Mathematical Methods
for Information Engineering

Paola Loreti
Sapienza Università di Roma
Dipartimento di Scienze di Base
e Applicate per l'Ingegneria
First lectures

Send e-mail to paola.loreti@uniroma1.it to fix a meeting on line.

Table of contents

Vectorial Spaces

A vectorial space over a field K is a set V with two applications, sum and product with a scalar number λ, characterized by the following properties

- the sum of two vectors u, v gives a new vector denoted by $u + v$,

$$(u, v) \to u + v$$

- the product of the vector u with a scalar number $\lambda \in K$ gives a new vector denoted by λu

$$(u, \lambda) \to \lambda u$$

The following properties are requested

- $(V, +)$ is an abelian group:
- $\lambda(u + v) = \lambda u + \lambda v \ \ \forall \lambda \in K \ \forall u, v \in V$
- $(\lambda + \lambda_1)v = \lambda v + \lambda_1 v \ \ \forall \lambda, \lambda_1 \in K \ \forall v \in V$
- $(\lambda \lambda_1)v = \lambda(\lambda_1 v) \ \ \forall \lambda, \lambda_1 \in K \ \forall v \in V$
- $1v = v \ \forall v \in V$

$V = \mathbb{R}^m \ K = \mathbb{R}$.

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_m + y_m)$$

$$\lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_m)$$

Let V a vectorial space, a subset W of V is a vectorial subspace if is a vectorial space with respect to the same applications:

$$\forall \lambda, \lambda_1 \in K, \ \forall u, v \in W \implies \lambda u + \lambda_1 v \in W$$

Notation $V(K)$, V over K

Normed Spaces

A vectorial space $X(\mathbb{R})$ endowed with norm is a vectorial normed space

$\forall x, y, z \in X$ e $\lambda \in \mathbb{R}$, the properties hold

- $\|x\| \geq 0$,
- $\|x\| = 0 \iff x = 0$,
- $\|\lambda x\| = |\lambda| \cdot \|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$.

Let us start our discussion recalling the properties of the modulus.
$\forall\, x, y \in \mathbb{R}$ the following properties hold true

- $|x| \geq 0$
- $x \neq 0$ if and only if $|x| > 0$
- $|x| = |-x|$
- $|xy| = |x||y|$
- $|x + y| \leq |x| + |y|$
- $||x| - |y|| \leq |x - y|$

Norms $\mathbb{R}^m$ and $p \geq 1$. The formula

$$\|x\|_p = (|x_1|^p + \cdots + |x_m|^p)^{1/p}.$$

defines a norm in $\mathbb{R}^m$.

We need to show the following properties $\forall x, y, z \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}$:

- $\|x\|_p \geq 0$,
- $\|x\|_p = 0 \iff x = 0$,
- $\|\lambda x\|_p = |\lambda| \cdot \|x\|_p$,
- $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

The inequality
$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$
will be shown later, thanks to Minkowski inequality. Particular cases
$$p = 1$$
$$p = 2$$
$$p = +\infty$$

The triangular inequality.
A particular case $p = 1$.

Example

▶ The formula

$$\|x\|_1 = |x_1| + \cdots + |x_m|, \quad x = (x_1, \ldots, x_m) \in \mathbb{R}^m$$

defines a norm on $\mathbb{R}^m$.
Indeed

$$\|x + y\|_1 = |x_1 + y_1| + \cdots + |x_m + y_m| \leq |x_1| + |y_1| + \cdots + |x_m| + |y_m|$$

$$= |x_1| + \cdots + |x_m| + \cdots + |y_1| + \cdots + |y_m| = \|x\|_1 + \|y\|_1$$

A particular case $p = \infty$.

Example

► The formula

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_m|\}$$

defines a norm on $\mathbb{R}^m$.

Scalar Product The scalar product in $\mathbb{R}^m$ ìs real number given by

$$x \cdot y = x_1 y_1 + \cdots + x_m y_m \qquad \text{for all } x, y \in \mathbb{R}^m$$

We need to verify that the following properties hold
for all $x, y, z \in \mathbb{R}^m$ $\lambda \in \mathbb{R}$

- $x \cdot y = y \cdot x$,
- $(x + y) \cdot z = x \cdot z + y \cdot z$,
- $\lambda(x \cdot y) = \lambda x \cdot y$.

We have

$$(x, x) = \|x\|^2$$

Ricordiamo $x^T x$: l'espressione (dove x è un vettore colonna)
rappresenta il prodotto scalare di un vettore per se stesso,
equivalente al quadrato della sua norma euclidea.

Young inequality Given $p > 1$, $p \in \mathbb{R}$ we define the conjugate of p the real number q such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Theorem
Young inequality: given two real positive numbers a e b, and given two numbers real and conjugate p, q, we have
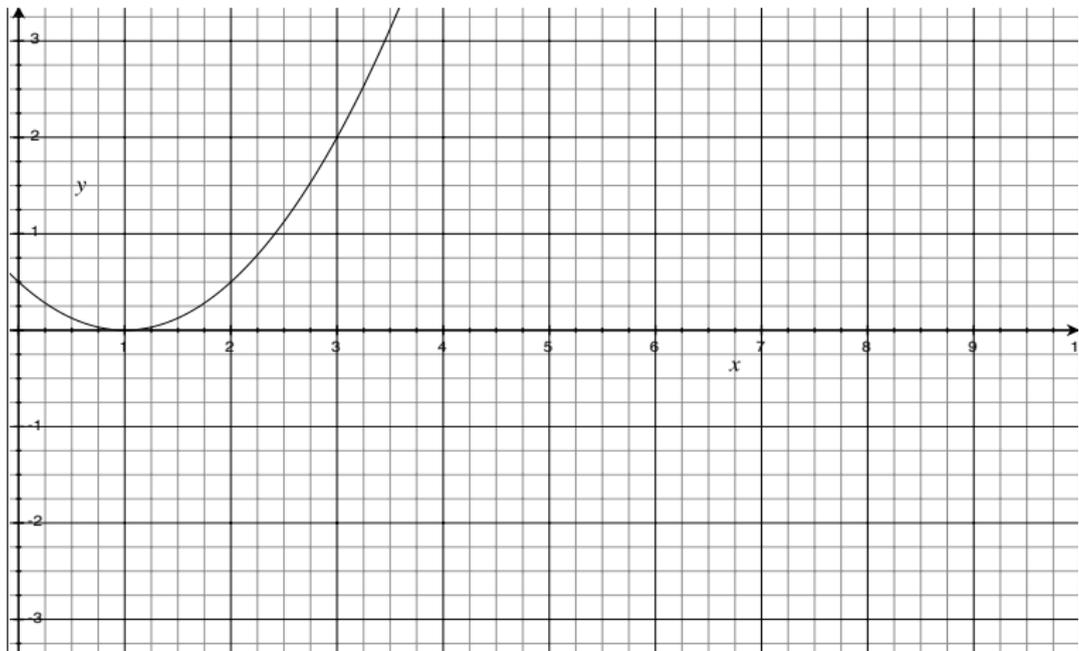
$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Equality

$$a^p = b^q$$

Indeed if $a^p = b^q$

$$ab = a(b^q)^{\frac{1}{q}} = aa^{\frac{p}{q}} = a^{\frac{p}{q}+1} = a^p = \frac{1}{p}a^p + (1 - \frac{1}{p})a^p = \frac{1}{p}a^p + \frac{1}{q}b^q$$

Let $b > 0$ and fixed and we define

$$f : [0, +\infty) \to \mathbb{R} \qquad f(t) = \frac{t^p}{p} + \frac{b^q}{q} - tb$$

Since

$$\lim_{t \to +\infty} \frac{t^p}{p} + \frac{b^q}{q} - tb = +\infty \qquad f(0) = \frac{b^q}{q} > 0$$

if we are to show that there exists a unique point $\hat{t} > 0$ such that $f'(\hat{t}) = 0$ and $f(\hat{t}) = 0$ then $\hat{t}$ will be the absolute minimum point

$$f'(t) = t^{p-1} - b$$

$$t^{p-1} = b \iff \hat{t} = b^{\frac{1}{p-1}} \qquad f''(b^{\frac{1}{p-1}}) > 0$$

$$f(b^{\frac{1}{p-1}}) = \frac{b^{\frac{p}{p-1}}}{p} + \frac{b^q}{q} - b^{\frac{1}{p-1}}b = \left(\frac{1}{p} + \frac{1}{q} - 1\right)b^q = 0$$

Then for any $a \geq 0$

$$f(a) \geq 0,$$

this means

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Proof using convexity. For $a > 0, p > 1$ the function $f(a) = \frac{a^p}{p}$ is a convex function and $-f(a)$ is a concave function. The Legendre transform is

$$f^*(b) = \max_a \left(ba - \frac{a^p}{p} \right)$$

Consider

$$F(a, b) = ab - \frac{a^p}{p}$$

Making the derivative with respect to a

$$F'(a, b) = 0 \iff a = b^{\frac{1}{p-1}}$$

$$f^*(b) = b^{\frac{p}{p-1}} \left(1 - \frac{1}{p} \right) = \frac{b^q}{q}$$

Then $ab - \frac{a^p}{p} \leq \max_a \left(ba - \frac{a^p}{p} \right) = \frac{b^q}{q}$

Theorem (Hölder Inequality)

Let p, q such that $p, q \in [1, +\infty)$ and conjugate, then $\forall x, y \in \mathbb{R}^m$ we have

$$|x \cdot y| \leq \|x\|_p \|y\|_q.$$

$$a_i = \frac{|x_i|}{\|x\|_p}, \qquad b_i = \frac{|y_i|}{\|y\|_q}$$

Follow, by Young inequality

$$a_i b_i \leq \frac{1}{p} \frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|y_i|^q}{\|y\|_q^q}$$

Taking the sum over the index i

$$\sum_{i=1}^m a_i b_i \leq \frac{1}{p} \frac{\sum_{i=1}^m |x_i|^p}{\|x\|_p^p} + \frac{1}{q} \frac{\sum_{i=1}^m |y_i|^q}{\|y\|_q^q} = 1$$

Then we get
$$\sum_{i=1}^{m} a_i b_i = \sum_{i=1}^{m} \frac{|x_i|}{\|x\|_p} \frac{|y_i|}{\|y\|_q} \leq 1$$
and Hölder inequality follows
$$|x \cdot y| \leq \|x\|_p \|y\|_q.$$

Theorem (Minkowski inequality)

Let $p \in [1, +\infty)$ and $\forall \ x, y \in \mathbb{R}^m$ then

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p. \tag{1}$$

We have

$$|x_i + y_i|^p = |x_i + y_i|^{p-1}|x_i + y_i| \leq$$
$$|x_i + y_i|^{p-1}(|x_i| + |y_i|)$$

Taking the sum

$$\sum_{i=1}^{m} |x_i + y_i|^p \leq \sum_{i=1}^{m} |x_i + y_i|^{p-1}|x_i| + \sum_{i=1}^{m} |x_i + y_i|^{p-1}|y_i|$$

we obtain

$$\sum_{i=1}^{m} |x_i + y_i|^{p-1}|x_i| \leq \|x\|_p \left(\sum_{i=1}^{m} |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}}$$

$$\sum_{i=1}^{m} |x_i + y_i|^{p-1}|y_i| \leq \|y\|_p \left(\sum_{i=1}^{m} |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}}$$

Then since $(p-1)q = p$

$$\|x + y\|_p^p \leq \|x + y\|_p^{p-1}(\|x\|_p + \|y\|_p)$$

then making the quotient with $\|x + y\|_p^{p-1}$ (that we assume not 0) we obtain the Minkowski inequality

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Example

$\mathbb{R}^m(\mathbb{R})$ with the euclidean norm. Given $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ then

$$\|x\|_2 = \left(x_1^2 + \cdots + x_m^2\right)^{1/2}.$$

Properties. It is possible to show

$$\lim_{p \to +\infty} \|x\|_p = \|x\|_\infty$$

Proof.
Indeed by the comparison with norms for any $p \geq 1$

$$\|x\|_\infty \leq \|x\|_p \leq m^{\frac{1}{p}} \|x\|_\infty,$$

and the result follows passing to the limit $p \to +\infty$. $\quad\quad\quad\square$
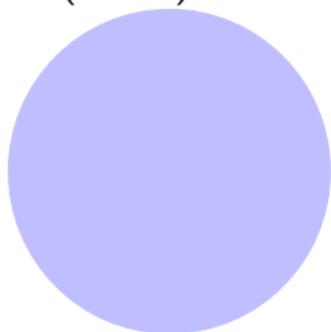
Recall

$$\|x\|_\infty = |x_{i_0}|,$$

for some i_0.

$$\|x\|_\infty^p = |x_{i_0}|^p \leq \sum_{i=1}^m |x_i|^p \leq m|x_{i_0}|^p = m \|x\|_\infty^p$$

Exercises. Consider

$$\|x\|_2 \leq 1.$$

This is the ball with respect to the euclidean norm: we draw the ball in the plane ($n = 2$).
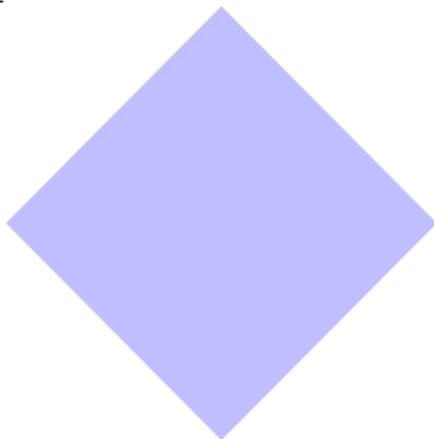
$\|x\|_2 \leq 1$

Now we consider the *the ball* with respect to $\|x\|_\infty$: in the plane this is the square.
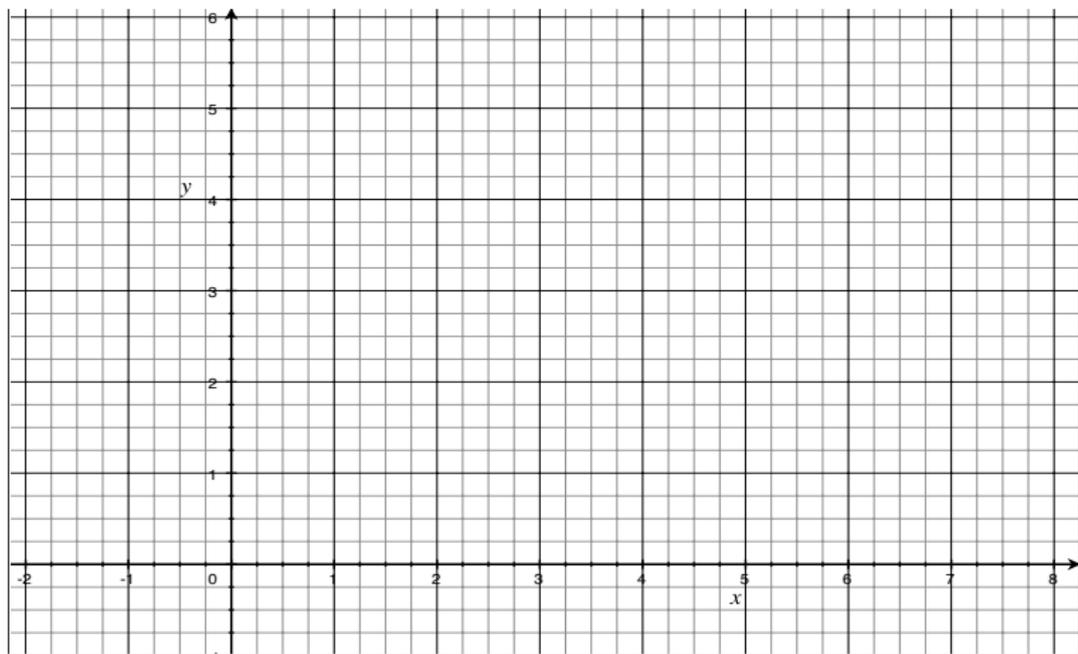
$$\|x\|_\infty \leq 1$$

Now we consider the *the ball* with respect to $\|x\|_1$: we draw in the plane $\|x\|_1 \leq 1$.

$\|x\|_1 \leq 1$

$$d_1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|,$$

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \text{ and } \mathbf{y} = (y_1, y_2, \ldots, y_n)$$

$\|x\|_1$:this is the taxicab norm or Manhattan norm. The name relates to the distance a taxi has to drive in a rectangular street grid to get from the origin to the point x. The distance derived from this norm is called the Manhattan distance.

$$d_1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|,$$

A taxicab (Manhattan distance or Manhattan length) geometry is a form of geometry in which the metric of Euclidean geometry is replaced by a new metric in which the distance between two points is the sum of the absolute differences of their Cartesian coordinates. The name alludes to the grid layout of most streets on the island of Manhattan, which causes the shortest path a car could take.

Metric Spaces.

Consider at first $\mathbb{R}^m$: this is a normed space with the $\|x\|_2$.

Definition

We define the *distance* between two points of $\mathbb{R}^m$ as

$$d(x, y) := \|x - y\|$$

$$d(x, y) := \|x - y\| = \sqrt{\sum_{i=1}^{m}(x_i - y_i)^2}$$

- $d(x, y) \geq 0$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(z, y)$

The canonical base in $\mathbb{R}^m$ is given by the vectors
$e^1 = (1, 0, \ldots, 0)$, $e^2 = (0, 1, \ldots, 0)$, $e^m = (0, 0, \ldots, 1)$.

$$e^j = (0, \ldots 1, 0 \ldots 0)$$
$$e^k = (0, \ldots 0, 1 \ldots 0).$$

We may compute the distance

$$d(e^j, e^k) = \sqrt{2} \qquad j \neq k$$

$\mathbb{R}^m$ with $\|x\|_2$ may be endowed of a metric, then $(\mathbb{R}^m, d)$ is a metric space.

$$(X, d)$$

Generally, X is a set and d the metric

- $d(x, y) \geq 0$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(z, y)$

Two norms $\|x\|_a$ $\|x\|_b$ are equivalent if there exist two constant m and M such that

$$m \|x\|_b \leq \|x\|_a \leq M \|x\|_b .$$

The norms p for $p \geq 1$ are equivalent

Equivalence of norms in $\mathbb{R}^m$.

- ▶ Norm equivalence is a transitive property (by changing the constants). We take $\|x\|_1$.
- ▶ We can divide by it (taking it different from 0). Then we have to show

$$m \le \|u\|_a \le M.$$

Observe that $u = x/\|x\|_1$ has norm $\|u\|_1 = 1$.

▶ Continuity. Given any norm $\|x\|_a$ on a vector space of dimension m, for any basis $(e_1, \ldots, e_m)$ of E, observe that for any vector $x = x_1 e_1 + \ldots x_m e_m$, we have

$$\|x\|_a \leq C \|x\|_1$$

with $C = \max_i \|e_i\|_a$. The above implies that

$$\left| \|u\|_a - \|v\|_a \right| \leq \|u - v\|_a \leq C \|u - v\|_1,$$

which means that the map $\rightarrow \|u\|_a$ is continuous with respect to the norm $\|u\|_1$.

- $\|u\|_a$ in the compact set $\|u\|_1 = 1$ achieves the maximum and the minimum value on the set.