Homework 2
High-Dimensional Probability for Data Science, Fall 2023

Hints are in the back of the homework set.

Problems in this set are about *random vectors*, i.e. vectors $X = (X_1, \ldots, X_n)$ in $\mathbb{R}^n$ whose coordinates X_i are random variables. The coordinates X_i are not necessarily independent. The expected value of a random vector is defined coordinate-wise, that is $\mathbb{E}\, X = (\mathbb{E}\, X_1, \ldots, \mathbb{E}\, X_n)$.

We will use the standard notation of linear algebra, where $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ denotes the inner product on $\mathbb{R}^n$, and $\|x\|_2 = \sqrt{\langle x, x\rangle} = (\sum_{i=1}^n x_i^2)^{1/2}$ denotes the Euclidean norm on $\mathbb{R}^n$.

Let us check that the standard properties of expectation of random variables, which you studied in an introductory probability course, pass on to random vectors.

Problem 1 (Expectation of random vectors)

Let X and Y be random vectors in $\mathbb{R}^n$.

(a) (Linearity) Prove that for any constants $a, b \in \mathbb{R}$, we have

$$\mathbb{E}(aX + bY) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y).$$

(b) (Multiplicativity) Prove that if X and Y are independent, then

$$\mathbb{E}\langle X, Y\rangle = \langle \mathbb{E}\, X, \mathbb{E}\, Y\rangle.$$

This generalizes the identity $\mathbb{E}(XY) = (\mathbb{E}\, X)(\mathbb{E}\, Y)$ for independent random variables.

Let us extend Problem 1 from Homework 1 for random vectors.

Problem 2 (Expectation minimizes the MSE)

Let X be a random vector with finite expectation. Show that the function $f(a) = \mathbb{E}\|X - a\|_2^2$ is minimized at $a = \mathbb{E}\, X$.

Recall that the variance of a random variable X satisfies

$$\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}\, X)^2 = \mathbb{E}(X^2) - (\mathbb{E}\, X)^2 = \frac{1}{2}\,\mathbb{E}(X - X')^2 \tag{1}$$

where X' is an independent copy of X. (The latter means that the random variables X and X' are independent and identically distributed.) The first equality in (1) is the definition of variance, the second equality follows by expanding the square and simplifying (try it!) and the third equality was proved in Lecture 2. Let us extend (1) for random vectors.

Problem 3 (Extending variance for random vectors)

For a random vector in $\mathbb{R}^n$, define

$$V(X) = \mathbb{E}\|X - \mathbb{E}\,X\|_2^2 \,. \tag{2}$$

(a) Prove that

$$V(X) = \mathbb{E}\|X\|_2^2 - \|\mathbb{E}\,X\|_2^2 \,.$$

(b) Prove that

$$V(X) = \frac{1}{2}\,\mathbb{E}\,\|X - X'\|_2^2 \,,$$

where X' is an independent copy of X.

Recall the additivity property of variance: if random variables $X_1, \ldots, X_k$ are independent, then

$$\mathrm{Var}(X_1 + \cdots + X_k) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_k).$$

Let us extend this property for random vectors.

Problem 4 (Additivity of variance)

(a) Check that if X and Y are independent random vectors in $\mathbb{R}^n$ with zero means, then

$$\mathbb{E}\|X + Y\|_2^2 = \mathbb{E}\|X\|_2^2 + \mathbb{E}\|Y\|_2^2 \,.$$

(b) Prove that if random vectors $X_1, \ldots, X_k$ in $\mathbb{R}^n$ are independent, then

$$V(X_1 + \cdots + X_k) = V(X_1) + \cdots + V(X_k),$$

where $V(\cdot)$ is defined in (2).

The diameter of a set $T \subset \mathbb{R}^n$ is defined as the largest distance between any two points in T. The radius of T is defined as the smallest radius of a Euclidean ball that contains T. In other words,

$$\mathrm{diam}(T) = \sup_{x,y \in T} \|x - y\|_2 \,, \quad \mathrm{rad}(T) = \inf_{a \in \mathbb{R}^n} \sup_{x \in T} \|x - a\|_2 \,.$$

Problem 5 (Radius and diameter)

(a) Check that
$$\mathrm{rad}(T) \le \mathrm{diam}(T) \le 2\,\mathrm{rad}(T).$$
(b) Prove a version of the Approximate Caratheodory Theorem (Lecture 2) under the assumption that $\mathrm{rad}(T) \le 1$ instead of $\mathrm{diam}(T) \le 1$ and with the bound $1/\sqrt{k}$ instead of $1/(2\sqrt{k})$.

TURN OVER FOR HINTS

Hints

Hints for Problem 2.

Extend the identity mentioned in the hint for Problem 1 (Homework 1) for random vectors.

Hints for Problem 3.

(a) Since $\|x\|_2^2 = \langle x, x \rangle$ for any $x \in \mathbb{R}^n$, we have $\|X - \mathbb{E}\, X\|_2^2 = \langle X - \mathbb{E}\, X, X - \mathbb{E}\, X \rangle$. Expand this inner product into four terms and simplify.

(b) We proved this for random variables in Lecture 2; proceed similarly for random vectors.

Hints for Problem 4.

If $k = 2$ and all X_i have zero mean, part (b) reduces to part (a). To deduce part (b) in full generality, you may use induction on k, and consider $X_i' = X_i - \mathbb{E}\, X_i$ that have zero mean.

Hints for Problem 5.

(a) To prove the bound $\operatorname{diam}(T) \le 2 \operatorname{rad}(T)$, use triangle inequality.

(b) The only difference comes in the end of the proof, where we need to bound $\mathbb{E}\|Z - \mathbb{E}\, Z\|_2^2$. Use the result of Problem 2.