Math Foundations of ML, Fall 2018

Homework #2

Due Friday September 7, at the beginning of class

As stated in the syllabus, unauthorized use of previous semester course materials is strictly prohibited in this course.

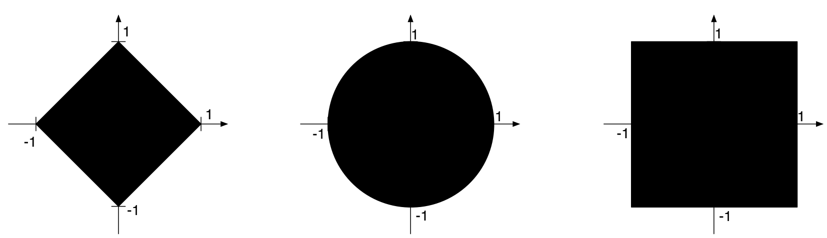
1. Using you class notes, prepare a 1-2 paragraph summary of what we talked about in class in the last week. I do not want just a bulleted list of topics, I want you to use complete sentences and establish context (Why is what we have learned relevant? How does it connect with other things you have learned here or in other classes?). The more insight you give, the better.

2. Prove the "reverse triangle inequality": show that in a normed linear space

$$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \ \leq\ \|\boldsymbol{x} - \boldsymbol{y}\|.$$

What can we say about $|\ \|\boldsymbol{x}\| - \|\boldsymbol{y}\|\ |$ in relation to $\|\boldsymbol{x} - \boldsymbol{y}\|$?

3. One way to visualize a norm in $\mathbb{R}^2$ is by its *unit ball*, the set of all vectors such that $\|\boldsymbol{x}\| \leq 1$. For example, we have seen that the unit balls for the ℓ_1, ℓ_2, and ℓ_∞ norms look like:



$B_1 = \{\boldsymbol{x} : \|\boldsymbol{x}\|_1 \leq 1\}$ $B_2 = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 \leq 1\}$ $B_\infty = \{\boldsymbol{x} : \|\boldsymbol{x}\|_\infty \leq 1\}$

Given an appropriate subset of the plane, $B \subset \mathbb{R}^2$, it might be possible to define a corresponding norm using

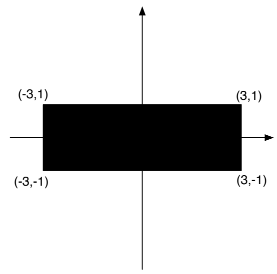
$$\|\boldsymbol{x}\|_B = \text{minimum value } r \geq 0 \text{ such that } \boldsymbol{x} \in rB, \tag{1}$$

where rB is just a scaling of the set B:

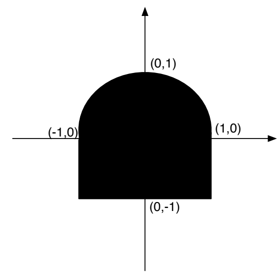
$$\boldsymbol{x} \in B \quad \Rightarrow \quad r \cdot \boldsymbol{x} \in rB.$$

(a) Let $\boldsymbol{x} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$. For $p = 1, 2, \infty$, find $r = \|\boldsymbol{x}\|_p$, and sketch $\boldsymbol{x}$ and rB_p (use different axes for each of the three values of p).

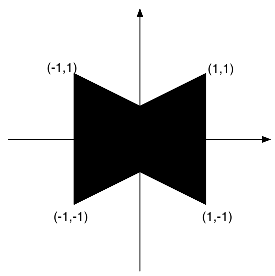
(b) Consider the 5 shapes below.



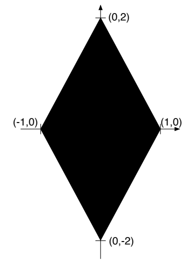
B_a



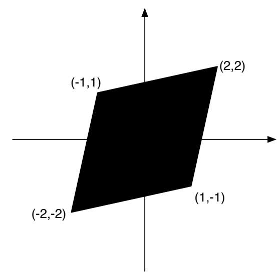
B_b



B_c



B_d



B_e

Explain why $\|\cdot\|_{B_b}$ and $\|\cdot\|_{B_c}$ are **not** valid norms. The most convincing way to do this is to find vectors for which one of the three properties of a valid norm are violated.

(c) Give a concrete method for computing $\|\boldsymbol{x}\|_{B_a}, \|\boldsymbol{x}\|_{B_d}$, and $\|\boldsymbol{x}\|_{B_e}$ for any given vector $\boldsymbol{x}$. (For example: for B_1, which corresponds to the ℓ_1 norm, we would write $\|\boldsymbol{x}\|_1 = |x_1| + |x_2|$.) Using you expressions, show that these are indeed valid norms. This will require a little bit of thought.

4. Below, $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^N$.

(a) Prove that $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\|_\infty \cdot \|\boldsymbol{y}\|_1$.

(b) Prove that $\|\boldsymbol{x}\|_1 \leq \sqrt{N} \cdot \|\boldsymbol{x}\|_2$. (Hint: Cauchy-Schwarz)

(c) Let B_2 be the unit ball for the ℓ_2 norm in $\mathbb{R}^N$. Fill in the right hand side below with an expression that depends only on $\boldsymbol{y}$:

$$\max_{\boldsymbol{x} \in B_2} \langle \boldsymbol{x}, \boldsymbol{y} \rangle = ????$$

Describe the vector $\boldsymbol{x}$ which achieves the maximum. Of course, this vector will depend on $\boldsymbol{y}$. (Hint: Cauchy-Schwarz)

(d) Let B_∞ be the unit ball for the ℓ_∞ norm in $\mathbb{R}^N$. Fill in the right hand side below with an expression that depends only on $\boldsymbol{y}$:

$$\max_{\boldsymbol{x} \in B_\infty} \langle \boldsymbol{x}, \boldsymbol{y} \rangle = ????$$

Describe the vector $\boldsymbol{x}$ which achieves the maximum. Of course, this vector will depend on $\boldsymbol{y}$. (Hint: Part (a))

(e) Let B_1 be the unit ball for the ℓ_1 norm in $\mathbb{R}^N$. Fill in the right hand side below with an expression that depends only on $\boldsymbol{y}$:

$$\max_{\boldsymbol{x} \in B_1} \langle \boldsymbol{x}, \boldsymbol{y} \rangle \quad = \quad ????$$

Describe the vector $\boldsymbol{x}$ which achieves the maximum. Of course, this vector will depend on $\boldsymbol{y}$. (Hint: Part (a))

(These last two might require some thought. If you solve them for $N = 2$, it should be easy to generalize.)

5. Let $\boldsymbol{A}$ be the 2×2 matrix

$$\boldsymbol{A} = \begin{bmatrix} 3 & 3 \\ -1/2 & 1/2 \end{bmatrix}.$$

For $\boldsymbol{x} \in \mathbb{R}^2$, define $\|\boldsymbol{x}\|_A = \|\boldsymbol{A}\boldsymbol{x}\|_2$. It should be clear that $\|\cdot\|_A$ is a valid norm; if it isn't, then convince yourself that it meets the required properties.

(a) Sketch the unit ball $B_A = \{\boldsymbol{x} : \|\boldsymbol{x}\|_A \leq 1\}$ corresponding to $\|\cdot\|_A$. Feel free to use MATLAB.

(b) For what kinds of matrices $\boldsymbol{A}$ is $\|\boldsymbol{x}\|_A = \|\boldsymbol{A}\boldsymbol{x}\|_2$ not a valid norm?

6. For a given matrix $N \times N$ matrix $\boldsymbol{Q}$, set

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_Q = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{x},$$

for vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$.

(a) Prove that if $\boldsymbol{Q}$ has an entry along its diagonal that is ≤ 0, then $\langle \cdot, \cdot \rangle_Q$ cannot be a valid inner product on $\mathbb{R}^N$.

(b) Prove that if $\boldsymbol{Q}$ is not symmetric, $\boldsymbol{Q}^{\mathrm{T}} \neq \boldsymbol{Q}$, then $\langle \cdot, \cdot \rangle_Q$ cannot be valid inner product on $\mathbb{R}^N$.

(c) A *symmetric positive definite matrix* (sym+def) is an $N \times N$ matrix $\boldsymbol{Q}$ that is symmetric ($\boldsymbol{Q} = \boldsymbol{Q}^{\mathrm{T}}$) and obeys

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{x} > 0, \quad \text{for all } \; \boldsymbol{x} \in \mathbb{R}^N, \; \boldsymbol{x} \neq \boldsymbol{0}.$$

Prove that $\langle \cdot, \cdot \rangle_Q$ is a valid inner product on $\mathbb{R}^N$ if and only if $\boldsymbol{Q}$ is symmetric positive definite.