Homework 4

SDS 391P.6, Spring 2026

Pratik Patil

Due: Apr 06 (Monday)

This homework is (also) a work in progress and is provided as-is for instructional purposes. The problems are drawn from various sources and do not (yet) have sufficient references to the original material. Additionally, errors may be present. so some cauution is advised! The document will be updated if corrections are necessary. Last updated: 2026-03-22.

0 Guidelines

- Please start early. If you have questions about the statements, notation, or possible typos, email us as soon as possible. When emailing about the course, please begin the subject line with [SDS 391P.6].
- Please begin your answer to each *main* question on a separate page. If you use any code, include it in an appendix. Submit a single combined PDF to Canvas. (If you encounter any submission issues, please let us know.)
- The problems and motivations draw on multiple sources (past course material, textbooks, and other standard references). If you use any external resources that materially guide your solution (beyond routine lookups), please cite them in your write-up.
- These questions are designed to build intuition and technique. You are welcome to go beyond what is explicitly asked. If you introduce additional assumptions (while keeping the spirit of the problem), state them clearly. If you discover something interesting along the way, feel free to include it as a brief remark; we may share especially instructive observations with the class.
- Parts labeled [Bonus] are optional: they are not required for full credit. They are intended as extra practice. You may skip them without penalty. If you attempt them, please label your solutions clearly with [Bonus].
- Many parts include hints intended to help you get started. You are not required to follow the suggested route, and you are encouraged to try alternative approaches when appropriate.
- We will grade primarily for correctness and clear reasoning. Do not over-optimize for minor presentation details. The spirit of the homework is for you to learn something new!

1 Entropy method beyond uniform gradient bounds

In Lecture 7 (`https://pratikpatil.io/teaching/sds391p6-s26/lec7.pdf`), we saw that entropy can be used as an exponential analog of variance. In particular, if a centered random variable Z satisfies

$$\mathrm{Ent}(e^{\lambda Z}) \leqslant \frac{\lambda^2 v}{2}\,\mathbb{E}e^{\lambda Z} \qquad \text{for all } \lambda \geqslant 0,$$

then Herbst's argument gives the quadratic centered log-MGF bound

$$\psi_Z(\lambda) \;:=\; \log \mathbb{E} e^{\lambda Z} \;\leqslant\; \frac{\lambda^2 v}{2},$$

and therefore the Gaussian upper tail

$$\mathbb{P}\{Z \geqslant t\} \leqslant \exp\Big(-\frac{t^2}{2v}\Big) \qquad \text{for all } t \geqslant 0.$$

In Lecture 8 (`https://pratikpatil.io/teaching/sds391p6-s26/lec8.pdf`), we then used modified log-Sobolev inequalities (MLSI) to bound entropy by a squared-gradient energy.

The goal of this exercise is to go beyond the case where the gradient is uniformly bounded (e.g., as in the Tsirelson-Ibragimov-Sudakov inequality). Instead, we will keep the random energy

$$G := \|\nabla f(X)\|_2^2$$

inside the argument and derive a *nonlinear Bernstein-type concentration inequality* whose tail depends on the exponential moments of G. At the end, as a bonus, you will apply this result to positive semidefinite quadratic forms, to generalize results of Homework 3, Exercise 5 (and Mid Exam, Exercise 2) on Gaussian quadratic forms to random vectors satisfying following MLSI:

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

$$\mathrm{Ent}(e^{u(X)}) \;\leqslant\; C\,\mathbb{E}\Big[\|\nabla u(X)\|_2^2\, e^{u(X)}\Big] \qquad \text{for all smooth } u : \mathbb{R}^n \to \mathbb{R},$$

where $C > 0$ is a fixed constant.

For a nonnegative random variable Y, recall the concentration entropy

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

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

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

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

$$\frac{\mathrm{Ent}(e^{\theta Z})}{\mathbb{E}e^{\theta Z}} = \theta\,\psi_Z'(\theta) - \psi_Z(\theta) = \theta^2 \left(\frac{\psi_Z(\theta)}{\theta}\right)'. \tag{1}$$

In particular, an upper bound on

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

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

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

$$\mathbb{E}[WY] \;\leqslant\; \log \mathbb{E}e^{W} + \mathrm{Ent}(Y).$$

Equivalently, in our notation:

$$\mathbb{E}[WY] \;\leqslant\; \kappa_W(1) + \mathrm{Ent}(Y).$$

Identify the equality condition.

(*Hint:* Compare Y with the tilted random density $e^W/\mathbb{E}e^W$ in terms of relative entropy, and use positivity of relative entropy.)

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

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

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

$$\mathrm{Ent}(e^{\theta Z}) \;\leqslant\; C\theta^2\, \mathbb{E}\Big[G\, e^{\theta Z}\Big].$$

(*Hint:* Apply MLSI to the function $u(X) = \theta Z$.)

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

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

and show that

$$\frac{\mathbb{E}\big[G\, e^{\theta Z}\big]}{\mathbb{E}e^{\theta Z}} \;\leqslant\; \frac{\kappa_G(\eta)}{\eta} + \frac{\mathrm{Ent}(e^{\theta Z})}{\eta\, \mathbb{E}e^{\theta Z}}.$$

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

$$\frac{\mathrm{Ent}(e^{\theta Z})}{\theta^2\, \mathbb{E}e^{\theta Z}} \;\leqslant\; \frac{C\, \kappa_G(\eta)}{\eta - C\theta^2}.$$

(b4) Use the Herbst identity (1) to show that whenever $\eta > C\theta^2$,

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

(*Hint:* For $0 \leqslant s \leqslant \theta$, bound $\eta - Cs^2$ below by $\eta - C\theta^2$.)

(b5) Deduce the one-sided tail bound

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

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 $G = \|\nabla f(X)\|_2^2$.

(*Hint:* Use the Chernoff method.)

(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

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

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

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

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

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

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

we have

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

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

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

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

(*Hint:* Optimize the Chernoff bound using the estimate from part (c2), and choose $\theta = \min\{t/(4Cb),\, 1/(4Ca)\}$.)

(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

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

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

$$\|\nabla f(X)\|_2^2 = 4X^\top A^2 X \leqslant 4\|A\|\, X^\top AX = 4\|A\|\big(f(X) + \operatorname{tr}(A)\big).$$

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

$$a = 4\|A\|, \qquad b = 4\|A\|\ \operatorname{tr}(A).$$

(*Hint:* Use $A^2 \preceq \|A\|A$, which holds because $A \succeq 0$.)

(d2) Conclude from part (c) that

$$\mathbb{P}\Big\{X^\top AX - \operatorname{tr}(A) \geqslant t\Big\} \leqslant \exp\left(-c\min\left\{\frac{t^2}{\|A\|\ \operatorname{tr}(A)}, \frac{t}{\|A\|}\right\}\right) \qquad \text{for all } t \geqslant 0,$$

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

(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 $\|A\|_F^2$ instead of $\|A\| \operatorname{tr}(A)$.

(*Hint:* Standard Gaussian vectors X satisfy the MLSI above. For such standard Gaussian X, you obtained a sharper bound involving $\|A\|_F^2$ instead of $\|A\| \operatorname{tr}(A)$.)

2 Practice with covering and packing

In Lecture 11 (https://pratikpatil.io/teaching/sds391p6-s26/lec11.pdf), we introduced several geometric tools that help us discretize complicated index sets: ϵ-nets, covering numbers, packing numbers, metric entropy, among others. This exercise is meant to give you practice with the definitions and with the basic arguments behind these notions.

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

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

  $$d(x, y) \leqslant \epsilon.$$

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

- The *covering number* of K is

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

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

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

  The *packing number* of K is

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

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

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

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

(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 ϵ, but not monotone in the set K.

(a1) Show that the functions

$$\epsilon \mapsto \mathcal{N}(K, d, \epsilon) \qquad \text{and} \qquad \epsilon \mapsto \mathcal{P}(K, d, \epsilon)$$

are (weakly) decreasing in ϵ.

(*Hint:* If $\epsilon_1 \leqslant \epsilon_2$, then every ϵ_1-net is automatically an ϵ_2-net. Likewise, every ϵ_2-separated set is automatically ϵ_1-separated.)

(a2) Show by example that, in general,

$$L \subset K \qquad \not\Rightarrow \qquad \mathcal{N}(L, d, \epsilon) \leqslant \mathcal{N}(K, d, \epsilon).$$

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

(*Hint:* Work in $T = \mathbb{R}$ with the Euclidean metric and try a two-point set inside an interval.)

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

$$\mathcal{N}(L, d, \epsilon) \leqslant \mathcal{N}(K, d, \epsilon/2).$$

(*Hint:* Start with an $(\epsilon/2)$-net of K. For each net point whose $(\epsilon/2)$-ball meets L, choose one representative point from L inside that ball.)

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

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

(*Hint:* If $\mathcal{M}$ were not an ϵ-net, then there would exist a point of K that could be added to $\mathcal{M}$ without violating ϵ-separation.)

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

$$\mathcal{P}(K, d, 2\epsilon) \ \leqslant\ \mathcal{N}(K, d, \epsilon) \ \leqslant\ \mathcal{P}(K, d, \epsilon).$$

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.

(*Hint:* For the upper bound, use part (b1). For the lower bound, show that an ϵ-ball can contain at most one point of a 2ϵ-separated set.)

(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

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

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

$$\frac{\mathrm{vol}(K)}{\mathrm{vol}(\epsilon B_2^n)} \ \leqslant\ \mathcal{N}(K, \epsilon) \ \leqslant\ \mathcal{P}(K, \epsilon) \ \leqslant\ \frac{\mathrm{vol}\big(K + (\epsilon/2)B_2^n\big)}{\mathrm{vol}((\epsilon/2)B_2^n)}.$$

(*Hint:* For the left inequality, compare the volume of K to the total volume of a covering by ϵ-balls. For the right inequality, start from an ϵ-separated set and use the fact that the $(\epsilon/2)$-balls around its points are disjoint.)

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

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

In particular, for $0 < \epsilon \leqslant 1$,

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

(*Hint:* Use $\mathrm{vol}(rB_2^n) = r^n \mathrm{vol}(B_2^n)$.)

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

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

(*Hint:* The naive set-inclusion argument is not enough here, since covering numbers are not monotone in the set. Instead, mimic the packing/volume argument directly on the sphere, observing that disjoint $(\epsilon/2)$-balls centered on the sphere fit inside $(1 + \epsilon/2)B_2^n$.)

(c4) Conclude that for $0 < \epsilon \leqslant 1/2$,

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

up to absolute constants.

3 Covariance estimation with sub-Gaussian random vectors

In Lecture 12 (https://pratikpatil.io/teaching/sds391p6-s26/lec12.pdf), we used metric entropy to prove operator-norm covariance estimation for isotropic sub-Gaussian data. This exercise develops that argument systematically for anisotropic sub-Gaussian data (defined precisely below). We begin by introducing sub-Gaussian random vectors, then we work through some basic examples, and finally we prove the covariance-estimation bound similar to the one in Lecture 12.

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

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

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

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

(*Hint:* For the lower bound, test against the standard basis vectors. For the upper bound, write $\langle X, u \rangle = \sum_{i=1}^d u_i X_i$ and use the bound for sums of independent sub-Gaussian random variables.)

(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

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

(*Hint:* Take all coordinates to be identical.)

(a3) [Bonus] Canonical examples.

(a31) Show that the uniform distribution on the cube $[-1, 1]^d$ is sub-Gaussian with

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

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

(*Hint:* Each coordinate is bounded by 1, hence is sub-Gaussian with absolute-constant ψ_2 norm. Then use part (a1).)

(a32) Let $X \sim \mathcal{N}(0, \Sigma)$ in $\mathbb{R}^d$, where $\Sigma \succeq 0$. Show that X is sub-Gaussian and

$$\|X\|_{\psi_2} \leqslant C \sqrt{\|\Sigma\|}.$$

(*Hint:* For each $u \in S^{d-1}$, identify the law of $\langle X, u \rangle$ and use the variational characterization of $\|\Sigma\|$.)

(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

$$\|W^2 - \mathbb{E}W^2\|_{\psi_1} \leqslant C\|W\|_{\psi_2}^2.$$

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

$$\|A\| \leqslant 2 \max_{u \in \mathcal{N}} |u^\top A u|.$$

(b1) Assume first that X is isotropic and $\|X\|_{\psi_2} \leqslant K$. 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

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

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

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

(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

$$\mathbb{P}\left\{ \|\widehat{\Sigma} - I_d\| \geqslant CK^2 \sqrt{\frac{d}{N}} \right\} \leqslant 2e^{-d} \qquad \text{provided } N \geqslant d.$$

(*Hint:* Apply part (c1) to all points of the net. Then choose $t \asymp K^2 \sqrt{d/N}$ and make the constant large enough to absorb the net cardinality 12^d.)

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

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

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

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

Show that

$$\mathbb{P}\left\{ \|\widehat{\Sigma} - \Sigma\| \geqslant CK^2 \|\Sigma\| \sqrt{\frac{d}{N}} \right\} \leqslant 2e^{-d} \qquad \text{provided } N \geqslant d.$$

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

$$\mathbb{P}\left\{ \|\widehat{\Sigma} - \Sigma\| \geqslant C \|\Sigma\| \sqrt{\frac{d}{N}} \right\} \leqslant 2e^{-d} \qquad \text{provided } N \geqslant d.$$

(*Hint:* Whiten the vector by setting $Y = \Sigma^{-1/2} X$, apply part (c2) to Y, and then transfer the bound back to X.)

4 Learning a spike model

In the previous Problem 3, we obtained operator-norm bounds for the sample covariance matrix. This is already enough to control the *eigenvalues* of the sample covariance by Weyl's inequality. To control the *eigenvectors*, however, we need perturbation theory.

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 Σ is exactly u, so if we can show that the sample covariance $\widehat{\Sigma}$ is close to Σ 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

$$\frac{1}{2}\,\|u - sv\|_2 \;\leqslant\; \|P_u - P_v\| \;\leqslant\; 2\,\|u - sv\|_2.$$

(*Hint:* First choose the sign s so that $\langle u, sv\rangle \geqslant 0$. Then u and sv lie in a two-dimensional subspace, so write

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

for some $w \perp u$ with $\|w\|_2 = 1$ and some $\theta \in [0, \pi/2]$. Work in the orthonormal basis (u, w) of $\operatorname{span}\{u, v\}$, compute the 2×2 matrix of $P_u - P_v$, and compare its norm with $\|u - sv\|_2$. If you prefer a softer upper bound, you can also use

$$uu^\top - vv^\top = u(u - sv)^\top + (u - sv)(sv)^\top$$

together with the fact from Homework 1, Problem 2(b1) that $\|xy^\top\| = \|x\|_2\|y\|_2$ for $x, y \in \mathbb{R}^d$.)

(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 (https://pratikpatil.io/teaching/sds391p6-s26/lec12.pdf), deduce that there exists a sign $s \in \{-1, 1\}$ such that

$$\|v_1(A) - s\, v_1(B)\|_2 \;\leqslant\; C\,\frac{\|A - B\|}{\delta}.$$

(*Hint:* Apply the projector version of Davis–Kahan to

$$P_A = v_1(A)v_1(A)^\top, \qquad P_B = v_1(B)v_1(B)^\top.$$

This gives a bound on $\|P_A - P_B\|$. Then use part (a1) to pass from projections to vectors. If $\|A - B\|$ is not small compared with δ, then the desired inequality is trivial after increasing the constant, since the left-hand side is always at most 2.)

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

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

Let $X_1, \dots, X_n$ be i.i.d. mean-zero random vectors in $\mathbb{R}^d$ with covariance matrix Σ, 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 $\|X_i\|_{\psi_2} \leqslant K$, then

$$\mathbb{P}\left\{ \|\widehat{\Sigma} - \Sigma\| \geqslant CK^2 \left(\sqrt{\frac{d}{n}} + \frac{d}{n} \right) \right\} \leqslant 2e^{-d}.$$

(b1) Show that the two largest eigenvalues of Σ are

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

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

(*Hint:* Check separately what Σ does to the vector u and to vectors orthogonal to u.)

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

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

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

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

then

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

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

(*Hint:* Combine the covariance-estimation bound with part (b1) and the perturbation bound from part (a2). A useful preliminary observation is that β is automatically bounded by an absolute constant: indeed,

$$1 + \beta = u^\top \Sigma u = \mathbb{E}\langle X_i, u\rangle^2 = \|\langle X_i, u\rangle\|_{L^2}^2 \lesssim \|\langle X_i, u\rangle\|_{\psi_2}^2 \leqslant 100.$$

So once $n \geqslant Cd/\beta^2$ with C sufficiently large, both terms $\sqrt{d/n}$ and d/n are small enough compared with the eigengap β.)

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 $X_1, \ldots, X_n$ 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 Σ_Y, and if

  $$\|\langle Y, v\rangle\|_{\psi_2} \leqslant K\, \|\langle Y, v\rangle\|_{L^2} \qquad \text{for all } v \in S^{d-1},$$

  then

  $$\mathbb{P}\left\{ \|\widehat{\Sigma}_Y - \Sigma_Y\| \geqslant CK^2 \left(\sqrt{\frac{d}{n}} + \frac{d}{n} \right) \|\Sigma_Y\| \right\} \leqslant 2e^{-d}.$$

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

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

Assume throughout that

$$\|u\|_2 = 1 \qquad \text{and} \qquad t \geqslant 0.1.$$

(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

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

  (*Hint:* Write $X = G + \theta t u$, expand $XX^\top$, and use that $\mathbb{E}G = 0$, $\mathbb{E}\theta = 0$, and θ is independent of G.)

  (a2) Deduce that the two largest eigenvalues of Σ are

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

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

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

(*Hint:* Check separately what Σ does to the vector u and to vectors orthogonal to u.)

(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 $v \in S^{d-1}$,

$$\|\langle X, v\rangle\|_{\psi_2} \leqslant K \,\|\langle X, v\rangle\|_{L^2}.$$

(*Hint:* Write

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

Use that $\langle G, v\rangle \sim \mathcal{N}(0, 1)$, that the bounded random variable $\theta t\langle u, v\rangle$ is sub-Gaussian, and that

$$\|\langle X, v\rangle\|_{L^2}^2 = \mathrm{Var}(\langle X, v\rangle) = 1 + t^2\langle u, v\rangle^2.$$

Finally, compare $1 + t|\langle u, v\rangle|$ with $\sqrt{1 + t^2\langle u, v\rangle^2}$.)

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

$$n \geqslant C\, d$$

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

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

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

(*Hint:* Use part (b1) together with the covariance-estimation theorem above, then apply Davis–Kahan using the eigengap from part (a2). Since $t \geqslant 0.1$, the eigengap $\lambda_1(\Sigma) - \lambda_2(\Sigma) = t^2$ is bounded below by an absolute constant, and also

$$\|\Sigma\| = 1 + t^2 \leqslant Ct^2$$

for an absolute constant C. So the covariance-estimation error is of order

$$\sqrt{\frac{d}{n}} + \frac{d}{n},$$

up to absolute constants.)

Source material

Parts of this homework were inspired by exercises from Tropp (2023); van Handel (2016), in addition to the author's accumulated experience working on related topics.

References

Tropp, J. A. (2023). Probability in high dimensions. Caltech CMS Lecture Notes 2021-01.

van Handel, R. (2016). Probability in high dimension. Lecture Notes (Princeton University). APC 550.