CS724: Topics in Algorithms
Norms and Inner Products - I
Slide Set 4

Prof. Dan A. Simovici

Lemma

Let $p, q \in \mathbb{R} - \{0, 1\}$ *such that* $\frac{1}{p} + \frac{1}{q} = 1$. *Then we have* $p > 1$ *if and only if* $q > 1$. *Furthermore, one of the numbers* p, q *belongs to the interval* $(0, 1)$ *if and only if the other number is negative.*

Lemma

Let $p, q \in \mathbb{R} - \{0, 1\}$ *be two numbers such that* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $p > 1$. *Then, for every* $a, b \in \mathbb{R}_{\geqslant 0}$, *we have*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q},$$

where the equality holds if and only if $a = b^{-\frac{1}{1-p}}$.

Proof

We have $q > 1$. Consider the function $f(x) = \frac{x^p}{p} + \frac{1}{q} - x$ for $x \geqslant 0$. We have $f'(x) = x^{p-1} - 1$, so the minimum is achieved when $x = 1$ and $f(1) = 0$. Thus,

$$f\left(ab^{-\frac{1}{p-1}}\right) \geqslant f(1) = 0,$$

which amounts to

$$\frac{a^p b^{-\frac{p}{p-1}}}{p} + \frac{1}{q} - ab^{-\frac{1}{p-1}} \geqslant 0.$$

By multiplying both sides of this inequality by $b^{\frac{p}{p-1}}$, we obtain the desired inequality.

Observe that if $\frac{1}{p} + \frac{1}{q} = 1$ and $p < 1$, then $q < 0$. In this case, we have the reverse inequality

$$ab \geqslant \frac{a^p}{p} + \frac{b^q}{q}. \tag{1}$$

which can be shown by observing that the function f has a maximum in $x = 1$. The same inequality holds when $q < 1$ and therefore $p < 0$.

Theorem

(The Hölder Inequality) *Let* $a_1, \ldots, a_n$ *and* $b_1, \ldots, b_n$ *be* $2n$ *nonnegative numbers, and let* p *and* q *be two numbers such that* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $p > 1$. *We have*

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p \right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} b_i^q \right)^{\frac{1}{q}} .$$

Proof

If $a_1 = \cdots = a_n = 0$ or if $b_1 = \cdots = b_n = 0$, then the inequality is clearly satisfied. Therefore, we may assume that at least one of $a_1, \ldots, a_n$ and at least one of $b_1, \ldots, b_n$ is non-zero. Define the numbers

$$x_i = \frac{a_i}{\left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}}} \text{ and } y_i = \frac{b_i}{\left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}}}$$

for $1 \leqslant i \leqslant n$. Lemma on Slide 3 applied to x_i, y_i yields

$$\frac{a_i b_i}{\left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}}} \leqslant \frac{1}{p} \frac{a_i^p}{\sum_{i=1}^{n} a_i^p} + \frac{1}{q} \frac{b_i^p}{\sum_{i=1}^{n} b_i^p}.$$

Adding these inequalities, we obtain

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}}$$

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

The nonnegativity of the numbers $a_1, \ldots, a_n, b_1, \ldots, b_n$ can be relaxed by using absolute values.

Theorem

Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be $2n$ numbers and let p and q be two numbers such that $\frac{1}{p} + \frac{1}{q} = 1$ and $p > 1$. We have

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leqslant \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} |b_i|^q\right)^{\frac{1}{q}}.$$

UMASS BOSTON

Proof

By a previous theorem, we have:

$$\sum_{i=1}^{n} |a_i||b_i| \leqslant \left(\sum_{i=1}^{n} |a_i|^p \right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} |b_i|^q \right)^{\frac{1}{q}}.$$

The needed equality follows from the fact that

$$\left| \sum_{i=1}^{n} a_i b_i \right| \leqslant \sum_{i=1}^{n} |a_i||b_i|.$$

Corollary

(The Cauchy-Schwarz Inequality for $\mathbb{R}^n$) Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be $2n$ real numbers. We have

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leqslant \sqrt{\sum_{i=1}^{n} a_i^2} \cdot \sqrt{\sum_{i=1}^{n} b_i^2}.$$

Proof.

The inequality follows immediately by taking $p = q = 2$. □

UMASS BOSTON

Theorem

(Minkowski's Inequality) Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be $2n$ nonnegative real numbers. If $p \geqslant 1$, we have

$$\left(\sum_{i=1}^{n}(a_i + b_i)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}.$$

If $p < 1$, the inequality sign is reversed.

Proof

For $p = 1$, the inequality is immediate. Therefore, we can assume that $p > 1$. Note that

$$\sum_{i=1}^{n} (a_i + b_i)^p = \sum_{i=1}^{n} a_i (a_i + b_i)^{p-1} + \sum_{i=1}^{n} b_i (a_i + b_i)^{p-1}.$$

By Hölder's inequality for p, q such that $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\begin{aligned}
\sum_{i=1}^{n} a_i (a_i + b_i)^{p-1} &\leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (a_i + b_i)^{(p-1)q}\right)^{\frac{1}{q}} \\
&= \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (a_i + b_i)^{p}\right)^{\frac{1}{q}}.
\end{aligned}$$

Proof cont'd

Similarly, we can write

$$\sum_{i=1}^{n} b_i(a_i + b_i)^{p-1} \leqslant \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n}(a_i + b_i)^p\right)^{\frac{1}{q}}.$$

Adding the last two inequalities yields

$$\sum_{i=1}^{n}(a_i + b_i)^p \leqslant \left(\left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}\right) \left(\sum_{i=1}^{n}(a_i + b_i)^p\right)^{\frac{1}{q}},$$

which is equivalent to inequality

$$\left(\sum_{i=1}^{n}(a_i + b_i)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} b_i^p\right)^{\frac{1}{p}}.$$

Definition

A function $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ is a *metric* if it has the following properties:

- $d(x, y) = 0$ if and only if $x = y$ for $x, y \in S$;
- $d(x, y) = d(y, x)$ for $x, y \in S$;
- $d(x, y) \leqslant d(x, z) + d(z, y)$ for $x, y, z \in S$.

The pair (S, d) will be referred to as a *metric space*.

If property (i) is replaced by the weaker requirement that $d(x, x) = 0$ for $x \in S$, then we refer to d as a *semimetric* on S. Thus, if d is a semimetric $d(x, y) = 0$ does not necessarily imply $x = y$ and we can have for two distinct elements x, y of S, $d(x, y) = 0$. If d is a semimetric, then we refer to the pair (S, d) as a *semimetric space*.

Example

Let S be a nonempty set. Define the mapping $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ by

$$d(u,v) = \begin{cases} 1 & \text{if } u \neq v, \\ 0 & \text{otherwise,} \end{cases}$$

for $x, y \in S$. It is easy to see that d satisfies the definiteness property. To prove that d satisfies the triangular inequality, we need to show that

$$d(x,y) \leqslant d(x,z) + d(z,y)$$

for all $x, y, z \in S$. This is clearly the case if $x = y$. Suppose that $x \neq y$, so $d(x,y) = 1$. Then, for every $z \in S$, we have at least one of the inequalities $x \neq z$ or $z \neq y$, so at least one of the numbers $d(x,z)$ or $d(z,y)$ equals 1. Thus d satisfies the triangular inequality. The metric d introduced here is the *discrete metric* on S.

Example

Consider the mapping $d : (\mathbf{Seq}_n(S))^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ defined by

$$d(\boldsymbol{p}, \boldsymbol{q}) = |\{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \boldsymbol{p}(i) \neq \boldsymbol{q}(i)\}|$$

for all sequences $\boldsymbol{p}, \boldsymbol{q}$ of length n on the set S.
It is easy to see that d is a metric. We justify here only the triangular inequality. Let $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$ be three sequences of length n on the set S. If $\boldsymbol{p}(i) \neq \boldsymbol{q}(i)$, then $\boldsymbol{r}(i)$ must be distinct from at least one of $\boldsymbol{p}(i)$ and $\boldsymbol{q}(i)$. Therefore,

$$\begin{aligned} &\{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \boldsymbol{p}(i) \neq \boldsymbol{q}(i)\} \\ &\quad \subseteq \quad \{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \boldsymbol{p}(i) \neq \boldsymbol{r}(i)\} \cup \{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \boldsymbol{r}(i) \neq \boldsymbol{q}(\end{aligned}$$

which implies the triangular inequality.

Example

For $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^n$ the *Euclidean metric* is the mapping

$$d_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

The first two conditions of Definition 7 are obviously satisfied.
To prove the third inequality, let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$. Choosing $a_i = x_i - y_i$ and $b_i = y_i - z_i$ for $1 \leqslant i \leqslant n$ in Minkowski's inequality implies

$$\sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} \leqslant \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} + \sqrt{\sum_{i=1}^{n} (y_i - z_i)^2},$$

which amounts to $d(\boldsymbol{x}, \boldsymbol{z}) \leqslant d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$. Thus, we conclude that d is indeed a metric on $\mathbb{R}^n$.

We use frequently use the notions of closed sphere and open sphere.

Definition

Let (S, d) be a metric space. The *closed sphere* centered in $x \in S$ of radius r is the set

$$B_d[x, r] = \{y \in S | d(x, y) \leqslant r\}.$$

The *open sphere* centered in $x \in S$ of radius r is the set

$$B_d(x, r) = \{y \in S | d(x, y) < r\}.$$

Definition

Let (S, d) be a metric space. The *diameter* of a subset U of S is the number $diam_{S,d}(U) = \sup\{d(x, y) \mid x, y \in U\}$. The set U is *bounded* if $diam_{S,d}(U)$ is finite.
The *diameter* of the metric space (S, d) is the number

$$diam_{S,d} = \sup\{d(x, y) \mid x, y \in S\}.$$

If the metric space is clear from the context, then we denote the diameter of a subset U just by $diam(U)$.
If (S, d) is a finite metric space, then $diam_{S,d} = \max\{d(x, y) \mid x, y \in S\}$.

A mapping $d : S \times S \longrightarrow \hat{\mathbb{R}}_{\geqslant 0}$ can be extended to the set of subsets of S by defining $d(U, V)$ as

$$d(U, V) = \inf\{d(u, v) \mid u \in U \text{ and } v \in V\}$$

for $U, V \in \mathcal{P}(S)$.

Observe that, even if d is a metric, then its extension is not, in general, a metric on $\mathcal{P}(S)$ because it does not satisfy the triangular inequality. Instead, we can show that for every U, V, W we have

$$d(U, W) \leqslant d(U, V) + diam(V) + d(V, W).$$

Indeed, by the definition of $d(U,V)$ and $d(V,W)$, for every $\epsilon > 0$, there exist $u \in U$, $v, v' \in V$, and $w \in W$ such that

$$d(U,V) \leqslant d(u,v) \leqslant d(U,V) + \frac{\epsilon}{2},$$
$$d(V,W) \leqslant d(v',w) \leqslant d(V,W) + \frac{\epsilon}{2}.$$

By the triangular axiom, we have

$$d(u,w) \leqslant d(u,v) + d(v,v') + d(v',w).$$

Hence,

$$d(u,w) \leqslant d(U,V) + diam(V) + d(V,W) + \epsilon,$$

which implies

$$d(U,W) \leqslant d(U,V) + diam(V) + d(V,W) + \epsilon$$

for every $\epsilon > 0$. This yields the needed inequality.

Definition

Let (S, d) be a metric space. The sets $U, V \in \mathcal{P}(S)$ are *separate* if $d(U, V) > 0$.

We denote the number $d(\{u\}, V) = \inf\{d(u, v) \mid v \in V\}$ by $d(u, V)$. It is clear that $u \in V$ implies $d(u, V) = 0$.

The notion of dissimilarity is a generalization of the notion of metric.

Definition

A *dissimilarity on a set S* is a function $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ satisfying the following conditions:

- $d(x, x) = 0$ for all $x \in S$;
- $d(x, y) = d(y, x)$ for all $x, y \in S$.

The pair (S, d) is a *dissimilarity space*.

A related concept is the notion of similarity.

Definition

A *similarity on a set* S is a function $s : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ satisfying the following conditions:

- $s(x, y) \leqslant s(x, x) = 1$ for all $x, y \in S$;
- $s(x, y) = s(y, x)$ for all $x, y \in S$.

The pair (S, s) is a *similarity space*.

Example

Let $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ be a metric on the set S. Then $s : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ defined by $s(x, y) = 2^{-d(x,y)}$ for $x, y \in S$ is a dissimilarity, such that $s(x, x) = 1$ for every $x, y \in S$.

Definition

A *seminorm* on an F-linear space V is a mapping $\nu : V \longrightarrow \mathbb{R}$ that satisfies the following conditions:

- $\nu(\boldsymbol{x}+\boldsymbol{y}) \leqslant \nu(\boldsymbol{x}) + \nu(\boldsymbol{y})$ (subadditivity), and
- $\nu(a\boldsymbol{x}) = |a|\nu(\boldsymbol{x})$ (positive homogeneity),

for $\boldsymbol{x}, \boldsymbol{y} \in V$ and $a \in F$.

By taking $a = 0$ in the second condition of the definition we have $\nu(\mathbf{0}) = 0$ for every seminorm on a real or complex space.

Theorem

If V is a real or complex linear space and $\nu : V \longrightarrow \mathbb{R}$ is a seminorm on V, then

$$\nu(\boldsymbol{x} - \boldsymbol{y}) \geqslant |\nu(\boldsymbol{x}) - \nu(\boldsymbol{y})|,$$

for $\boldsymbol{x}, \boldsymbol{y} \in V$.

Proof.

We have $\nu(\boldsymbol{x}) \leqslant \nu(\boldsymbol{x} - \boldsymbol{y}) + \nu(\boldsymbol{y})$, so

$$\nu(\boldsymbol{x}) - \nu(\boldsymbol{y}) \leqslant \nu(\boldsymbol{x} - \boldsymbol{y}). \tag{2}$$

Since $\nu(\boldsymbol{x} - \boldsymbol{y}) = |-1|\nu(\boldsymbol{y} - \boldsymbol{x}) \geqslant \nu(\boldsymbol{y}) - \nu(\boldsymbol{x})$ we have

$$-(\nu(\boldsymbol{x}) - \nu(\boldsymbol{y})) \leqslant \nu(\boldsymbol{x}) - \nu(\boldsymbol{y}). \tag{3}$$

The Inequalities (2) and (3) give the desired inequality. □

UMASS BOSTON

Corollary

If $p : V \longrightarrow \mathbb{R}$ is a seminorm on V, then $p(\mathbf{x}) \geqslant 0$ for $\mathbf{x} \in V$.

Proof.

By choosing $\mathbf{y} = \mathbf{0}$ we have $\nu(\mathbf{x}) \geqslant |\nu(\mathbf{x})| \geqslant 0$. $\square$

Definition

Let $\mathcal{F} = (F, \{0, 1, +, -, \cdot\})$ be the real or the complex field. A *norm* on an F-linear space V is a seminorm $\nu : V \longrightarrow \mathbb{R}$ such that $\nu(\boldsymbol{x}) = 0$ implies $\boldsymbol{x} = \mathbf{0}$ for $\boldsymbol{x} \in V$.

The pair (V, ν) is referred to as a *normed linear space*.

Example

The set of real-valued continuous functions defined on the interval $[-1, 1]$ is a real linear space. The addition of two such functions f, g, is defined by $(f + g)(x) = f(x) + g(x)$ for $x \in [-1, 1]$; the multiplication of f by a scalar $a \in \mathbb{R}$ is $(af)(x) = af(x)$ for $x \in [-1, 1]$.
Define $\nu(f) = \sup\{|f(x)| \mid x \in [-1, 1]\}$. Since $|f(x)| \leqslant \nu(f)$ and $|g(x)| \leqslant \nu(g)$ for $x \in [-1, 1]\}$, it follows that
$|(f + g)(x)| \leqslant |f(x)| + |g(x)| \leqslant \nu(f) + \nu(g)$. Thus,
$\nu(f + g) \leqslant \nu(f) + \nu(g)$.
We denote $\nu(f)$ by $\| f \|$.

Theorem

For $p \geqslant 1$, the function $\nu_p : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geqslant 0}$ defined by

$$\nu_p(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}},$$

where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, is a norm on $\mathbb{R}^n$.

Proof

We must prove that ν_p satisfies the conditions of the definition of norms and that $\nu_p(\boldsymbol{x}) = 0$ implies $\boldsymbol{x} = \boldsymbol{0}$.
Let $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Minkowski's inequality applied to the nonnegative numbers $a_i = |x_i|$ and $b_i = |y_i|$ amounts to

$$\left(\sum_{i=1}^{n}(|x_i| + |y_i|)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}}.$$

Since $|x_i + y_i| \leqslant |x_i| + |y_i|$ for every i, we have

$$\left(\sum_{i=1}^{n}(|x_i + y_i|)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}},$$

that is, $\nu_p(\boldsymbol{x} + \boldsymbol{y}) \leqslant \nu_p(\boldsymbol{x}) + \nu_p(\boldsymbol{y})$. Thus, ν_p is a norm on $\mathbb{R}^n$.

Example

The mapping $\nu_1 : \mathbb{R}^n \longrightarrow \mathbb{R}$ given by

$$\nu_1(\boldsymbol{x}) = |x_1| + |x_2| + \cdots + |x_n|,$$

for $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. ν_1 is a norm on $\mathbb{R}^n$.

Example

A special norm on $\mathbb{R}^n$ is the function $\nu_\infty : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geqslant 0}$ given by

$$\nu_\infty(\boldsymbol{x}) = \max\{|x_i| \mid 1 \leqslant i \leqslant n\}$$

for $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.
We start from the inequality

$$|x_i + y_i| \leqslant |x_i| + |y_i| \leqslant \nu_\infty(\boldsymbol{x}) + \nu_\infty(\boldsymbol{y})$$

for every i, $1 \leqslant i \leqslant n$. This implies

$$\nu_\infty(\boldsymbol{x} + \boldsymbol{y}) = \max\{|x_i + y_i| \mid 1 \leqslant i \leqslant n\} \leqslant \nu_\infty(\boldsymbol{x}) + \nu_\infty(\boldsymbol{y}),$$

which gives the desired inequality.

Example

This norm can be regarded as a limit case of the norms ν_p. Indeed, let $\boldsymbol{x} \in \mathbb{R}^n$ and let $M = \max\{|x_i| \mid 1 \leqslant i \leqslant n\} = |x_{\ell_1}| = \cdots = |x_{\ell_k}|$ for some $\ell_1, \ldots, \ell_k$, where $1 \leqslant \ell_1, \ldots, \ell_k \leqslant n$. Here $x_{\ell_1}, \ldots, x_{\ell_k}$ are the components of $\boldsymbol{x}$ that have the maximal absolute value and $k \geqslant 1$. We can write

$$\lim_{p\to\infty} \nu_p(\boldsymbol{x}) = \lim_{p\to\infty} M \left(\sum_{i=1}^{n} \left(\frac{|x_i|}{M} \right)^p \right)^{\frac{1}{p}} = \lim_{p\to\infty} M(k)^{\frac{1}{p}} = M,$$

which justifies the notation ν_∞.

We use the alternative notation $\| \boldsymbol{x} \|_p$ for $\nu_p(\boldsymbol{x})$. We refer $\| \boldsymbol{x} \|_2$ as the *Euclidean norm* of $\boldsymbol{x}$ and we denote this norm simply by $\| \boldsymbol{x} \|$ when there is no risk of confusion.

Example

For $p \geqslant 1$, let ℓ_p be the set that consists of sequences of real numbers $\boldsymbol{x} = (x_0, x_1, \ldots)$ such that the series $\sum_{i=0}^{\infty} |x_i|^p$ is convergent. We can show that ℓ_p is a linear space.

Let $\boldsymbol{x}, \boldsymbol{y} \in \ell_p$ be two sequences in ℓ_p. Using Minkowski's inequality we have

$$\sum_{i=0}^{n} |x_i + y_i|^p \leqslant \sum_{i=0}^{n} (|x_i| + |y_i|)^p \leqslant \sum_{i=0}^{n} |x_i|^p + \sum_{i=0}^{n} |y_i|^p,$$

which shows that $\boldsymbol{x} + \boldsymbol{y} \in \ell_p$. It is immediate that $\boldsymbol{x} \in \ell_p$ implies $a\boldsymbol{x} \in \ell_p$ for every $a \in \mathbb{R}$ and $\boldsymbol{x} \in \ell_p$.

The following statement shows that any norm defined on a linear space generates a metric on the space.

Theorem

Each norm $\nu : V \longrightarrow \mathbb{R}_{\geqslant 0}$ on a real linear space V generates a metric on the set V defined by $d_\nu(\mathbf{x}, \mathbf{y}) = \nu(\mathbf{x} - \mathbf{y})$ for $\mathbf{x}, \mathbf{y} \in V$.

Proof.

Note that if $d_\nu(\mathbf{x}, \mathbf{y}) = \nu(\mathbf{x} - \mathbf{y}) = 0$, it follows that $\mathbf{x} - \mathbf{y} = \mathbf{0}$; that is, $\mathbf{x} = \mathbf{y}$.
The symmetry of d_ν is obvious and so we need to verify only the triangular axiom. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in L$. Applying the subaditivity of norms we have we have

$$\nu(\mathbf{x} - \mathbf{z}) = \nu(\mathbf{x} - \mathbf{y} + \mathbf{y} - \mathbf{z}) \leqslant \nu(\mathbf{x} - \mathbf{y}) + \nu(\mathbf{y} - \mathbf{z})$$

or, equivalently, $d_\nu(\mathbf{x}, \mathbf{z}) \leqslant d_\nu(\mathbf{x}, \mathbf{y}) + d_\nu(\mathbf{y}, \mathbf{z})$, for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in L$, which concludes the argument. □

UMASS BOSTON

Observe that the norm ν can be expressed using d_ν as

$$\nu(\boldsymbol{x}) = d_\nu(\boldsymbol{x}, \boldsymbol{0})$$

for $\boldsymbol{x} \in V$.
For $p \geqslant 1$, then d_p denotes the metric d_{ν_p} induced by the norm ν_p on the linear space $\mathbb{R}^n$ known as the *Minkowski metric*.
If $p = 2$, we have the *Euclidean metric* on $\mathbb{R}^n$ given by

$$d_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

UMASS BOSTON

For $p = 1$, we have

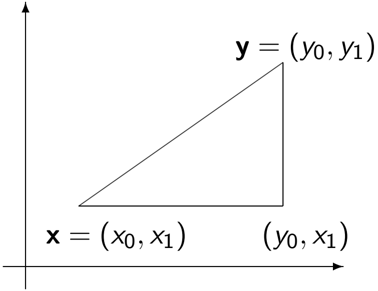
$$d_1(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} |x_i - y_i|.$$

This metric is known also as the *city-block metric*.
The norm ν_∞ generates the metric d_∞ given by

$$d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max\{|x_i - y_i| \mid 1 \leqslant i \leqslant n\},$$

also known as the *Chebyshev metric*.

A representation of these metrics can be seen below for the special case of $\mathbb{R}^2$. If $\boldsymbol{x} = (x_0, x_1)$ and $\boldsymbol{y} = (y_0, y_1)$, then $d_2(\boldsymbol{x}, \boldsymbol{y})$ is the length of the hypotenuse of the right triangle and $d_1(\boldsymbol{x}, \boldsymbol{y})$ is the sum of the lengths of the two legs of the triangle.



Theorem

(Projections on Closed Sets Theorem) *Let U be a closed subset of $\mathbb{R}^n$ such that $U \neq \emptyset$ and let $\mathbf{x}_0 \in \mathbb{R}^n - U$. Then, there exists $\mathbf{x}_1 \in U$ such that $\| \mathbf{x} - \mathbf{x}_0 \|_2 \geqslant \| \mathbf{x}_1 - \mathbf{x}_0 \|_2$ for every $\mathbf{x} \in U$.*

Proof

Let $d = \inf\{\| \boldsymbol{x} - \boldsymbol{x}_0 \|_2 \mid \boldsymbol{x} \in U\}$ and let $U_n = U \cap B\left(\boldsymbol{x}_0, d + \frac{1}{n}\right)$. Note that the sets form a descending sequence of bounded and closed sets $U_1 \supseteq U_2 \supseteq \cdots \supseteq U_n \supseteq \cdots$. Since U_1 is compact, $\bigcap_{n \geqslant 1} U_n \neq \emptyset$. Let $\boldsymbol{x}_1 \in \bigcap_{n \geqslant 1} U_n$. Since $U_n \subseteq U$ for every n, it follows that $\boldsymbol{x}_1 \in U$. Note that $\| \boldsymbol{x}_1 - \boldsymbol{x}_0 \|_2 \leqslant d + \frac{1}{n}$ for every n because $\boldsymbol{x}_1 \in U_n = U \cap B\left(\boldsymbol{x}_0, d + \frac{1}{n}\right)$. This implies $\| \boldsymbol{x}_1 - \boldsymbol{x}_0 \|_2 \leqslant d \leqslant \| \boldsymbol{x} - \boldsymbol{x}_0 \|_2$ for every $\boldsymbol{x} \in U$.

Lemma

Let $a_1, \ldots, a_n$ be n positive numbers. If p and q are two positive numbers such that $p \leqslant q$, then

$$\left(a_1^p + \cdots + a_n^p\right)^{\frac{1}{p}} \geqslant \left(a_1^q + \cdots + a_n^q\right)^{\frac{1}{q}}.$$

Proof: Let $f : \mathbb{R}^{>0} \longrightarrow \mathbb{R}$ be the function defined by

$$f(r) = (a_1^r + \cdots + a_n^r)^{\frac{1}{r}}.$$

Since

$$\ln f(r) = \frac{\ln(a_1^r + \cdots + a_n^r)}{r},$$

it follows that

$$\frac{f'(r)}{f(r)} = -\frac{1}{r^2}\ln(a_1^r + \cdots + a_n^r) + \frac{1}{r} \cdot \frac{a_1^r \ln a_1 + \cdots + a_n^r \ln a_r}{a_1^r + \cdots + a_n^r}.$$

Proof cont'd

To prove that $f'(r) < 0$, it suffices to show that

$$\frac{a_1^r \ln a_1 + \cdots + a_n^r \ln a_r}{a_1^r + \cdots + a_n^r} \leqslant \frac{\ln\left(a_1^r + \cdots + a_n^r\right)}{r}.$$

This last inequality is easily seen to be equivalent to

$$\sum_{i=1}^{n} \frac{a_i^r}{a_1^r + \cdots + a_n^r} \ln \frac{a_i^r}{a_1^r + \cdots + a_n^r} \leqslant 0,$$

which holds because

$$\frac{a_i^r}{a_1^r + \cdots + a_n^r} \leqslant 1$$

for $1 \leqslant i \leqslant n$.

Theorem

Let p and q be two positive numbers such that $p \leqslant q$. For every $\mathbf{u} \in \mathbb{R}^n$, we have $\| \mathbf{u} \|_p \geqslant \| \mathbf{u} \|_q$.

Proof.

This statement follows immediately from previous Lemma. □

Corollary

Let p, q be two positive numbers such that $p \leqslant q$. For every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $d_p(\mathbf{x}, \mathbf{y}) \geqslant d_q(\mathbf{x}, \mathbf{y})$.

Proof.

This statement follows immediately from the previous Theorem. □

Example

For $p = 1$ and $q = 2$ the inequality of the Theorem becomes

$$\sum_{i=1}^{n} |u_i| \leqslant \sqrt{\sum_{i=1}^{n} |u_i|^2},$$

which is equivalent to

$$\frac{\sum_{i=1}^{n} |u_i|}{n} \leqslant \sqrt{\frac{\sum_{i=1}^{n} |u_i|^2}{n}}.$$

Theorem

Let $p \geqslant 1$. For every $\mathbf{x} \in \mathbb{R}^n$ we have

$$\| \mathbf{x} \|_\infty \leqslant \| \mathbf{x} \|_p \leqslant n \| \mathbf{x} \|_\infty .$$

Proof.

Starting from the definition of ν_p we have

$$\| \mathbf{x} \|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \leqslant n^{\frac{1}{p}} \max_{1 \leqslant i \leqslant n} |x_i| = n^{\frac{1}{p}} \| \mathbf{x} \|_\infty .$$

The first inequality is immediate. □

Corollary

Let p and q be two numbers such that $p, q \geqslant 1$. There exist two constants $c, d \in \mathbb{R}_{>0}$ such that

$$c \parallel \mathbf{x} \parallel_q \leqslant \parallel \mathbf{x} \parallel_p \leqslant d \parallel \mathbf{x} \parallel_q$$

for $\mathbf{x} \in \mathbb{R}^n$.

Proof.

Since $\parallel \mathbf{x} \parallel_\infty \leqslant \parallel \mathbf{x} \parallel_p$ and $\parallel x \parallel_q \leqslant n \parallel \mathbf{x} \parallel_\infty$, it follows that $\parallel x \parallel_q \leqslant n \parallel \mathbf{x} \parallel_p$. Exchanging the roles of p and q, we have $\parallel x \parallel_p \leqslant n \parallel \mathbf{x} \parallel_q$, so

$$\frac{1}{n} \parallel \mathbf{x} \parallel_q \leqslant \parallel \mathbf{x} \parallel_p \leqslant n \parallel \mathbf{x} \parallel_q$$

for every $\mathbf{x} \in \mathbb{R}^n$. □

UMASS BOSTON

Corollary

For every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $p \geqslant 1$*, we have* $d_\infty(\mathbf{x}, \mathbf{y}) \leqslant d_p(\mathbf{x}, \mathbf{y}) \leqslant n d_\infty(\mathbf{x}, \mathbf{y})$*.*
Further, for $p, q > 1$*, there exist* $c, d \in \mathbb{R}_{>0}$ *such that*

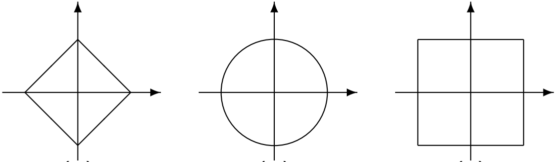
$$c d_q(\mathbf{x}, \mathbf{y}) \leqslant d_p(\mathbf{x}, \mathbf{y}) \leqslant c d_q(\mathbf{x}, \mathbf{y})$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*.*

If $p \leqslant q$, then the closed sphere $B_{d_p}(\mathbf{x}, r)$ is included in the closed sphere $B_{d_q}(\mathbf{x}, r)$. For example, we have

$$B_{d_1}(\mathbf{0}, 1) \subseteq B_{d_2}(\mathbf{0}, 1) \subseteq B_{d_\infty}(\mathbf{0}, 1).$$

In (a) - (c) we represent the closed spheres $B_{d_1}(\mathbf{0}, 1)$, $B_{d_2}(\mathbf{0}, 1)$, and $B_{d_\infty}(\mathbf{0}, 1)$.



(a) (b) (c)

Theorem

Let $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ be $2m$ nonnegative numbers such that $\sum_{i=1}^m x_i = \sum_{i=1}^m y_i = 1$ and let p and q be two positive numbers such that $\frac{1}{p} + \frac{1}{q} = 1$. We have

$$\sum_{j=1}^m x_j^{\frac{1}{p}} y_j^{\frac{1}{q}} \leqslant 1.$$

Proof.

The Hölder inequality applied to $x_1^{\frac{1}{p}}, \ldots, x_m^{\frac{1}{p}}$ and $y_1^{\frac{1}{q}}, \ldots, y_m^{\frac{1}{q}}$ yields the needed inequality

$$\sum_{j=1}^m x_j^{\frac{1}{p}} y_j^{\frac{1}{q}} \leqslant \sum_{j=1}^m x_j \sum_{j=1}^m y_j = 1$$

□