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

*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 $x_1, ..., x_m \in \mathbb{R}^n$ if

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

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

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

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 $\mathrm{diam}(T) = \sup\{\|x-y\|_2 : x, y \in T\}$.

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

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

Proof. Suppose $|T| = m$. WLOG we can assume T is bounded by 1 in $\|\cdot\|_2$. 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 $i = 1, ..., m$. We can verify that $\mathbb{E}X = \sum_{i=1}^{m} \lambda_i x_i = x$. Taking $X_1, ..., X_k \overset{\text{iid}}{\sim} X$. It remains to analyse the quantity $\mathbb{E}\|x - \frac{1}{k}\sum_{j=1}^{k} X_j\|_2^2$.

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

Applying the result of Exercise 0.0.3 (b), we obtain

$$\mathbb{E}\|X - x\|_2^2 = \mathbb{E}\|X\|_2^2 - \|\mathbb{E}X\|_2^2 \leq \mathbb{E}\|X\|_2^2 \leq 1$$

Plugging this in above, we obtain the desired bound in expectation, hence there must exist a realization of the X_j, $x_1, ..., x_k$, such that the bound holds.

□

Exercise 0.0.3. Check the following identities for random vectors.

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

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

Solution.

$$\begin{aligned}
\mathbb{E}\left\|\sum_{j=1}^{k} X_j\right\|_2^2 = \sum_{i=1}^{n}\mathbb{E}\left(\sum_{j=1}^{m} X_j^{(i)}\right)^2 &= \sum_{i=1}^{n}\mathrm{Var}\left(\sum_{j=1}^{m} X_j^{(i)}\right) && \text{(by mean zero)}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}\mathrm{Var}\left(X_j^{(i)}\right) && \text{(by independence)}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}\mathbb{E}\left(X_j^{(i)}\right)^2 && \text{(by mean zero)}\\
&= \mathbb{E}\sum_{j=1}^{m}\|X_j\|_2^2
\end{aligned}$$

□

Among other things, this result implies that the expected squared distance of a random walk (starting from the origin) is equal to sum of the expected squared distances of each step.

(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.

$$\begin{aligned}
\mathbb{E}\|X - \mathbb{E}X\|_2^2 = \mathbb{E}\sum_{i=1}^{n}\left(X^{(i)} - (\mathbb{E}X)^{(i)}\right)^2 = \sum_{i=1}^{n}\mathrm{Var}(X^{(i)}) &= \sum_{i=1}^{n}\mathbb{E}\left(X^{(i)}\right)^2 - \left(\mathbb{E}X^{(i)}\right)^2\\
&= \mathbb{E}\|X\|_2^2 - \|\mathbb{E}X\|_2^2
\end{aligned}$$

□

1 Preliminaries on random variables

1.1 Basic quantities

The **expection** of a random variable X is denoted as $\mathbb{E}X$, and **variance** is denoted as $\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2$. (We note that the expectation operator $\mathbb{E}$ can be directly defined as the Lebesgue integral of the random variable $X : \Omega \to \mathbb{R}$ 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

$$\|X\|_\infty = \operatorname{ess\,sup} X$$

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 $\sqrt{\mathrm{Var}(X)} = \|X - \mathbb{E}X\|_2$, and the **covariance** of random variable X and Y as

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

Remark I. n 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.