Some basic inequalities

Definition. Let V be a vector space over the complex numbers. An *inner product* is given by a function $\langle\cdot,\cdot\rangle$

$$\begin{aligned} V \times V &\to \mathbb{C} \\ (x,y) &\mapsto \langle x,y\rangle \end{aligned}$$

satisfying the following properties (for all $x \in V$, $y \in V$ and $c \in \mathbb{C}$)

(1) $\langle x+\tilde{x},y\rangle = \langle x,y\rangle + \langle \tilde{x},y\rangle$
(2) $\langle cx,y\rangle = c\langle x,y\rangle$
(3) $\langle y,x\rangle = \overline{\langle x,y\rangle}$
(4) $\langle x,x\rangle \ge 0$ and $\langle x,x\rangle = 0$ if and only of $x=0$.

Note that if $\langle\cdot,\cdot\rangle$ is an inner product then for each y the function $x \mapsto \langle x,y\rangle$ is a linear function. Also we have $\langle x,cy\rangle = \bar{c}\langle x,y\rangle$ and $\langle x,y+\tilde{y}\rangle = \langle x,y\rangle + \langle x,\tilde{y}\rangle$.

Remark: We can also define inner products for vector spaces over $\mathbb{R}$, but then the third axiom is changed to the symmetry axiom $\langle y,x\rangle = \langle x,y\rangle$ for all $x,y \in V$. Thus if V is a vector space over the real numbers then then for each y the function $x \mapsto \langle x,y\rangle$ is a linear function, and for each x the function $y \mapsto \langle x,y\rangle$ is a linear function. The latter statement for $y \mapsto \langle x,y\rangle$ fails in vector spaces over $\mathbb{C}$.

Definition. A semi-norm on a vector space over $\mathbb{C}$ (or over $\mathbb{R}$) is a function $\|\cdot\| : V \to [0,\infty)$ satisfying the following properties for all $x,y \in V$.

(1) $\|x\| \ge 0$
(2) For scalars c, $\|cx\| = |c|\|x\|$.
(3) $\|x+y\| \le \|x\| + \|y\|$ (the triangle inequality).

If in addition we also have the property that and $\|x\| = 0$ only if $x=0$ then we call $\|\cdot\|$ a norm.

1. The Cauchy-Schwarz inequality

Theorem. *Let $\langle\cdot,\cdot\rangle$ be an inner product on V. Then for all $x,y \in V$*

$$|\langle x,y\rangle| \le \sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}.$$

Proof. The inequality is immediate if one of the two vectors is 0. We may thus assume that $y \ne 0$ and therefore $\langle y,y\rangle > 0$. We shall first show the weaker inequality

$$\operatorname{Re}\langle x,y\rangle \le \sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle} \tag{1.1}$$

Let $t \in \mathbb{R}$. We shall use that

$$\langle x+ty, x+ty\rangle \ge 0.$$

Then compute

$$\begin{aligned}\langle x+ty, x+ty\rangle &= \langle x,x\rangle + t\langle x,y\rangle + t\langle y,x\rangle + t^2\langle y,y\rangle \\ &= \langle x,x\rangle + 2t\,\mathrm{Re}\,\langle x,y\rangle + t^2\langle y,y\rangle.\end{aligned}$$

Here we used that for the complex number $z = \langle x,y\rangle$ we have $z+\overline{z} = 2\,\mathrm{Re}\,(z)$.

We have seen that for all $t \in \mathbb{R}$

$$\langle x,x\rangle + 2t\,\mathrm{Re}\,\langle x,y\rangle + t^2\langle y,y\rangle \geq 0.$$

We use this inequality for the special choice $t = -\frac{\mathrm{Re}\,(\langle x,y\rangle)}{\langle y,y\rangle}$ (which happens to be the choice of t that minimizes the quadratic polynomial). Plugging in this value of t yields the inequality

$$\langle x,x\rangle - \frac{(\mathrm{Re}\,\langle x,y\rangle)^2}{\langle y,y\rangle} \geq 0$$

which gives

$$(\mathrm{Re}\,\langle x,y\rangle)^2 \leq \langle x,x\rangle\langle y,y\rangle$$

and (1.1) follows.

Finally let $z := \langle x,y\rangle$. If $z=0$ there is nothing to prove, so assume $z \neq 0$. Then we can write z in polar form, i.e. $z = |z|(\cos\phi + i\sin\phi)$ for some angle ϕ. Let $c = \cos\phi - i\sin\phi$. Then $cz = |z|$ and cz is real and positive. [1] Also $|c| = 1$. Hence we get

$$|\langle x,y\rangle| = c\langle x,y\rangle = \langle cx,y\rangle = \mathrm{Re}\,\langle cx,y\rangle.$$

Applying the already proved inequality (1.1) for the vectors cx and y we see that the last expression is

$$\leq \sqrt{\langle cx,cx\rangle}\sqrt{\langle y,y\rangle} = \sqrt{c\overline{c}\langle x,x\rangle}\sqrt{\langle y,y\rangle} = \sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}\,.$$

This finishes the proof. $\square$

Exercise: Show that equality in Cauchy-Schwarz, $|\langle x,y\rangle| = \sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}$, only happens if x and y are linearly dependent (i.e. one of the two is a scalar multiple of the other).

Definition. We set $\|x\| = \sqrt{\langle x,x\rangle}$.

Theorem The map $x \mapsto \sqrt{\langle x,x\rangle}$ defines a norm on V.

Proof. Setting $\|x\| := \sqrt{\langle x,x\rangle}$ we clearly have that $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$, by property (4) for the inner product. Also $\sqrt{\langle cx,cx\rangle} = \sqrt{c\overline{c}\langle x,x\rangle} = |c|\sqrt{\langle x,x\rangle}$. It remains to prove the triangle inequality.

---

[1] If you prefer not to use polar notation, another equivalent way to define c, given $z = a+bi$ with $z \neq 0$ is to set $c = \frac{a-ib}{\sqrt{a^2+b^2}}$, i.e. $c = \overline{z}/|z|$. Note that $cz = z\overline{z}/|z| = |z|^2/|z| = |z|$.

We compute

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + 2\,\mathrm{Re}\,(\langle x,y\rangle) + \|y\|^2\end{aligned}$$

and by the Cauchy-Schwarz inequality the last expression is

$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

So we have shown $\|x+y\|^2 \leq (\|x\| + \|y\|)^2$ and the triangle inequality follows. $\square$

2. Generalized arithmetic and geometric means

Given two nonnegative numbers a, b we call $\sqrt{ab}$ the geometric mean of a and b. The geometric significance is that the rectangle with sides of length a and b has the same area as the square with sidelength $\sqrt{ab}$. The arithmetic mean is $\frac{a+b}{2}$. The arithmetic mean exceeds the geometric mean:

$$\sqrt{ab} \leq \frac{a+b}{2}\,.$$

This follows immediately from $(\sqrt{a} - \sqrt{b})^2 \geq 0$, i.e. $a + b - 2\sqrt{a}\sqrt{b} \geq 0$ (for nonnegative a, b).

A useful generalization is

Theorem. *Let a, b be nonnegative numbers and let $0 < \vartheta < 1$. Then*

$$a^{1-\vartheta} b^{\vartheta} \leq (1-\vartheta)a + \vartheta b\,. \tag{2.1}$$

Proof. If one of a, b is zero then the inequality is immediate. Let's assume that $a \neq 0$. Then setting $c = b/a$ the assertion is equivalent with

$$c^{\vartheta} \leq (1-\vartheta) + \vartheta c, \text{ for } c \geq 0. \tag{2.2}$$

To prove (2.2) we set

$$f(c) := (1-\vartheta) + \vartheta c - c^{\vartheta}$$

and observe that $f'(c) = \vartheta(1 - c^{\vartheta-1})$. Since by assumption $0 < \vartheta < 1$ we see that $f'(c) \leq 0$ for $0 \leq c \leq 1$ and $f'(c) \geq 0$ for $c \geq 1$. Hence f must have a minimum at $c = 1$. Clearly $f(1) = 0$ and therefore $f(c) \geq 0$ for all $c \geq 0$. Thus (2.2) holds. $\square$

3. The inequalities by Hölder and Minkowski

For vectors $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$ (or in $\mathbb{C}^n$) we define

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}.$$

It is our intention to show that $\|x\|_p$ defines a norm ehen $p > 1$. We shall use the following result (Hölder's inequality) to prove this.

For $p > 1$ we define the conjugate number p' by

$$\frac{1}{p} + \frac{1}{p'} = 1$$

i.e. $p' = \frac{p}{p-1}$.

Theorem: (Hölder's inequality): *Let* $1 < p < \infty$, $1/p + 1/p' = 1$. *For* $x, y \in \mathbb{C}^n$,

$$\Big|\sum_{i=1}^{n} x_i y_i\Big| \le \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} \Big(\sum_{i=1}^{n} |y_i|^{p'}\Big)^{1/p'}$$

or in the above notation

$$\Big|\sum_{i=1}^{n} x_i y_i\Big| \le \|x\|_p \|y\|_{p'}.$$

Remark. When $p = 2$, then $p' = 2$ and Hölder's inequality becomes the Cauchy-Schwarz inequality for the standard scalar product $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$ on $\mathbb{R}^n$ (or the standard scalar product $\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$ on $\mathbb{C}^n$).

Proof of Hölder's inequality. If we replace x with $x/\|x\|_p$ and y with $y/\|y\|_{p'}$ then we see that it is enough to show that

$$\Big|\sum_{i=1}^{n} x_i y_i\Big| \le 1 \text{ provided that } \|x\|_p = 1 \text{ and } \|y\|_{p'} = 1 \tag{3.1}$$

Also it is clearly sufficient to do this for vectors x and y with nonnegative entries (simply replace x_i with $|x_i|$ etc.)

Thus for the rest of the proof we assume that x, y are vectors with nonnegative entries satisfying $\|x\|_p = 1$, $\|y\|_{p'} = 1$.

Set $a_i = x_i^p$, $b_i = y_i^{p'}$. And set $\vartheta = 1 - 1/p$. Since we assume $p > 1$ we see that $0 < \vartheta < 1$. By the inequality for the generalized arithmetic and geometric means we have $a_i^{1-\vartheta} b_i^{\vartheta} \le (1-\vartheta)a_i + \vartheta b_i$ i.e.

$$x_i y_i = a_i^{1/p} b_i^{1-1/p} \le \frac{1}{p} a_i + (1 - \frac{1}{p}) b_i = \frac{1}{p} x_i^p + (1 - \frac{1}{p}) y_i^{p'}$$

Thus

$$\begin{aligned}\sum_{i=1}^{n} x_i y_i &\le \frac{1}{p} \sum_{i=1}^{n} x_i^p + (1 - \frac{1}{p}) \sum_{i=1}^{n} y_i^{p'} \\ &= \frac{1}{p} \|x\|_p^p + (1 - \frac{1}{p}) \|y\|_{p'}^{p'} = \frac{1}{p} + (1 - \frac{1}{p}) = 1;\end{aligned}$$

here we have used that $\|x\|_p = 1$, $\|y\|_{p'} = 1$. □

Remark: Hölder's inequality has extensions to ther settings. One is in Problem 6 on the first homework assignment. Here note that Riemann integrals can be approximated by sums, and so the Hölder inequality with n summands may be useful for similar versions for integrals as well.

The following result called *Minkowski's inequality*[2] establishes the triangle inequality for $\|\cdot\|_p$.

Theorem: *For $x, y \in \mathbb{C}^n$*

$$\Big(\sum_{i=1}^n |x_i + y_i|^p\Big)^{1/p} \le \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^n |y_i|^p\Big)^{1/p}$$

or shortly, $\|x+y\|_p \le \|x\|_p + \|y\|_p$.

Proof. If $x + y = 0$ the inequality is trivial, thus we assume that $x + y \ne 0$ and hence $\|x + y\|_p > 0$

Write

$$\begin{aligned}\|x+y\|_p^p &= \sum_{i=1}^n |x_i + y_i|^p = \sum_{i=1}^n |x_i+y_i|^{p-1}|x_i+y_i| \\ &\le \sum_{i=1}^n |x_i+y_i|^{p-1}(|x_i|+|y_i|) = \sum_{i=1}^n |x_i||x_i+y_i|^{p-1} + \sum_{i=1}^n |y_i||x_i+y_i|^{p-1}\end{aligned}$$

By Hölder's inequality

$$\begin{aligned}\sum_{i=1}^n |x_i||x_i+y_i|^{p-1} &\le \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}\Big(\sum_{i=1}^n |x_i+y_i|^{(p-1)p'}\Big)^{1/p'} \\ &= \|x\|_p\|x+y\|_p^{p-1}\end{aligned}$$

since $(p-1)p' = p$. The same calculation yields

$$\sum_{i=1}^n |y_i||x_i+y_i|^{p-1} \le \|y\|_p\|x+y\|_p^{p-1}\,.$$

We add the two inequalities and we get

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

Divide by $\|x+y\|_p^{p-1}$ and the asserted inequality follows. □

Corollary. Let $1 \le p < \infty$. The expression $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ defines a norm on $\mathbb{C}^n$ (or $\mathbb{R}^n$).

[2]Minkowski is pronounced "Minkoffski"