Homework 2 - High Dimensional Probability

Bernardi Alberto, Cracco Gianmarco,
Manente Alessandro, Mazzieri Riccardo

December 2020

Exercise 1

Deduce from Theorem 3.1.1[1], that:

$$\sqrt{n} - CK^2 \leq \mathbb{E}[\|X\|_2] \leq \sqrt{n} + CK^2. \tag{1}$$

Solution:

From Thm. 3.1.1 it follows:

$$\mathbb{E}\left[\exp\frac{(\|X\|_2 - \sqrt{n})^2}{(CK^2)^2}\right] \leq 2$$

$$\log\left(\mathbb{E}\left[\exp\frac{(\|X\|_2 - \sqrt{n})^2}{(CK^2)^2}\right]\right) \leq \log 2 \leq 1. \quad \star$$

On the other hand,

$$\begin{aligned}\log\left(\mathbb{E}\left[\exp\frac{(\|X\|_2 - \sqrt{n})^2}{(CK^2)^2}\right]\right) &\geq \mathbb{E}\left[\frac{(\|X\|_2 - \sqrt{n})^2}{(CK^2)^2}\right] \\ &\geq \mathbb{E}^2\left[\frac{\|X\|_2 - \sqrt{n}}{CK^2}\right] \\ &= \frac{1}{(CK^2)^2}\left(\mathbb{E}\left[\|X\|_2 - \sqrt{n}\right]\right)^2\end{aligned}$$

where for the first inequality we used Jensen's inequality for concave functions, while for the second one we used Jensen's inequality for convex functions.

[1]Let $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ be a random vector with independent, sub-gaussian coordinates X_i that satisfy $\mathbb{E}[X_i^2] = 1$ then: $\left\|\|X\|_2 - \sqrt{n}\right\|_{\psi_2} \leq CK^2$ where $K = \max_i \|X_i\|_{\psi_2}$ and C is an absolute constant.

Going back to $\star$,

$$\frac{1}{(CK^2)^2}\left(\mathbb{E}\left[\|X\|_2-\sqrt{n}\right]\right)^2 \leq 1$$
$$\Updownarrow$$
$$-1 \leq \frac{1}{CK^2}\left(\mathbb{E}\left[\|X\|_2-\sqrt{n}\right]\right) \leq 1$$
$$\Updownarrow$$
$$\sqrt{n}-CK^2 \leq \mathbb{E}[\|X\|_2] \leq \sqrt{n}+CK^2.$$

■

Exercise 2

Deduce from Theorem 3.1.1, that:

$$\operatorname{Var}\left(\|X\|_2\right) \leq CK^4.$$

Solution:

$$\begin{aligned}
\left\|\|X\|_2-\mathbb{E}\|X\|_2\right\|_{\psi_2} &= \left\|\|X\|_2-\sqrt{n}-\left(\mathbb{E}\|X\|_2-\sqrt{n}\right)\right\|_{\psi_2} \\
&\leq \left\|\|X\|_2-\sqrt{n})\right\|_{\psi_2}+\left\|\left(\mathbb{E}\|X\|_2-\sqrt{n}\right)\right\|_{\psi_2} \\
&\overset{\star}{\leq} \left(1+\frac{1}{\sqrt{\log 2}}\right)CK^2 = CK^2,
\end{aligned}$$

where the last inequality follows from $\left\|\|X\|_2-\sqrt{n}\right\|_{\psi_2} \leq CK^2$ (from Thm. 3.1.1) and $\left\|\mathbb{E}\left[\|X\|_2\right]-\sqrt{n}\right\|_{\psi_2} \leq \frac{CK^2}{\sqrt{\log 2}}$.
This last one, in particular, comes from:

$$2 \geq \mathbb{E}\left[\exp\left(\frac{\left(\mathbb{E}\|X\|_2-\sqrt{n}\right)^2}{t^2}\right)\right] = \exp\left(\frac{\left(\mathbb{E}\|X\|_2\sqrt{n}\right)^2}{t^2}\right)$$

$$\implies t \geq \frac{|\mathbb{E}\|X\|_2-\sqrt{n}|}{\sqrt{\log 2}}$$

$$\implies \left\|\mathbb{E}\|X\|_2-\sqrt{n}\right\|_{\psi_2} = \frac{|\mathbb{E}\|X\|_2-\sqrt{n}|}{\sqrt{\log 2}} \leq \frac{CK^2}{\sqrt{\log 2}}$$

using the result from Ex. 1.

From $\star$, the definition of $\|\cdot\|_{\psi_2}$ and the Jensen's inequality we get:

$$2 \geq \mathbb{E}\left[\exp\left(\frac{(\|X\|_2 - \mathbb{E}\|X\|_2)^2}{CK^4}\right)\right] \geq \exp\left(\frac{\mathbb{E}\left[(\|X\|_2 - \mathbb{E}\|X\|_2)^2\right]}{CK^4}\right)$$

$$\implies \mathrm{Var}\left(\|X\|_2\right) = \mathbb{E}\left[(\|X\|_2 - \mathbb{E}\|X\|_2)^2\right] \leq \log 2 \cdot CK^4 = CK^4.$$

■

Exercise 3

Let $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ be a random vector with independent coordinates X_i that satisfy $\mathbb{E}X_i^2 = 1$ and $\mathbb{E}X_i^4 \leq K^4$. Show that

$$\mathrm{Var}\left(\|X\|_2\right) \leq CK^4.$$

Solution:

We first show that:

$$\mathbb{E}\left(\|X\|_2^2 - n\right)^2 \leq K^4 n.$$

Indeed, since we have:

$$\mathbb{E}\|X\|_2^4 = \mathbb{E}\left(\sum_{i=1}^n X_i^2\right)^2 = \mathbb{E}\left[\sum_{i=1}^n X_i^4 + 2\sum_{i=1}^n\sum_{j=1}^{i-1} X_i^2 X_j^2\right] =$$

$$= \sum_i \mathbb{E}X_i^4 + 2\sum_{i=1}^n\sum_{j=1}^{i-1} \mathbb{E}X_i^2\mathbb{E}X_j^2 \leq nK^4 + 2\binom{n}{2} = nK^4 + n^2 - n,$$

we obtain:

$$\mathbb{E}\left(\|X\|_2^2 - n\right)^2 = \mathbb{E}\left[\|X\|_2^4 + n^2 - 2n\|X\|_2^2\right] =$$

$$= \mathbb{E}\|X\|_2^4 + n^2 - 2n\mathbb{E}\|X\|_2^2 = \mathbb{E}\|X\|_2^4 - n^2 \leq nK^4 - n \leq nK^4.$$

From this we can write:

$$\frac{\mathbb{E}\left(\|X\|_2^2 - n\right)^2}{n^2} \leq \frac{nK^4}{n^2} \iff \mathbb{E}\left(\frac{\|X\|_2^2}{n} - 1\right)^2 \leq \frac{K^4}{n}.$$

Now we can use the fact that $|x-1| \leq |x^2 - 1| \quad \forall x \geq 0$ to obtain:

$$\mathbb{E}\left(\frac{\|X\|_2}{\sqrt{n}} - 1\right)^2 = \frac{1}{n}\mathbb{E}\left(\|X\|_2 - \sqrt{n}\right)^2 \leq \mathbb{E}\left(\frac{\|X\|_2^2}{n} - 1\right)^2 \leq \frac{K^4}{n},$$

that implies

$$\mathbb{E}\left(\|X\|_2 - \sqrt{n}\right)^2 \leq K^4.$$

In order to conclude, we just need to show that:

$$\mathrm{Var}\left(\|X\|_2\right) \leq \mathbb{E}\left(\|X\|_2 - \sqrt{n}\right)^2.$$

However, this holds true since:

$$\begin{aligned}
\mathbb{E}\left(\|X\|_2 - \mathbb{E}\|X\|_2\right)^2 &= \mathbb{E}\|X\|_2^2 - 2\mathbb{E}\|X\|_2\mathbb{E}\|X\|_2 + \mathbb{E}^2\|X\|_2 = \mathbb{E}\|X\|_2^2 - \mathbb{E}^2\|X\|_2 \\
\mathbb{E}\left(\|X\|_2 - \sqrt{n}\right)^2 &= \mathbb{E}\left[\|X\|_2^2 - 2\sqrt{n}\|X\|_2 + n\right] = \mathbb{E}\|X\|_2^2 - 2\sqrt{n}\mathbb{E}\|X\|_2 + n \\
\implies \mathrm{Var}\left(\|X\|_2\right) &\leq \mathbb{E}\left(\|X\|_2 - \sqrt{n}\right)^2 \iff \mathbb{E}^2\|X\|_2 - 2\sqrt{n}\mathbb{E}\|X\|_2 + n \geq 0,
\end{aligned}$$

which is true since the last term is a square. The proof is hence complete, with $C = 1$.

■

Exercise 4

Let $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ be a random vector with independent coordinates X_i with continuous distributions. Assume that the densities of X_i are uniformly bounded by 1. Show that, for any $\epsilon > 0$, we have

$$\mathbb{P}\left\{\|X\|_2 \leq \epsilon\sqrt{n}\right\} \leq (C\epsilon)^n.$$

Solution:

$$\begin{aligned}
\mathbb{P}\left\{\|X\|_2 \leq \epsilon^2 n\right\} &= \mathbb{P}\left\{\exp\left(-\lambda\sum_{i=1}^n \frac{{X_i}^2}{\epsilon^2}\right) \geq \exp\left(-\lambda n\right)\right\} \\
&\leq e^{\lambda n}\mathbb{E}\left[\exp\left(-\lambda\sum_{i=1}^n \frac{{X_i}^2}{\epsilon^2}\right)\right] \\
&= e^{\lambda n}\prod_{i=1}^n \mathbb{E}\left[\exp\left(-\lambda\frac{{X_i}^2}{\epsilon^2}\right)\right].
\end{aligned}$$

Now notice that, for any $i = 1, \ldots, n$:

$$\mathbb{E}\left[\exp\left(-\lambda\frac{{X_i}^2}{\epsilon^2}\right)\right] = \int_{-\infty}^{+\infty} e^{-\lambda\frac{x^2}{\epsilon^2}} p(x)dx = \int_{-\infty}^{+\infty} e^{-\lambda\frac{x^2}{\epsilon^2}} dx = \epsilon\sqrt{\frac{\pi}{\lambda}}.$$

Setting $\lambda = \pi$ and using the first chain of inequalities we considered, we get:

$$\mathbb{P}\left\{\|X\|_2 \leq \epsilon^2 n\right\} \leq e^{\pi n} \cdot \epsilon^n = (e^\pi \cdot \epsilon)^n = (C \cdot \epsilon)^n.$$

■

Exercise 5

Let X and Y be independent, mean-zero, isotropic random vectors in $\mathbb{R}^n$. Check that:

$$\mathbb{E}\left[\|X-Y\|_2^2\right] = 2n.$$

Solution:

$$\begin{aligned}
\mathbb{E}\left[\|X-Y\|_2^2\right] = \mathbb{E}\left[\sum_{i=1}^n (X_i - Y_i)^2\right] &= \sum_{i=1}^n \mathbb{E}\left[(X_i-Y_i)^2\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[X_i^2 + Y_i^2 - 2X_iY_i\right] \\
&= \sum_{i=1}^n \underbrace{\mathbb{E}\left[X_i^2\right]}_{=1} + \underbrace{\mathbb{E}\left[Y_i^2\right]}_{=1} - 2 \underbrace{\mathbb{E}\left[X_iY_i\right]}_{=\mathbb{E}[X_i]\cdot\mathbb{E}[Y_i]=0} = 2n,
\end{aligned}$$

where in the last equality we applied both isotropy and the independence of X and Y. ■

Exercise 6

Prove that $X \sim \mathcal{N}(0, \mathbf{1}_n)$ is isotropic.

Solution:

$$\mathrm{Cov(X)} = \mathbf{1}_n = \mathbb{E}\left[XX^\top\right] - \mathbb{E}[X]^2 = \mathbb{E}\left[XX^\top\right] - 0 = \mathbb{E}\left[XX^\top\right]$$

■

Exercise 7

Let X be a random vector in $\mathbb{R}^n$. Prove that $X \sim \mathcal{N}(\mu, \Sigma)$ if and only if $\langle X, \theta\rangle$ has a normal distribution for all $\theta \in \mathbb{R}^n$.

Solution:

($\Longrightarrow$) X normal vector $\Longrightarrow$ $X = \mu + AZ$ where A is the matrix such that $\Sigma = AA^T$ and $Z \sim \mathcal{N}(0, \mathbf{1}_n)$. As a consequence, picking $\theta \in \mathbb{R}^n$

$$\langle X, \theta\rangle = \sum_{i=1}^n \theta_i X_i = \sum_{i=1}^n \theta_i(\mu_i + A_i Z) = \sum_{i=1}^n \mu_i\theta_i + \sum_{i=1}^n\sum_{j=1}^n \theta_i A_{ij} Z_j,$$

where A_i denotes the i-th column of A. We know that if $Z \sim \mathcal{N}(0, \mathbf{1}_n)$, then Z_i and Z_j are independent for any $i \neq j$ and hence the random variable $\sum_{i=1}^n \mu_i\theta_i + \sum_{i=1}^n \sum_{j=1}^n \theta_i A_{ij} Z_j$ will be normal, since it is the sum of a constant and a normal random variable (indeed, the linear combination of independent normals is still normal). Therefore, from the arbitrariety of the choice of θ we can conclude that $\langle X, \theta\rangle$ is normally distributed for any $\theta \in \mathbb{R}^n$.

($\Longleftarrow$) Let $\langle X, \theta \rangle$ be a random variable normally distributed $\forall \theta \in \mathbb{R}^n$. Assume you have Y normal random vector, then $\langle Y, \theta \rangle$ normal $\forall \theta \in \mathbb{R}^n$. Then, from the Cramer-Wold Theorem, we have that X is normally distributed.

■

Exercise 8

Let $X \sim \mathcal{N}(0, \mathbf{1}_n)$, then $X = \|X\|_2 \frac{X}{\|X\|_2}$. Prove that:

(i) $\|X\|_2$ and $Y = \frac{X}{\|X\|_2}$ are independent.

(ii) $\sqrt{n}Y \sim Unif(\sqrt{n}S^{n-1})$.

Solution:

(i) We observe that, for $r \in \mathbb{R}$ and $\theta \in \mathbb{R}^n$:

$$P\{\|X\|_2 = r, Y = \theta\} = P\{X = r\theta\} = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(\frac{-\|X\|_2^2}{2}\right) dx = \ldots$$

switching to polar coordinates we have:

$$\cdots = \frac{|S_{n-1}|}{(2\pi)^{n/2}} \exp\left(\frac{-r^2}{2}\right) r^{n-1} dr \frac{d\theta}{|S_{n-1}|},$$

where $d\theta$ is the surface element of S_{n-1} and $|S_{n-1}|$ is the surface of the sphere. We now claim that:

$$P\{\|X\|_2 = r\} = \frac{|S_{n-1}|}{(2\pi)^{n/2}} \exp\left(\frac{-r^2}{2}\right) r^{n-1} dr \tag{2}$$

and

$$P\{Y = \theta\} = \frac{d\theta}{|S_{n-1}|} \tag{3}$$

are densities. In such a case, in fact, we would have

$$P\{\|X\|_2 = r, Y = \theta\} = P\{\|X\|_2 = r\} P\{Y = \theta\},$$

which is the definition of independence.
For the first one we have to prove that:

$$\frac{|S_{n-1}|}{(2\pi)^{n/2}} \int_0^\infty \exp\left(\frac{-r^2}{2}\right) r^{n-1} dr = 1.$$

With the change of variable $x = \frac{r^2}{2}, dr = \frac{1}{\sqrt{2x}} dx$, we have

$$\frac{|S_{n-1}|}{(2\pi)^{n/2}}\int_0^\infty e^{-x}\,(2x)^{\frac{n-1}{2}}\,(2x)^{-\frac{1}{2}}\,dx = \frac{|S_{n-1}|}{(2\pi)^{n/2}}\int_0^\infty e^{-x}\,(2x)^{\frac{n}{2}-1}\,dx =$$
$$= \frac{|S_{n-1}|}{(2\pi)^{n/2}}2^{\frac{n}{2}-1}\underbrace{\int_0^\infty e^{-x}x^{\frac{n}{2}-1}dx}_{\Gamma(\frac{n}{2})=\frac{2\pi^{\frac{n}{2}}}{|S_{n-1}|}} = \frac{2^{\frac{n}{2}-1}\cdot 2}{2^{\frac{n}{2}}} = 1,$$

which proves that (2) is a density.
For the second one is easy to see that, being $d\theta$ the surface element of S_{n-1},

$$\int_{S_{n-1}}\frac{d\theta}{|S_{n-1}|} = 1,$$

which proves that also (3) is a density. This concludes the proof for point (i).

(ii) Statement follows from the previous result. In fact we have that, given Y random variable such that $P\{Y=\theta\} = \frac{d\theta}{|S_{n-1}|}$, we then have:

$$P\{\sqrt{n}Y=\theta\} = P\{Y=\frac{\theta}{\sqrt{n}}\} = \frac{d\frac{\theta}{\sqrt{n}}}{|S_{n-1}|} = \frac{d\theta}{\sqrt{n}|S_{n-1}|} = \frac{d\theta}{|\sqrt{n}S_{n-1}|},$$

which means that Y is uniformly distributed on $\sqrt{n}S_{n-1}$, which ends the proof.

■

Exercise 9

Let $X=(X_1,\dots,X_n)^T$ be a random vector with X_i sub-gaussian. Show that X is sub-gaussian.

Solution:
Given X and Y sub-gaussian random variables, we can prove that the sum $X+Y$ is still sub-gaussian. Indeed, setting $\hat{k} = \|X\|_{\psi_2} + \|Y\|_{\psi_2} = t_1 + t_2$, we have:

$$\begin{aligned}
\mathbb{P}\left\{|X+Y|\geq t\right\} &\leq \mathbb{P}\left\{\exp\left(\frac{(X+Y)^2}{\hat{k}^2}\right)\geq\exp\left(\frac{t^2}{\hat{k}^2}\right)\right\}\leq e^{-\frac{t^2}{\hat{k}^2}}\mathbb{E}\left[\exp\left(\frac{(X+Y)^2}{\hat{k}^2}\right)\right]\\
&\leq e^{-\frac{t^2}{\hat{k}^2}}\mathbb{E}\left[\frac{t_1}{t_1+t_2}\mathbb{E}\left[\exp\left(\frac{X^2}{{t_1}^2}\right)\right]+\frac{t_2}{t_1+t_2}\mathbb{E}\left[\exp\left(\frac{(\mathbb{E}X)^2}{{t_2}^2}\right)\right]\right]\\
&\leq 2\exp\left(-\frac{t^2}{\hat{k}^2}\right)\\
&\Longrightarrow \mathbb{P}\left\{|X+Y|\geq t\right\}\leq 2e^{-\frac{t^2}{\hat{k}^2}}.
\end{aligned}$$

Moreover, if X is a sub-gaussian random variable, then also αX is sub-gaussian for any $\alpha \in \mathbb{R}$. In fact:

$$\mathbb{P}\left\{|\alpha X| \geq t\right\} = \mathbb{P}\left\{|X| \geq \frac{t}{|\alpha|}\right\} \leq 2\exp\left(-\frac{ct^2}{|\alpha|^2 k^2}\right),$$

where in the last inequality we applied the definition of sub-gaussian random variable valid for X. Stated these two properties, it easily follows that $X = (X_1, \ldots, X_n)$ is a sub-gaussian random vector. Indeed, given $x \in \mathbb{R}^n$, we have $\langle X, x\rangle = \sum_{i=1}^n X_i x_i$, which is a linear combination of sub-gaussian random variables and hence it is a sub-gaussian random variable. We can conclude by the arbitrary choice of x.

■