

High-Dimensional Probability: Answers, Theorems, and Definitions

Last revised on May 28, 2021

- Companion notes for *High-Dimensional Probability*, by Roman Vershynin. Link to book (PDF available online): www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.html.
- **Disclaimer:** These notes compile my answers to the exercises, and lift the required theorems and definitions from the book. I wrote these notes to aid my personal study of the book. Read them at your own risk!*

Contents

0	Appetizer: Using probability to cover a geometric set	2
1	Preliminaries on random variables	3
1.1	Basic quantities	3
1.2	Inequalities	3
1.3	Limits of random variables	3

*Scribe: Alex Bie, alexbie98@gmail.com.

0 Appetizer: Using probability to cover a geometric set

A point $x \in \mathbb{R}^n$ is a **convex combination** of points $x_1, \dots, x_m \in \mathbb{R}^n$ if

$$x = \sum_{i=1}^m \lambda_i x_i \quad \text{with each } \lambda_i \geq 0 \quad \text{and} \quad \sum_{i=1}^m \lambda_i = 1.$$

The **convex hull** of $T \subseteq \mathbb{R}^n$, $\text{conv}(T)$, is the set of all convex combinations of T .

Theorem 0.0.1 (Carathéodory's Theorem). Let $x \in \text{conv}(T)$. There exists $k \leq n + 1$ points $x_1, \dots, x_k \in T$ such that x is a convex combination of x_1, \dots, x_k .

The result says we can obtain any point in the convex hull of T using at most a dimension-dependent number of points. Let the **diameter** of a set T be defined as $\text{diam}(T) = \sup\{\|x - y\|_2 : x, y \in T\}$.

Theorem 0.0.2 (Approximate Carathéodory's Theorem). Let $\text{diam}(T) = 1$. Let $x \in \text{conv}(T)$. For any k , there exists k points x_1, \dots, x_k such that

$$\left\| x - \frac{1}{k} \sum_{j=1}^k x_j \right\|_2 \leq \frac{1}{\sqrt{k}}$$

Proof. Suppose $|T| = m$. We write $x = \sum_{i=1}^m \lambda_i x_i$, and interpret λ_i as probabilities. We define the random variable

$$X = x_i \text{ with probability } \lambda_i$$

for $i = 1, \dots, m$. We can verify that $\mathbb{E} X = \sum_{i=1}^m \lambda_i x_i = x$. Taking $X_1, \dots, X_k \stackrel{\text{iid}}{\sim} X$. It remains to analyse the quantity $\mathbb{E} \|x - \frac{1}{k} \sum_{j=1}^k X_j\|_2^2$.

$$\begin{aligned} \mathbb{E} \left\| x - \frac{1}{k} \sum_{j=1}^k X_j \right\|_2^2 &\leq \frac{1}{k^2} \mathbb{E} \left\| \sum_{j=1}^k X_j - x \right\|_2^2 \\ &= \frac{1}{k^2} \sum_{j=1}^k \mathbb{E} \|X_j - x\|_2^2 \quad \text{by Exercise 0.0.3 (a)} \end{aligned}$$

□

Exercise 0.0.3. Check the following identities for random vectors.

(a) Let X_1, \dots, X_k be independent, mean zero random vectors in \mathbb{R}^n . Show that

$$\mathbb{E} \left\| \sum_{j=1}^k X_j \right\|_2^2 = \sum_{j=1}^k \mathbb{E} \|X_j\|_2^2$$

Solution.

□

(b) Let X be a random vector in \mathbb{R}^n . Show that

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

Solution.

□

1 Preliminaries on random variables

1.1 Basic quantities

The **expectation** of a random variable X is denoted as $\mathbb{E}X$, and **variance** is denoted as $\text{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2$. (We note that the expectation operator \mathbb{E} can be directly defined as the Lebesgue integral of the random variable $X : \Omega \rightarrow \mathbb{R}$ in the probability space (Ω, M, μ) .)

The **p-th moment** of X is given by $\mathbb{E}X^p$. We also let $\|X\|_p = (\mathbb{E}X^p)^{\frac{1}{p}}$ denote the **p-norm** of X . For $p = \infty$, we have

$$\|X\|_\infty = \text{ess sup } X$$

recalling that the **essential supremum** of a function f is the "smallest value γ such that $\{\omega \in \Omega : |f(\omega)| > \gamma\}$ has measure 0".

From this, we can define the **L^p spaces**[†], given a probability space (Ω, M, μ)

$$L^p = \{X : \|X\|_p < \infty\}$$

Results from measure and integration theory tell us that the $(L^p, \|\cdot\|_p)$ are complete. In the case of L^2 , we have that with the inner product

$$\begin{aligned} \langle X, Y \rangle &= \int_{\Omega} XY(\omega) \mu(\omega) \\ &= \mathbb{E}XY \end{aligned}$$

$(L^2, \langle \cdot, \cdot \rangle)$ is a Hilbert space. In this case we can express the **standard deviation** of X as $\sqrt{\text{Var}(X)} = \|X - \mathbb{E}X\|_2$, and the **covariance** of random variable X and Y as

$$\text{Cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y) = \langle X - \mathbb{E}X, Y - \mathbb{E}Y \rangle$$

Remark I. In this setting, considering random variables as vectors in L^2 , the covariance between X and Y can be interpreted as the *alignment* between the vectors $X - \mathbb{E}X$ and $Y - \mathbb{E}Y$.

1.2 Inequalities

1.3 Limits of random variables

[†]A technical note is that the objects of L_p are actually equivalence classes of functions $[X]$ with equality almost everywhere, otherwise $\|\cdot\|_p$ is only a semi-norm.