Homework 6
HDP KNU+ Fall 2022

Hints are in the back of this homework set.

As in the previous homework sets, $C, C_1, C_2, \dots$ and $c, c_1, c_2, \dots$ denote positive absolute constants of your choice.

Life in high dimensions is full of surprises. From linear algebra we know that the space $\mathbb{R}^n$ can not accommodate more than n orthogonal vectors. However, $\mathbb{R}^n$ can accommodate *exponentially* many *almost* orthogonal vectors, for large n. This is another manifestation of how much more room there is in high-dimensional worlds than in our three-dimensional world.

Problem 1 (Exponentially many almost orthogonal vectors)

(a) Let $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$ be independent symmetric Bernoulli random vectors in $\mathbb{R}^n$, i.e. all of the coordinates X_i and Y_i are independent random variables that take values ± 1 with probability $1/2$. Show that

$$\mathbb{P}\left\{ |\langle X, Y \rangle| \geq 0.001n \right\} \leq 2\exp(-c_1 n).$$

(b) Deduce that the angle $\angle(X, Y)$ between the vectors X and Y satisfies

$$\mathbb{P}\left\{ |\angle(X, Y) - \pi/2| > 0.01\pi \right\} \leq 2\exp(-c_1 n).$$

(c) Prove that for every dimension $n > C$, there exist $N \geq \frac{1}{2}\exp(cn)$ vectors $v_1, \dots, v_N$ in $\mathbb{R}^n$ so that all pairwise angles between these vectors are between 89° and 91°.

Our proof of Johnson-Lindenstrauss lemma, given in Lecture 15 (October 5), utilizes a *Gaussian random matrix* – a matrix whose entries are $N(0, 1)$ – that projects the data points onto a space of lower dimension. Here you will check that a *Bernoulli random matrix* – a matrix with ± 1 entries – works as well. Bernoulli matrices take less memory to store: one bit per entry, so they are preferred in practice. The result you are about to prove in part (c) was first established by D. Achlioptas[1] in 2003.

Problem 2 (Johnson-Lindenstrauss with binary coins)

Let G be an $n \times d$ Bernoulli random matrix – a matrix whose entries are i.i.d. symmetric Bernoulli random variables (i.e. each entry takes values ± 1 with probability $1/2$).

[1] D. Achlioptas, *Database-friendly random projections: Johnson-Lindenstrauss with binary coins*, Journal of Computer and System Sciences, Volume 66, Issue 4, June 2003, Pages 671-687.

(a) Let z be a fixed unit vector in $\mathbb{R}^d$. Show that the $X = Gz$ is a random vector in $\mathbb{R}^n$ whose all coordinates X_j are independent random variables, which satisfy

$$\mathbb{E}\, X_j = 0, \quad \mathrm{Var}(X_j) = 1, \quad \|X_j\|_{\psi_2} \le C_1.$$

(b) Prove a thin-shell inequality for $X = Gz$:

$$\mathbb{P}\left\{0.99\sqrt{n} \le \|X\|_2 \le 1.01\sqrt{n}\right\} \ge 1 - 2\exp(-cn).$$

(c) Let $x_1, \dots, x_N$ be any set of fixed vectors in $\mathbb{R}^d$. Let G be an $n \times d$ Bernoulli random matrix, and set $T = \frac{1}{\sqrt{n}} G$. Prove that if $n = C \log N$, then the map T approximately preserves the pairwise geometry of the data, namely that following event holds with positive probability:

$$0.99\left\|x_i - x_j\right\|_2 \le \left\|Tx_i - Tx_j\right\|_2 \le 1.01\left\|x_i - x_j\right\|_2 \quad \text{for all } i, j = 1, \dots, N.$$

The *Gram matrix* of a system of vectors $v_1, \dots, v_n$ in $\mathbb{R}^d$ is defined as the $n \times n$ matrix G whose entries are the inner products between the vectors, i.e. $G_{ij} = \langle v_j, v_j \rangle$.

Problem 3 (Gram matrices)

(a) Check that the Gram matrix G of any system of vectors is positive semidefinite.

(b) Conversely, prove that any $n \times n$ positive semidefinite matrix G is a Gram matrix of some system of vectors $v_1, \dots, v_n$ in $\mathbb{R}^n$.

The following key fact was used in the proof of Grothendieck's identity (Lecture 17, October 9).

Problem 4 (Grothendieck's identity)

Let θ be a random vector uniformly distributed on the unit circle $S^1 = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$. Prove that for any pair of vectors $u, v \in S^1$, we have

$$\mathbb{E}\, \mathrm{sign}\left(\langle u, \theta \rangle\right) \mathrm{sign}\left(\langle v, \theta \rangle\right) = \frac{2}{\pi} \arcsin\left(\langle u, v \rangle\right).$$

TURN OVER FOR HINTS

Hints

Hint for Problem 1. (a) Express $\langle X, Y\rangle$ as a sum. Note that the terms of the sum are independent symmetric Bernoulli random variables. Apply Hoeffding's inequality.

(b) Express the cosine of the angle via the inner product.

(c) Take the union bound over all N^2 pairs of vectors and use part (b). This logic is similar to step 2 of our proof of Johnson-Lindenstrauss lemma (Lecture 15, October 5).

Hint for Problem 2. (a) Matrix-vector multiplication allows us to express X_j as a sum. The terms of the sum are independent symmetric Bernoulli random variables multiplies by weights z_j. Note that each term has subgaussian norm bounded by $\left|z_j\right|$, and apply subgaussian Hoeffding's inequality (Lecture 14, October 3.)

(b) Argue like in the proof of thin-shell inequality for normal distribution (Lecture 15, October 5).

(c) Argue like in our proof of Johnson-Lindenstrauss lemma (Lecture 15, October 5).

Hint for Problem 3. There are several ways to prove (b) using linear algebra. For example, consider the spectral decomposition $G = \sum_{i=1}^n \lambda_i u_i u_i^\mathsf{T}$ and define the matrix $V = \sum_{i=1}^n \sqrt{\lambda_i} u_i u_i^\mathsf{T}$. (Why can we take the square root?) Then check that $G = V^2$; for this reason V is commonly called the square root of V Then G is theGram matrix of the rows of V (check).

Hint for Problem 4. The circle S^1 decomposes into four arcs depending on the value of $f(\theta) = \operatorname{sign}(\langle u, \theta\rangle) \operatorname{sign}(\langle v, \theta\rangle)$. Two arcs give value 1 and the other two, -1. Since θ is uniformly distributed on S^1, the expectation is the sum of the (normalized) length of the first two arcs minus the (normalized) length of the other two arcs.