Unit 2B: Norms and Balls

Prof.Dr.P.M.Bajracharya

School of Mathematical Sciences
T.U., Kirtipur

February 19, 2024

Summary

Norms in $\mathbb{R}^2$

Recall:

Euclidean norm

Given $x = (x_1, x_2) \in \mathbb{R}^2$, we define that

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2}.$$

It is also known as the L_2 norm. From the definition, it is clear that

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2}$$

is the distance from the origin to the point $x = (x_1, x_2)$. So, it is the length of the vector x. Here are the key properties of the length $\|x\|_2$.

Positivity/non-negativity. The length of a vector is always greater than 0, unless it is the zero vector, in which case its length is equal to 0.

Positive scalability. The length of product of a vector and a scalar real number is the length of the vector multiplied by the absolute value of the scalar.

Triangle inequality. The length of one side of a triangle is not larger than sum of the lengths of the other two sides of that triangle.

It is these three properties that motivated the definition of a *norm*, which is a generalization of the Euclidean norm in $\mathbb{R}^2$.

Norm

A mapping $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is called a **norm**, if it satisfies the following properties:

N1. $\|x\| \geq 0$ for all $x \in \mathbb{R}^n$.

N2. $\|x\| = 0$ if and only if $x = 0$,

N3. $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in \mathbb{R}^n$ and for all $\alpha \in \mathbb{R}$,

N4. $\|x + y\| \leq \|x\| + \|y\|$. (Triangle inequality)

We have defined L_2 norm on $\mathbb{R}^2$. But we can define a norm on $\mathbb{R}^2$ in many ways. The most popular norms on $\mathbb{R}^2$ are L_1 and L_∞. They are defined as follows:

$$L_1 : \|x\|_1 = \sum_{1}^{2} |x_i| = |x_1| + |x_2|.$$

$$L_\infty : \|x\|_\infty = \max_{i \in \{1,2\}} |x_i| = \max\{|x_1|, |x_2|\}.$$

These provide two different ways other than the L_2 norm to measure the size of a vector.

Let $x = (1, 2)$. Then

- $\|x\|_1 = 1 + 2 = 3$
- $\|x\|_2 = (1 + 4)^{1/2} \approx 2.24$
- $\|x\|_\infty = \max\{1, 2\} = 2.$

This example illustrates the more general property that

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1.$$

This means that the numerical value of the size of a vector depends on the norm used to measure its size.

Fact. For all of these inequalities, there exists vectors $x \in \mathbb{R}^2$ such that the inequality holds as an equality.

Question. Can you give an example of a vector $x \in \mathbb{R}^2$ such that each of the above inequalities holds as an equality?

Let $x = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$. Then

- $\|x\|_1 = |\cos\theta| + |\sin\theta|$.
- $\|x\|_2 = \cos^2\theta + \sin^2\theta$.
- $\|x\|_\infty = \max\{\cos\theta, \sin\theta\}$.

Similar statements hold for the vector
$x = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$.
These properties are not peculiar to $\mathbb{R}^2$.
Consider, e.g., a vector in $\mathbb{R}^3$.

Example

If we take $x = (1, 2, 3)$, then we have

- $\|x\|_1 = 6$.
- $\|x\|_2 = (1^2 + 2^2 + 3^2)^{1/2} \approx 3.71$.
- $\|x\|_\infty = 3$.

Again, this illustrates that $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$.

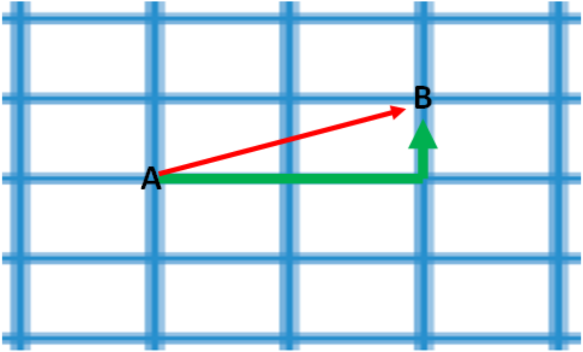


Figure: Roadmap of a city

Example

While we have defined L_1, L_2 and L_3 norms, it might be less easy to see – initially at least – why these other notions of size or norm are interesting. To see one example of this, consider the above figure, which illustrates a grid-like street map of a city.

Example ...

- Perhaps a bird can ignore the streets and fly directly from point A to point B, but to walk from point A to point B means that one must go along streets, horizontally and vertically, as opposed to along the hypotenuse.
- This is captured by the L_1 norm, and thus this notion of norm is a more meaningful notion of distance when walking around city blocks such as those shown in the above figure.

- The L_2 norm is important because it is associated with angles and dot products.
- This means, when we are measuring these quantities on the plane, there are two complementary perspectives: *algebra*; and *geometry*.
- Understanding these connections forms most of the basis for understanding high dimensional data.
- This connection will generalize to high dimensional Euclidean spaces.

Norms on $\mathbb{R}^n$

We now generaliaze L_2 or L_3 norms to $\mathbb{R}^n$ for a positve integer $n \geq 1$.

Points in $\mathbb{R}^n$

A point in $\mathbb{R}^n$ is an ordered list of n numbers, that is, n- tuple of real numbers and can be represented as

$$(x_1, x_2, ..., x_n) \text{ or } \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix}.$$

This is often interpreted as a point representing position.

We now generaliaze L_1, L_2 and L_∞ norms to $\mathbb{R}^n$ for a positve integer $n \geq 1$.

L_1, L_2 and L_∞ norms on $\mathbb{R}^n$

$$\|x\|_1 = \sum_1^n |x_i| = |x_1| + |x_2| + ... + |x_n|.$$

$$\|x\|_2 = \left(\sum_1^n x_i^2\right) = (x_1^2 + x_2^2 + ... + x_n^2)^{1/2}.$$

$$\|x\|_\infty = \max_{i\in\{1,2,...,n\}} |x_i| = \max\{|x_1|, |x_2|, ..., |x_n|\}.$$

- These three norms provide three different ways to measure the “size” of a vector in $\mathbb{R}^n$.
- Which norm is better is not the problem.
- Instead, you should ask what does a particular norm capture and when is it useful.
- All of these norms are useful in different situations.

Observe that all three norms reduce to the same thing in $\mathbb{R}$. More precisely, if $x \in \mathbb{R}$, then

$$\|x\|_1 = \|x\|_2 = \|x\|_\infty = |x|.$$

Example

1. Let $x = (1, 1, ..., 1) \in \mathbb{R}^n$. Then
$$\|x\|_1 = n, \ \|x\|_2 = \sqrt{n}, \ \|x\|_\infty = 1.$$
2. Let $x = (1, 0, ..., 0) \in \mathbb{R}^n$. Then
$$\|x\|_1 = \|x\|_2 = \|x\|_\infty = 1.$$

These two examples are interesting since they represent two extreme cases—one in which the "mass" of the vector is spread out evenly over all the components of that vector, and the other in which the "mass" of the vector is spread out very unevenly.

The L_1, L_2 and L_∞ norms are related to each other as follows:

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2, \tag{1}$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty, \tag{2}$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty. \tag{3}$$

Hints for the proof: Let $M = \|x\|_\infty$. Then

$$\left(\sum_1^n |x_i|\right)^2 \geq \sum_1^n |x_i|^2 \geq M^2,$$

$$\sum_1^n |x_i| \leq nM.$$

Remark

Whenever you see an inequality such as one of (1), (2) and (3), you should ask whether it is "tight," i.e., whether it can be "saturated." By that, we mean does there exist a vector $x \in \mathbb{R}^n$ such that the inequality holds as an equality. If the answer is yes, then it is a much more informative inequality than if the answer is no.

Remark

Note that the L_1, L_2 and L_∞ norms are the special cases of the L_p norm defined as follows:

$$\|x\|_p = \left(\sum_1^n |x_i|^p \right)^{1/p},$$

where $1 \leq p \leq \infty$.

Balls

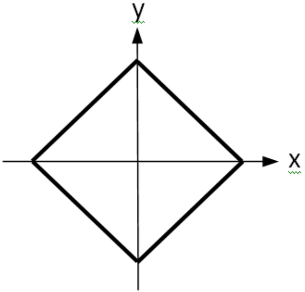
Having defined the size (norm) of a vector, we can now define a *ball* in $\mathbb{R}^n$ which is the generalization of a ball in two or three dimensions.

Definition.

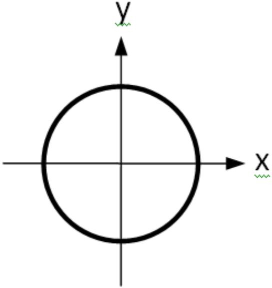
Let $x_0 \in \mathbb{R}^n$ with a norm $\|\cdot\|$. Then

- **Unit sphere:**
$$S(x_0) = \{x \in \mathbb{R}^n : \ \|x - x_0\| = 1\}.$$
- **Open unit ball:**
$$B(x_0) = \{x \in \mathbb{R}^n : \ \|x - x_0\| < 1\}.$$
- **Closed unit ball:**
$$\overline{B}(x_0) = \{x \in \mathbb{R}^n : \ \|x - x_0\| \leq 1\}.$$

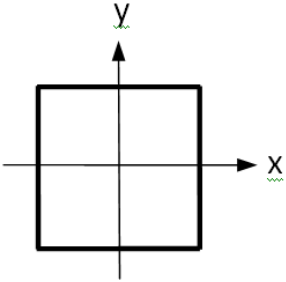
Compare L_1, L_2 and L_∞ unit balls in the plane and space.

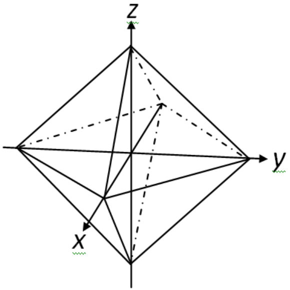


L_1 ball in the plane

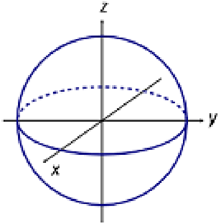


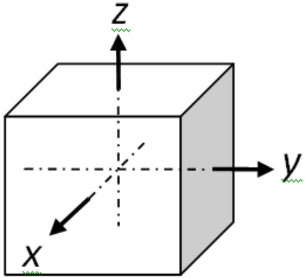
L_2 ball in the plane





L_∞ ball in the plane L_1 ball in the space





L_2 ball in the space L_∞ ball in the space

	L_1 ball	L_2 ball	L_∞ ball
$n=1$	$2r$	$2r$	$2r$
$n=2$	$\frac{(2r)^2}{2!}$	πr^2	$(2r)^2$
$n=3$	$\frac{(2r)^3}{3!}$	$\frac{4}{3}\pi r^3$	$(2r)^3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table: Volumes of balls of radius r for different norms in different dimensions.

High-dimensional vectors, i.e., vectors in $\mathbb{R}^n$, when n is 10^2 or 10^6 or even when it is only 10, have very different properties than vectors in $\mathbb{R}^n$, when $n = 1$ or 2 or 3.

How can we begin to understand this?

To get an understanding of the geometry of $\mathbb{R}^n$, let's ask:

What is the difference between an L_2 ball in $\mathbb{R}$ versus $\mathbb{R}^2$ versus $\mathbb{R}^3$?

More specifically,

What about the ratio of the L_2 ball and L_∞ ball ?

- $\mathbb{R}:$ $$\frac{L_2\text{-ball}}{L_\infty\text{-ball}} = \frac{2r}{2r} = 1$$
- $\mathbb{R}^2:$ $$\frac{L_2\text{-ball}}{L_\infty\text{-ball}} = \frac{\pi r^2}{(2r)^2} = \frac{\pi}{4} = 0.78$$
- $\mathbb{R}^3:$ $$\frac{L_2\text{-ball}}{L_\infty\text{-ball}} = \frac{4\pi r^3/3}{(2r)^3} = \frac{\pi}{6} = 0.55.$$

We raise the same question for the ratio of the L_1 ball and L_∞ ball.

- $\mathbb{R}$: $$\frac{L_1\text{-ball}}{L_\infty\text{-ball}} = \frac{2r}{2r} = 1$$
- $\mathbb{R}^2$: $$\frac{L_1\text{-ball}}{L_\infty\text{-ball}} = \frac{(2\pi)r^2/2!}{(2r)^2} = \frac{1}{2}$$
- $\mathbb{R}^3$: $$\frac{L_1\text{-ball}}{L_\infty\text{-ball}} = \frac{(2\pi)r^3/3!}{(2r)^3} = \frac{1}{6}.$$

Note that the ratio decreases as the dimension increases. We will see that this is true more generally, and it is this phenomenon that is responsible for many of the important and counterintuitive properties of $\mathbb{R}^n$, for $n > 1$.