MA3K0 High-Dimensional Probability
Example Sheet 3

2020, term 2
Stefan Adams

Students should hand in solutions to the questions by noon 12 pm Friday week 7 (21.02.2020) to the maths drop off box outside the undergraduate office.

Exercise 9: Uniform distribution on the Euclidean ball [25 points] Denote $B(0,\sqrt{n})$ the Euclidean ball around the origin with radius $\sqrt{n}$,

$$B(0,\sqrt{n}) = \big\{x \in \mathbb{R}^n \colon x_1^2 + \cdots + x_n^2 \le n\big\}.$$

Show that a random vector

$$X \sim \mathsf{Unif}(B(0,\sqrt{n}))$$

(X uniformly distributed in the Euclidean ball $B(0,\sqrt{n})$ is sub-Gaussian, and $\|X\|_{\psi_2} \le C$.

(*Hint: This is an extension of the Theorem in the lecture on uniformly distributed random vectors on the Euclidean sphere* $\sqrt{n}\mathsf{S}^{n-1}$.)

Exercise 10: ε-nets and Operator norm

(a) [7 points] Let $x \in \mathbb{R}^n$ and let $\mathcal{N}$ be an ε-net of the sphere S^{n-1} with $\varepsilon > 0$. Show that

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

where $\|\cdot\|_2$ denotes the Euclidean norm.

(b) [8+10 points] Suppose A is an $m\times n$ matrix and $\varepsilon \in [0,\frac{1}{2})$. Recall that the operator norm is given by maximising the quadratic form

$$\|A\| = \sup_{x\in\mathsf{S}^{n-1}, y\in\mathsf{S}^{m-1}}\{\langle Ax,y\rangle\}.$$

(i) 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

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

(ii) Show that, when $m=n$ and A is symmetric,

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

Exercise 11: Covariance estimation random matrices [20 points] Recall the *Covariance estimation theorem* [1] from the lecture. Suppose $X^{(1)}, \dots, X^{(N)}$ are samples of a random vector $X \in \mathbb{R}^n$. The *empirical* or *sample* covariance matrix Σ_N is defined as

$$\Sigma_N = \frac{1}{N} \sum_{i=1}^{N} X^{(i)} (X^{(i)})^T.$$

Under the assumptions of the theorem, show that, for any $u \geq 0$, we have

$$\|\Sigma_N - \Sigma\| \leq CK^2 \Big(\sqrt{\frac{n+u}{N}} + \frac{n+u}{N} \Big) \|\Sigma\|$$

with probability at least $1 - 2\mathrm{e}^{-u}$.

Exercise 12: Entropy Consider real-valued random variables X and suppose that the moment generating function $M_X(\lambda) = \mathbb{E}\big[\mathrm{e}^{\lambda X}\big]$ as well as the derivative $M_X'(\lambda)$ exists for all $\lambda \in \mathbb{R}$. Recall that $\mathsf{H}(X)$ is the entropy functional for the convex function $\varphi(u) = u \log u$ for $u > 0$, and $\varphi(0) := 0$.

(a) [7 points] Show that the entropy has the variational representation

$$\mathsf{H}\big(\mathrm{e}^{\lambda X}\big) = \inf_{t \in \mathbb{R}} \big\{ \mathbb{E}\big[\psi(\lambda(X - t)) \mathrm{e}^{\lambda X}\big] \big\},$$

where $\psi(u) := \mathrm{e}^{-u} - 1 + u$.

(b) [7 points] Let H_θ denote the entropy functional defined by the convex function $\theta(u) := u \log u - u$ for $u > 0$, and $\theta(0) := 0$. Show that

$$\mathsf{H}_\theta(\mathrm{e}^{\lambda X}) = \mathsf{H}(\mathrm{e}^{\lambda X}).$$

(c) [8+8 points] Recall the Bernstein bound from the lecture, that is, there are positive constants B and σ such that

$$(\diamond) \qquad \mathsf{H}(\mathrm{e}^{\lambda X}) \leq \lambda^2 \big[B M_X'(\lambda) + M_X(\lambda)(\sigma^2 - B\mathbb{E}[X]) \big] \quad \text{for all } \lambda \in [0, \frac{1}{B}).$$

(i) Show that a random variable X satisfies the Bernstein bound $(\diamond)$ if and only if $\widetilde{X} = X - \mathbb{E}[X]$ satisfies the inequality

$$(\clubsuit) \qquad \mathsf{H}(\mathrm{e}^{\lambda \widetilde{X}}) \leq \lambda^2 \big[B M_{\widetilde{X}}'(\lambda) + M_{\widetilde{X}}(\lambda) \sigma^2 \big] \quad \text{for all } \lambda \in [0, \frac{1}{B}).$$

(ii) Show that a mean zero random variable X satisfies inequality $(\clubsuit)$ if and only if $\widehat{X} = \frac{1}{B} X$ satisfies

$$(\spadesuit) \qquad \mathsf{H}(\mathrm{e}^{\lambda \widehat{X}}) \leq \lambda^2 \big[M_{\widehat{X}}'(\lambda) + \widehat{\sigma}^2 M_{\widehat{X}}(\lambda) \big] \quad \text{for all } \lambda \in [0, 1),$$

where $\widehat{\sigma}^2 = \sigma^2 / B^2$.

[1] Covariance estimation theorem: Suppose X is sub-Gaussian random vector in $\mathbb{R}^n$ and assume that there exists $K \geq 1$ such that

$$\|\langle X, x \rangle\|_{\psi_2} \leq K \|\langle X, x \rangle\| \quad \text{for any } x \in \mathbb{R}^n.$$

Then, for every $N \in \mathbb{N}$, we have

$$\mathbb{E}[\|\Sigma_N - \Sigma\|] \leq CK^2 \Big(\sqrt{\frac{n}{N}} + \frac{n}{N} \Big) \|\Sigma\|$$

for some absolute constant $C > 0$.

hand-in times – assignment sheets:

Sheet	week	date	discussion in Support class
1	3	24.01.2020 at noon 12 pm	week 4
2	5	07.02.2020 at noon 12 pm	week 6
3	7	21.02.2020 at noon 12 pm	week 8
4	9	06.03.2020 at noon 12 pm	week 10