to sign.

**Solution:**

Let  $v = v_1(\widehat{\Sigma})$ . From part (a2), the population covariance  $\Sigma = I_d + t^2 uu^\top$  has eigengap

$$\lambda_1(\Sigma) - \lambda_2(\Sigma) = t^2 > 0.$$

Therefore, by Exercise 4(a2) applied with  $A = \Sigma$  and  $B = \widehat{\Sigma}$ , there exists  $s \in \{-1, 1\}$  such that

$$\|u - sv\|_2 \leq C \frac{\|\widehat{\Sigma} - \Sigma\|_{\text{op}}}{t^2}.$$

Next, by part (b1),  $\exists K$  such that  $\|\langle X, w \rangle\|_{\psi_2} \leq K \|\langle X, w \rangle\|_{L_2}$  for all  $w \in S^{d-1}$ . So, using the covariance estimation result stated in the problem,

$$\|\widehat{\Sigma} - \Sigma\|_{\text{op}} \leq CK^2 \|\Sigma\|_{\text{op}} \left( \sqrt{\frac{d}{n}} + \frac{d}{n} \right), \quad \text{with probability at least } 1 - 2e^{-d}.$$

Since  $t \geq 0.1$ ,  $\|\Sigma\|_{\text{op}} = \lambda_1(\Sigma) = 1 + t^2 \leq C't^2$  for an absolute constant  $C'$ . Thus,

$$\|\widehat{\Sigma} - \Sigma\|_{\text{op}} \leq CK^2 t^2 \left( \sqrt{\frac{d}{n}} + \frac{d}{n} \right), \quad \text{with probability at least } 1 - 2e^{-d}.$$

Now choose the constant in the assumption  $n \geq C_0 d$  large enough so that

$$CK^2 \left( \sqrt{\frac{d}{n}} + \frac{d}{n} \right) \leq c$$

for a sufficiently small absolute constant  $c > 0$ . This gives

$$\|\widehat{\Sigma} - \Sigma\|_{\text{op}} \leq ct^2.$$

Substituting into the Davis–Kahan bound gives

$$\|u - sv\|_2 \leq C \frac{ct^2}{t^2} = Cc.$$

Choosing  $c$  small enough, equivalently  $C_0$  large enough, ensures that  $Cc \leq 0.1$ . Therefore,

$$\min_{s \in \{-1, 1\}} \|u - sv\|_2 \leq 0.1, \quad \text{with probability at least } 1 - 2e^{-d}. \quad \square$$

\*\*\* END OF SOLUTIONS \*\*\*