

Math of ML : Exercises 13 *

December 15, 2025

Before attempting the exercises, read the lecture notes, in case we did not have the time to finish the proof in class (which is probable).

Exercise 1 (Propagation of Lipschitz regularity). *Prove Lemma 3.2.*

Exercise 2 (Propagation of subgaussian tails). *Prove Lemma 3.3.*

Indications: For a subgaussian vector $X \in \mathbb{R}^d$, consider the (optimal) variance proxy seminorm

$$\|X\|_{vp} := \inf \left\{ s > 0 ; \mathbf{E}[e^{u^\top (X - \mathbf{E}[X])}] \leq e^{s^2 \|u\|_2^2 / 2}, \forall u \in \mathbb{R}^d \right\}.$$

We may admit or verify the following: $\|\cdot\|_{vp}$ is a seminorm (it satisfies positive homogeneity, triangle inequality, and is nonnegative but not definite). Moreover if $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is L -Lipschitz continuous, then there is an absolute $c > 0$ such that $\|f(X)\|_{vp} \leq cL\sqrt{d}\|X\|_{vp}$.

Exercise 3 (Multivariate Subgaussian Concentration via a Covering Argument). *In this exercise, we extend the subgaussian concentration result (from Lecture 3), to the multivariate case. This extension was used in Lectures 11 and 13. Let $X \in \mathbb{R}^d$ be a mean-zero subgaussian random vector in the sense that $\sigma^2 = \|X\|_{vp} < +\infty$ (see definition in previous exercise).*

(i) **One-dimensional concentration.** *Prove that for any fixed $u \in S^{d-1}$ (the ℓ_2 unit sphere in \mathbb{R}^d) and any $t \geq 0$,*

$$\mathbf{P} \left(X^\top u \geq t \right) \leq \exp \left(-\frac{t^2}{2\sigma^2} \right).$$

(ii) **ϵ -net.** *If $A \subset \mathbb{R}^d$ is a bounded set, we say that a finite subset $\mathcal{N}_\epsilon \subset A$ is an ϵ -net of A iff $A \subset \cup_{x \in \mathcal{N}_\epsilon} B(x, \epsilon)$ where $B(x, \epsilon)$ is the closed ℓ_2 -ball of center x and radius ϵ . For $A = S^{d-1}$ the unit sphere in \mathbb{R}^d , show that for any $\epsilon \in]0, 1[$,*

$$\|X\|_2 \leq \frac{1}{1 - \epsilon} \sup_{u \in \mathcal{N}_\epsilon} u^\top X$$

(iii) **Union bound.** *We admit that there exists a $(1/2)$ -net of S^{d-1} of cardinality bounded by 4^d . Deduce that*

$$\mathbf{P} (\|X\|_2 \geq t) \leq 4^d e^{-t^2/(8\sigma^2)}$$

(iv) **Conclusion.** *Deduce that*

$$\mathbf{P} (\|X\|_2 \geq t) \leq 2e^{-t^2/(16d\sigma^2)}.$$

In other words, $\|X\|_2$ satisfies the same two-sided bound as a $(8d\sigma^2)$ -subgaussian random variable (the dependency in d and σ here cannot be improved).

*Lénaïc Chizat EPFL lenaic.chizat@epfl.ch

Exercise 4. Grönwall lemma is ubiquitous in the analysis of dynamical systems, so it is important to be very comfortable with this result. Here we look at the discrete version.

Lemma 1 (Discrete Grönwall's lemma). Let (u_n) and (w_n) be nonnegative sequences satisfying

$$u_n \leq \alpha + \sum_{k=0}^{n-1} u_k w_k \quad \forall n \geq 0.$$

Then for all $n \geq 0$ it holds

$$u_n \leq \alpha \exp\left(\sum_{k=0}^{n-1} w_k\right).$$

Prove the lemma by the following steps:

(i) Verify the identity

$$1 + \sum_{k=0}^{n-1} \left(\prod_{l=0}^{k-1} (1 + w_l) \right) w_k \leq \prod_{k=0}^{n-1} (1 + w_k).$$

(ii) Prove by induction that for all n it holds

$$u_n \leq \alpha \prod_{k=0}^{n-1} (1 + w_k).$$

(iii) Deduce the lemma.