Exercise Problem Sets 5

Oct. 14. 2022

Problem 1. Let $\|\cdot\|:\mathbb{F}^n\to\mathbb{R}$, where $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, be defined by

$$\|\boldsymbol{x}\|_p \equiv \begin{cases} \Big(\sum\limits_{i=1}^n |x_i|^p\Big)^{\frac{1}{p}} & \text{if } 1\leqslant p<\infty\,,\\ \max\big\{|x_1|,\cdots,|x_n|\big\} & \text{if } p=\infty\,,\end{cases} \qquad \boldsymbol{x}=(x_1,\cdots,x_n)\,.$$

Complete the following.

1. Prove the Hölder inequality $\big|\langle \boldsymbol{x},\boldsymbol{y}\rangle\big| \leqslant \|\boldsymbol{x}\|_p\|\boldsymbol{y}\|_q$ for all $\boldsymbol{x},\boldsymbol{y}\in\mathbb{F}^n$, where $p,q\in[1,\infty]$ satisfy $\dfrac{1}{p}+\dfrac{1}{q}=1$.
2. Show that $\|\cdot\|_p$ is indeed a norm on $\mathbb{F}^n$ for all $1\leqslant p\leqslant\infty$.
3. Show that $\|\boldsymbol{x}\|_\infty = \lim\limits_{p\to\infty}\|\boldsymbol{x}\|_p$ for all $\boldsymbol{x}\in\mathbb{F}^n$.
4. Show that for each $1\leqslant p,q\leqslant\infty$ and $p\neq q$, $\|\cdot\|_p$ and $\|\cdot\|_q$ are equivalent norms.

Hint: 1. Prove first the Young inequality (if you do not know this inequality)

$$ab\leqslant \frac{1}{p}a^p+\frac{1}{q}b^q \qquad \forall\, a,b\geqslant 0 \text{ and } p,q\in(1,\infty) \text{ satisfying } \frac{1}{p}+\frac{1}{q}=1\,,$$

Proof. 1. First we prove the Young inequality. Suppose that $1<p<\infty$. Consider the function $y=f(x)=x^{p-1}$. The inverse function of f is $y=f^{-1}(x)=x^{\frac{1}{p-1}}$. For $a,b>0$, we do not necessarily have $a^{p-1}=b$; thus by the convexity of f we have

$$\int_0^a f(x)\,dx+\int_0^b f^{-1}(x)\,dx\geqslant ab\,.$$

The inequality above implies that

$$ab\leqslant\int_0^a x^{p-1}\,dx+\int_0^b x^{\frac{1}{p-1}}\,dx=\frac{1}{p}a^p+\frac{1}{1+\frac{1}{p-1}}b^{\frac{1}{p-1}+1}=\frac{1}{p}a^p+\frac{1}{q}b^q$$

since $q=\dfrac{p}{p-1}$.

Now suppose that $1<p<\infty$. Let $\boldsymbol{x}=(x_1,\cdots,x_n)$ and $\boldsymbol{y}=(y_1,\cdots,y_n)$ be given, and $q=\dfrac{p}{p-1}$ be the Hölder conjugate of p satisfying $\dfrac{1}{p}+\dfrac{1}{q}=1$. By Young's inequality, we find that

$$\frac{|x_i|}{\|\boldsymbol{x}\|_p}\frac{|y_i|}{\|\boldsymbol{y}\|_q}\leqslant\frac{1}{p}\Big(\frac{|x_i|}{\|\boldsymbol{x}\|_p}\Big)^p+\frac{1}{q}\Big(\frac{|y_i|}{\|\boldsymbol{y}\|_q}\Big)^q=\frac{1}{p}\frac{|x_i|^p}{\|\boldsymbol{x}\|_p^p}+\frac{1}{q}\frac{|y_i|^q}{\|\boldsymbol{y}\|_q^q}\,;$$

thus

$$\Big|\sum_{i=1}^n\frac{x_i}{\|\boldsymbol{x}\|_p}\frac{y_i}{\|\boldsymbol{y}\|_q}\Big|\leqslant\sum_{i=1}^n\frac{|x_i|}{\|\boldsymbol{x}\|_p}\frac{|y_i|}{\|\boldsymbol{y}\|_q}\leqslant\frac{1}{p}\frac{1}{\|\boldsymbol{x}\|_p^p}\sum_{i=1}^n|x_i|^p+\frac{1}{q}\frac{1}{\|\boldsymbol{y}\|_q^q}\sum_{i=1}^n|y_i|^q=\frac{1}{p}+\frac{1}{q}=1\,.$$

If $p = \infty$, then $q = 1$ and clearly we have

$$\Big|\sum_{i=1}^{n} x_i y_i\Big| \leqslant \sum_{i=1}^{n} |x_i||y_i| \leqslant \sum_{i=1}^{n} \Big(\max_{1\leqslant i\leqslant n} |x_i|\Big)|y_i| = \|\boldsymbol{x}\|_\infty \sum_{i=1}^{n} |y_i| = \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1 \,.$$

The case that $p = 1$ can be proved in a similar fashion.

2. The case for $p = 1$ and $p = \infty$ are trivial:

$$\|\boldsymbol{x} + \boldsymbol{y}\|_1 = \sum_{k=1}^{n} |x_k + y_k| \leqslant \sum_{k=1}^{n} \big(|x_k| + |y_k|\big) = \sum_{k=1}^{n} |x_k| + \sum_{k=1}^{n} |y_k| = \|\boldsymbol{x}\|_1 + \|\boldsymbol{y}\|_1 \,,$$

and for all $1 \leqslant j \leqslant n$,

$$|x_j + y_j| \leqslant |x_j| + |y_j| \leqslant \max\big\{|x_1|, \cdots, |x_n|\big\} + \max\big\{|y_1|, \cdots, |y_n|\big\} = \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty$$

so that by the fact that $\|\boldsymbol{x} + \boldsymbol{y}\|_\infty = |x_k + y_k|$ for some $1 \leqslant k \leqslant n$ we obtain that

$$\|\boldsymbol{x} + \boldsymbol{y}\|_\infty \leqslant \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty \,.$$

For the case $1 < p < \infty$, having established Hölder's inequality we find that

$$\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|_p^p &= \sum_{i=1}^{n} |x_i + y_i|^p \leqslant \sum_{i=1}^{n} |x_i + y_i|^{p-1}|x_i| + \sum_{i=1}^{n} |x_i + y_i|^{p-1}|y_i| \\
&\leqslant \Big[\sum_{i=1}^{n} \big(|x_i + y_i|^{p-1}\big)^{\frac{p}{p-1}}\Big]^{\frac{p-1}{p}} \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \\
&\quad + \Big[\sum_{i=1}^{n} \big(|x_i + y_i|^{p-1}\big)^{\frac{p}{p-1}}\Big]^{\frac{p-1}{p}} \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{\frac{1}{p}} \\
&= \Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{\frac{p-1}{p}} \big(\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p\big) = \|\boldsymbol{x} + \boldsymbol{y}\|_p^{p-1}\big(\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p\big) \,.
\end{aligned}$$

Therefore, $\|\boldsymbol{x} + \boldsymbol{y}\|_p \leqslant \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$ for $1 < p < \infty$ as well.

3. W.L.O.G. we can assume that $\boldsymbol{x} \neq \boldsymbol{0}$. Suppose that $\|\boldsymbol{x}\|_\infty = |x_k|$ for some $1 \leqslant k \leqslant n$. Then

$$\|\boldsymbol{x}\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \geqslant |x_k| = \|\boldsymbol{x}\|_\infty \,.$$

Moreover, $|x_j| \leqslant |x_k|$ for all $1 \leqslant j \leqslant n$; thus

$$\|\boldsymbol{x}\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} = |x_k|\Big[\sum_{i=1}^{n} \Big(\frac{|x_i|}{|x_k|}\Big)^p\Big]^{\frac{1}{p}} \leqslant |x_k|\Big(\sum_{i=1}^{n} 1^p\Big)^{\frac{1}{p}} = |x_k| n^{\frac{1}{p}} \,;$$

thus

$$\|\boldsymbol{x}\|_\infty \leqslant \|\boldsymbol{x}\|_p \leqslant \|\boldsymbol{x}\|_\infty n^{\frac{1}{p}} \,.$$

By the fact that $\lim\limits_{p\to\infty} n^{\frac{1}{p}} = 1$, the Sandwich Lemma implies that $\lim\limits_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty$.

4. It suffices to show that every p-norm is equivalent to the ∞-norm since if so, then for all $1 \leqslant p, q < \infty$ there exist C_1, C_2, C_3, C_4 such that

$$C_1 \|\boldsymbol{x}\|_p \leqslant \|\boldsymbol{x}\|_\infty \leqslant C_2 \|\boldsymbol{x}\|_p \quad \text{and} \quad C_3 \|\boldsymbol{x}\|_q \leqslant \|\boldsymbol{x}\|_\infty \leqslant C_4 \|\boldsymbol{x}\|_q \qquad \forall\, \boldsymbol{x} \in \mathbb{F}^n .$$

Therefore,

$$\frac{C_1}{C_4} \|\boldsymbol{x}\|_p \leqslant \|\boldsymbol{x}\|_q \leqslant \frac{C_2}{C_3} \|\boldsymbol{x}\|_p \qquad \forall\, \boldsymbol{x} \in \mathbb{F}^n .$$

Now we show that each p-norm is equivalent to the ∞-norm. Note that

$$\|\boldsymbol{x}\|_\infty \leqslant \|\boldsymbol{x}\|_p \qquad \forall\, 1 \leqslant p \leqslant \infty .$$

On the other hand,

$$\|\boldsymbol{x}\|_p = \Big(\sum_{i=1}^n |x_i|^p \Big)^{\frac{1}{p}} \leqslant \Big(\sum_{i=1}^n \|\boldsymbol{x}\|_\infty^p \Big)^{\frac{1}{p}} \leqslant n^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty .$$

Therefore,

$$n^{-\frac{1}{p}} \|\boldsymbol{x}\|_p \leqslant \|\boldsymbol{x}\|_\infty \leqslant \|\boldsymbol{x}\|_p \qquad \forall\, \boldsymbol{x} \in \mathbb{F}^n \text{ and } 1 \leqslant p \leqslant \infty .$$

□

Problem 2. Complete the following.

1. For $f \in \mathscr{C}([a, b]; \mathbb{R})$, define

$$\|f\|_p = \begin{cases} \Big(\displaystyle\int_a^b |f(x)|^p \, dx \Big)^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty , \\ \displaystyle\max_{x \in [a,b]} \big| f(x) \big| & \text{if } p = \infty . \end{cases}$$

Show that $\| \cdot \|_p$ is a norm on $\mathscr{C}([a, b]; \mathbb{R})$.

2. Show that $\|f\|_\infty = \lim\limits_{p \to \infty} \|f\|_p$ for all $f \in \mathscr{C}([a, b]; \mathbb{R})$.

3. Are $\| \cdot \|_p$ and $\| \cdot \|_q$ equivalent norms on $\mathscr{C}([a, b]; \mathbb{R})$ for any $1 \leqslant p, q \leqslant \infty$?

Proof. 1. For a continuous function $h : [a, b] \to \mathbb{R}$,

$$\int_a^b h(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^n h\big(a + i \frac{b-a}{n}\big) \frac{b-a}{n} .$$

Therefore, with c_i and d_i denoting $f\big(a + i\frac{b-a}{n}\big)$ and $g\big(a + i\frac{b-a}{n}\big)$, respectively, we have

$$\|f + g\|_p = \lim_{n \to \infty} \Big(\sum_{i=1}^n \Big| (f+g)\big(a + i\frac{b-a}{n}\big) \Big|^p \frac{b-a}{n} \Big)^{\frac{1}{p}} = (b-a)^{\frac{1}{p}} \lim_{n \to \infty} \Big[n^{-\frac{1}{p}} \Big(\sum_{i=1}^n |c_i + d_i|^p \Big)^{\frac{1}{p}} \Big] ,$$

and similarly,

$$\|f\|_p = (b-a)^{\frac{1}{p}} \lim_{n \to \infty} \Big[n^{-\frac{1}{p}} \Big(\sum_{i=1}^n |c_i|^p \Big)^{\frac{1}{p}} \Big] , \qquad \|g\|_p = (b-a)^{\frac{1}{p}} \lim_{n \to \infty} \Big[n^{-\frac{1}{p}} \Big(\sum_{i=1}^n |d_i|^p \Big)^{\frac{1}{p}} \Big] .$$

By Minkowski's inequality in Problem 1,

$$n^{-\frac{1}{p}} \Big(\sum_{i=1}^n |c_i + d_i|^p \Big)^{\frac{1}{p}} \leqslant n^{-\frac{1}{p}} \Big(\sum_{i=1}^n |c_i|^p \Big)^{\frac{1}{p}} + n^{-\frac{1}{p}} \Big(\sum_{i=1}^n |d_i|^p \Big)^{\frac{1}{p}} ;$$

thus the desired conclusion follows from passing to the limit as $n \to \infty$.

2. By the Extreme Value Theorem (Problem 3 of Exercise 4), there exists $c \in [a, b]$ such that

$$\big|f(c)\big| = \max_{x \in [a,b]} \big|f(x)\big| = \|f\|_\infty \,.$$

W.L.O.G. we can assume that $f(c) > 0$.

Let $n \in \mathbb{N}$ be given. Then by the continuity of f, there exists $\delta_n > 0$ such that

$$\big|f(x) - f(c)\big| < \frac{1}{n} \quad \text{whenever} \quad x \in \mathrm{I}_n \equiv (c - \delta_n, c + \delta_n) \cap [a, b] \,.$$

Then for $n \gg 1$,

$$\big|f(x)\big| > \big|f(c)\big| - \frac{1}{n} \quad \text{whenever} \quad x \in \mathrm{I}_n \,.$$

Therefore, for $n \gg 1$,

$$\begin{aligned}
\|f\|_p &= \Big(\int_a^b \big|f(x)\big|^p \, dx\Big)^{\frac{1}{p}} \geqslant \Big(\int_{\mathrm{I}_n} \big|f(x)\big|^p \, dx\Big)^{\frac{1}{p}} \geqslant \Big(|f(c)| - \frac{1}{n}\Big)\Big(\int_{\mathrm{I}_n} dx\Big)^{\frac{1}{p}} \\
&= \Big(\|f\|_\infty - \frac{1}{n}\Big)|\mathrm{I}_n|^{\frac{1}{p}} \,;
\end{aligned}$$

thus for all $n \gg 1$,

$$\Big(\|f\|_\infty - \frac{1}{n}\Big)|\mathrm{I}_n|^{\frac{1}{p}} \leqslant \|f\|_p \leqslant \|f\|_\infty (b - a)^{\frac{1}{p}} \,.$$

Therefore, passing to the limit as $p \to \infty$, we find that for $n \gg 1$,

$$\|f\|_\infty - \frac{1}{n} \leqslant \liminf_{p \to \infty} \|f\|_p \leqslant \limsup_{p \to \infty} \|f\|_p \leqslant \|f\|_\infty \,.$$

Therefore, passing to the limit as $n \to \infty$, we find that

$$\|f\|_\infty = \liminf_{p \to \infty} \|f\|_p = \limsup_{p \to \infty} \|f\|_p = \|f\|_\infty \,;$$

thus $\lim\limits_{p \to \infty} \|f\|_p = \|f\|_\infty$.

3. The 1-norma and the ∞-norm are not equivalent. For each $n \in \mathbb{N}$, consider the function $f_n : [0, 1] \to \mathbb{R}$ defined by

$$f_n(x) = \begin{cases} -n^2 x + n & \text{if } 0 \leqslant x \leqslant \dfrac{1}{n} \,, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\|f_n\|_1 = \dfrac{1}{2}$ but $\|f_n\|_\infty = n$. Therefore,

$$\frac{\|f_n\|_\infty}{\|f_n\|_1} = 2n$$

which does not belong to any given bounded interval $[C_1, C_2]$ when n is large. In fact, any p-norm and q-norm cannot be equivalent since for every $n > 0$ one can also find a function $f : [0, 1] \to \mathbb{R}$ such that $\|f\|_p = 1$ and $\|f\|_q > n$ if $p < q$. □

Problem 3. A set A in a vector space $\mathcal{V}$ is called ***convex*** if for all $x, y \in A$, the line segment joining x and y, denoted by $\overline{xy}$, lies in A.

1. Show that open r-balls and closed r-balls in any normed vector spaces are convex.

2. Show that an ellipsoid $E \equiv \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \boldsymbol{x}^{\mathrm{T}} Q \boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{c} \leqslant 0\}$ is convex, where Q is a positive semi-definite $n \times n$ matrix and $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$ are vectors.

3. Show that a polytope $P \equiv \{\boldsymbol{x} \in \mathbb{R}^n \,|\, A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ is convex, where A is an $m \times n$ matrix, $\boldsymbol{b} \in \mathbb{R}^m$ is a vector, and $\leqslant$ is defined by $\boldsymbol{c} \leqslant \boldsymbol{d}$ if and only if $c_i \leqslant d_i$ for all components.

4. Show that if C_α is convex for all $\alpha \in I$, then $\bigcap_{\alpha \in I} C_\alpha$ is convex.

5. Show that if $C_1, C_2, \cdots, C_N$ are convex sets and $\mu_1, \mu_2, \cdots, \mu_N$ are real numbers, then

$$\mu_1 C_1 + \mu_2 C_2 + \cdots + \mu_N C_N \equiv \{\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_N x_n \,|\, x_k \in C_k \text{ for } 1 \leqslant k \leqslant N\}$$

   is convex.

6. Show that if $C_1, C_2, \cdots, C_N$ are convex sets, then the Cartesian product $C_1 \times C_2 \times \cdots \times C_N$ is convex.

7. Let A be an $m \times n$ matrix, and C be a convex set in $\mathbb{R}^n$, D be a convex set in $\mathbb{R}^m$. Show that

$$A(C) \equiv \{A\boldsymbol{x} \,|\, \boldsymbol{x} \in C\} \qquad \text{and} \qquad A^{-1}(D) \equiv \{\boldsymbol{x} \,|\, A\boldsymbol{x} \in D\}$$

   are convex.

8. Let $\Delta_k \equiv \left\{(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k \,\middle|\, 0 \leqslant \lambda_i \leqslant 1 \text{ for all } 1 \leqslant i \leqslant k \text{ and } \sum_{i=1}^{k} \lambda_i = 1\right\}$. A convex combination of k vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$ is a sum of the form $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k$ for some $(\lambda_1, \cdots, \lambda_k) \in \Delta_k$. Show that the collection of all convex combinations of k given vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ is convex; that is, show that

$$\{\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k \,|\, (\lambda_1, \lambda_2, \cdots, \lambda_k) \in \Delta_k\}$$

   is convex.

9. Let S be a subset in $\mathbb{R}^n$. Show that the collection of all convex combinations of finitely many vectors from S is convex; that is, show that

$$\{\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k \,|\, k \in \mathbb{N}, \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S \text{ and } (\lambda_1, \cdots, \lambda_k) \in \Delta_k\}$$

   is convex. The set defined above is called the convex full of S and is sometimes denoted by $\mathrm{conv}(S)$.

Problem 4. Show that

$$\|A\|_1 = \max\left\{\sum_{i=1}^n |a_{i1}|, \sum_{i=1}^n |a_{i2}|, \cdots, \sum_{i=1}^n |a_{im}|\right\} \qquad \forall\, A \in \mathcal{M}_{n\times m}\,.$$

Hint: Mimic the computation of $\|A\|_\infty$ in Example 2.19 of the lecture note, or make use of Problem 6 of Exercise 2.

Proof. By Problem 6 of Exercise 2.

$$\|\boldsymbol{x}\|_1 = \sup_{\|\boldsymbol{y}\|_\infty=1} \boldsymbol{x}\cdot\boldsymbol{y} \qquad \text{and} \qquad \|\boldsymbol{y}\|_\infty = \sup_{\|\boldsymbol{x}\|_1=1} \boldsymbol{x}\cdot\boldsymbol{y}\,,$$

where $\boldsymbol{x}\cdot\boldsymbol{y}$ denotes the standard inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ in the Euclidean space. Therefore,

$$\|A\|_1 = \sup_{\|\boldsymbol{x}\|_1=1} \|A\boldsymbol{x}\|_1 = \sup_{\|\boldsymbol{x}\|=1}\sup_{\|\boldsymbol{y}\|_\infty=1} (A\boldsymbol{x})\cdot\boldsymbol{y} = \sup_{\|\boldsymbol{x}\|_1=1}\sup_{\|\boldsymbol{y}\|_\infty=1} \boldsymbol{x}\cdot(A^{\mathrm{T}}\boldsymbol{y})\,,$$

and Problem 5 of Exercise 2 further implies that

$$\|A\|_1 = \sup_{\|\boldsymbol{y}\|_\infty=1}\sup_{\|\boldsymbol{x}\|_1=1} (A^{\mathrm{T}}\boldsymbol{y})\cdot\boldsymbol{x} = \sup_{\|\boldsymbol{y}\|_\infty=1} \|A^{\mathrm{T}}\boldsymbol{y}\|_\infty = \|A^{\mathrm{T}}\|_\infty\,.$$

By the fact that the ∞-norm of an $n\times m$ real matrix is the maximum of the sum of the absolute value of entries of all row vectors, we find that

$$\|A\|_1 = \|A^{\mathrm{T}}\|_\infty = \max\left\{\sum_{i=1}^n |a_{i1}|, \sum_{i=1}^n |a_{i2}|, \cdots, \sum_{i=1}^n |a_{im}|\right\}\,. \qquad \square$$

Alternative proof. Let $\boldsymbol{x} = (x_1, \cdots, x_m) \in \mathbb{R}^m$ and $\|\boldsymbol{x}\|_1 = 1$. Then for $A = [a_{ij}] \in \mathcal{M}_{n\times m}$, we have

$$\begin{aligned}
\|A\boldsymbol{x}\|_1 &= \sum_{i=1}^n \Big|\sum_{j=1}^m a_{ij}x_j\Big| \leqslant \sum_{i=1}^n\sum_{j=1}^m |a_{ij}||x_j| = \sum_{j=1}^m\sum_{i=1}^n |a_{ij}||x_j| = \sum_{j=1}^m |x_j|\Big(\sum_{i=1}^n |a_{ij}|\Big) \\
&\leqslant \sum_{j=1}^m |x_j|\Big(\max_{1\leqslant j\leqslant m}\sum_{i=1}^n |a_{ij}|\Big) = \Big(\max_{1\leqslant j\leqslant m}\sum_{i=1}^n |a_{ij}|\Big)\sum_{j=1}^m |x_j| = \Big(\max_{1\leqslant j\leqslant m}\sum_{i=1}^n |a_{ij}|\Big)\|\boldsymbol{x}\|_1 \\
&= \max_{1\leqslant j\leqslant m}\sum_{i=1}^n |a_{ij}|\,.
\end{aligned}$$

Therefore, $\|A\|_1 = \sup\limits_{\|\boldsymbol{x}\|_1=1} \|A\boldsymbol{x}\|_1 \leqslant \max\limits_{1\leqslant j\leqslant m}\sum\limits_{i=1}^n |a_{ij}|$.

On the other hand, suppose that $\max\limits_{1\leqslant j\leqslant m}\sum\limits_{i=1}^n |a_{ij}| = \sum\limits_{i=1}^n |a_{ik}|$; that is, the maximum of the sum of absolute value of column entries of A occurs at the k-th column. Let $\boldsymbol{x} = (x_1, \cdots, x_m) \in \mathbb{R}^m$ be defined by

$$x_j = \begin{cases} 0 & \text{if } j \neq k\,, \\ 1 & \text{if } j = k\,. \end{cases}$$

Then

$$\|A\boldsymbol{x}\|_1 = \sum_{i=1}^n \Big|\sum_{j=1}^m a_{ij}x_j\Big| = \sum_{i=1}^n |a_{ik}| = \max_{1\leqslant j\leqslant m}\sum_{i=1}^n |a_{ij}|\,;$$

thus $\|A\|_1 = \sup\limits_{\|\boldsymbol{x}\|_1=1} \|A\boldsymbol{x}\|_1 \geqslant \max\limits_{1\leqslant j\leqslant m}\sum\limits_{i=1}^n |a_{ij}|$. $\square$

Problem 5. Let $\mathcal{M}_{n\times m}(\mathbb{F})$ be collection of $n \times m$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For $A \in \mathcal{M}_{n\times m}(\mathbb{F})$, define

$$\|A\|_p = \sup_{\|\boldsymbol{x}\|_p=1} \|A\boldsymbol{x}\|_p = \sup_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|A\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}.$$

1. Show that $\|\cdot\|_p$ is a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$.
2. Show that $\|A\|_2 = \sqrt{\text{the maximum eigenvalue of } A^\dagger A}$, where $A^\dagger$ is the conjugate transpose of A.
3. Show that $\|A\|_\infty = \max\Big\{\sum\limits_{k=1}^m |a_{1k}|, \sum\limits_{k=1}^m |a_{2k}|, \cdots, \sum\limits_{k=1}^m |a_{nk}|\Big\}$ if $A \in \mathcal{M}_{n\times m}(\mathbb{F})$.
4. Show that $\|A\|_1 = \max\Big\{\sum\limits_{k=1}^n |a_{k1}|, \sum\limits_{k=1}^n |a_{k2}|, \cdots, \sum\limits_{k=1}^n |a_{km}|\Big\}$ if $A \in \mathcal{M}_{n\times m}(\mathbb{F})$.
5. Show that $\|A\|_2^2 \leqslant \|A\|_1 \|A\|_\infty$ for all $A \in \mathcal{M}_{n\times m}(\mathbb{F})$.

Proof. The proofs of 1,2,4 are identical to the proof for the case of $\mathbb{F} = \mathbb{R}$ and are given in Example 2.19 and Problem 4.

3. It suffices to show the case $\mathbb{F} = \mathbb{C}$ and A is not zero matrix. Let $\boldsymbol{x} \in \mathbb{C}^m$. If $\|\boldsymbol{x}\|_\infty = 1$, then for each $1 \leqslant i \leqslant n$,

$$|a_{i1}x_1 + a_{i2}x_2 + \cdots a_{im}x_m| \leqslant \sum_{j=1}^m |a_{ij}| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^m |a_{ij}|;$$

thus the absolute value of each component of $A\boldsymbol{x}$, under the constraint $\|\boldsymbol{x}\|_\infty = 1$, has an upper bound $\max\limits_{1\leqslant i\leqslant n} \sum\limits_{j=1}^m |a_{ij}|$. Therefore,

$$\|A\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1} \|A\boldsymbol{x}\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1} \max_{1\leqslant i\leqslant n} |a_{i1}x_1 + a_{i2}x_2 + \cdots a_{im}x_m| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^m |a_{ij}|.$$

On the other hand, assume $\max\limits_{1\leqslant i\leqslant n} \sum\limits_{j=1}^m |a_{ij}| = \sum\limits_{j=1}^m |a_{kj}|$ for some $1 \leqslant k \leqslant n$. Let $\beta_j \in \mathbb{C}$ satisfy

$$\beta_j a_{jk} = |a_{kj}|$$

and define

$$\boldsymbol{x} = (\beta_1, \beta_2, \cdots, \beta_n)^{\mathrm{T}}.$$

Then $\|\boldsymbol{x}\|_\infty = 1$ (since A is not zero matrix so that $\max\{|b_1|, \cdots, |b_n|\} = 1$), and $\|A\boldsymbol{x}\|_\infty = \sum\limits_{j=1}^m |a_{kj}|$; thus

$$\|A\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1} \|A\boldsymbol{x}\|_\infty \geqslant \sum_{j=1}^m |a_{kj}| = \max_{1\leqslant i\leqslant n} \sum_{j=1}^m |a_{ij}|.$$

The combination of the two inequalities above implies the desired identity.

5. Let $\lambda \geqslant 0$ be the largest eigenvalue of $A^\dagger A$ with corresponding eigenvector $\boldsymbol{v}$. Then $A^\dagger A\boldsymbol{v} = \lambda\boldsymbol{v}$ so that 2 implies that

$$\|A\|_2^2\|\boldsymbol{v}\|_1 = \lambda\|\boldsymbol{v}\|_1 = \|A^\dagger A\boldsymbol{v}\|_1 \leqslant \|A^\dagger\|_1\|A\boldsymbol{v}\|_1 \leqslant \|A^\dagger\|_1\|A\|_1\|\boldsymbol{v}\|_1\,;$$

thus by the fact (from 3 and 4) that $\|A^\dagger\|_1 = \|A\|_\infty$ and $\|\boldsymbol{v}\|_1 \neq 0$, we conclude the desired inequality. □

Problem 6. Let $\mathcal{M}_{n\times m}(\mathbb{F})$ be the collection of all $n \times m$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define a function $\|\cdot\|_{p,q} : \mathcal{M}_{n\times m}(\mathbb{F}) \to \mathbb{R}$ by

$$\|A\|_{p,q} = \sup_{\|\boldsymbol{x}\|_p=1} \|A\boldsymbol{x}\|_q\,,$$

here we recall that $\|\cdot\|_p$ is the p-norm on $\mathbb{F}^n$ given in Problem 1. If $p = q$, we simply use $\|A\|_p$ to denote $\|A\|_{p,q}$. Complete the following.

1. Show that $\|A\|_{p,q} = \sup\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p}$ for all $p, q \geqslant 1$.

2. Show that $\|A\|_{p,q} = \inf\big\{M \in \mathbb{F} \,\big|\, \|A\boldsymbol{x}\|_q \leqslant M\|\boldsymbol{x}\|_p \ \forall\, \boldsymbol{x} \in \mathbb{F}^m\big\}$.

3. $\|A\boldsymbol{x}\|_q \leqslant \|A\|_{p,q}\|\boldsymbol{x}\|_p$ for all $\boldsymbol{x} \in \mathbb{F}^m$.

4. $\|\cdot\|_{p,q}$ defines a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$.

5. Let $\{A_k\}_{k=1}^\infty \subseteq \mathcal{M}_{n\times m}(\mathbb{F})$. Show that $\lim\limits_{k\to\infty} \|A_k\|_{p,q} = 0$ if and only if each entry of A_k converges to 0. In other words, by writing $A_k = \big[a_{ij}^{(k)}\big]_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$, show that $\lim\limits_{k\to\infty} \|A_k\|_{p,q} = 0$ if and only if $\lim\limits_{k\to\infty} a_{ij}^{(k)} = 0$ for all $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$. In particular, $A_k \to A$ in the sense that $\|A_k - A\|_{p,q} \to 0$ as $k \to \infty$ if and only if the (i, j)-th entry of A_k converges to (i, j)-th entry of A for all $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$.

Proof. 1. If $\boldsymbol{x} \neq \boldsymbol{0}$, then $\boldsymbol{y} = \dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_p}$ satisfies that $\|\boldsymbol{y}\|_p = 1$; thus if $\boldsymbol{x} \neq \boldsymbol{0}$,

$$\frac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p} = \|A\boldsymbol{y}\|_q \leqslant \sup_{\|\boldsymbol{x}\|_p=1} \|A\boldsymbol{x}\|_q = \|A\|_{p,q}\,.$$

Therefore, $\sup\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p} \leqslant \|A\|_{p,q}$.

On the other hand, if $\|\boldsymbol{x}\|_p = 1$, then $\boldsymbol{x} \neq \boldsymbol{0}$; thus if $\|\boldsymbol{x}\|_p = 1$,

$$\|A\boldsymbol{x}\|_q = \frac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p} \leqslant \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p}\,.$$

Therefore, $\|A\|_{p,q} = \sup\limits_{\|\boldsymbol{x}\|_p=1} \|A\boldsymbol{x}\|_q \leqslant \sup\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p}$.

2. 2 follows from Problem 4 of Exercise 2.

3. By 1, $\dfrac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p} \leqslant \|A\|_{p,q}$ for all $\boldsymbol{x} \neq \mathbf{0}$ or equivalently,

$$\|A\boldsymbol{x}\|_q \leqslant \|A\|_{p,q}\|\boldsymbol{x}\|_p \qquad \forall\, \boldsymbol{x} \neq \mathbf{0}\,.$$

Since the inequality above also holds for $\boldsymbol{x} = \mathbf{0}$, we conclude that

$$\|A\boldsymbol{x}\|_q \leqslant \|A\|_{p,q}\|\boldsymbol{x}\|_p \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^m\,.$$

4. The proof of 4 is similar to the proof of that $\|\cdot\|_p$ is a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$.

5. Let $B = [b_{ij}] \in M_{n\times m}$, and $|b_{k\ell}| = \max\limits_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m} |b_{ij}|$; that is, the maximum of the absolute value of entries of B occurs at the (k,ℓ)-entry. Let $\boldsymbol{e}_\ell$ be the unit vector whose ℓ-th component is 1. Since $B\boldsymbol{e}_\ell$ is the ℓ-th column of B, for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$,

$$|b_{ij}| \leqslant |b_{k\ell}| \leqslant \|B\boldsymbol{e}_\ell\|_q \leqslant \|B\|_{p,q}\|\boldsymbol{e}_\ell\|_p = \|B\|_{p,q}\,;$$

thus

$$|b_{ij}| \leqslant \|B\|_{p,q} \qquad \forall\, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\,. \tag{$\star$}$$

On the other hand, there exists $\boldsymbol{x} \in \mathbb{R}^m$ such that $\|\boldsymbol{x}\|_p = 1$ and $\|B\boldsymbol{x}\|_q \geqslant \dfrac{\|B\|_{p,q}}{2}$. Therefore, if $1 \leqslant q < \infty$,

$$\begin{aligned}\frac{\|B\|_{p,q}}{2} &\leqslant \|B\boldsymbol{x}\|_q = \Big(\sum_{i=1}^n \Big|\sum_{j=1}^m b_{ij}x_j\Big|^q\Big)^{\frac{1}{q}} \leqslant \Big[\sum_{i=1}^n \Big(\sum_{j=1}^m |b_{ij}|\Big)^q\Big]^{\frac{1}{q}} \leqslant m\Big[\sum_{i=1}^n \Big(\frac{1}{m}\sum_{j=1}^m |b_{ij}|\Big)^q\Big]^{\frac{1}{q}} \\ &\leqslant m\Big(\sum_{i=1}^n \frac{1}{m}\sum_{j=1}^m |b_{ij}|^q\Big)^{\frac{1}{q}} \leqslant m^{1-\frac{1}{q}}\Big(\sum_{i=1}^n\sum_{j=1}^m |b_{ij}|^q\Big)^{\frac{1}{q}} \leqslant m\Big(\sum_{i=1}^n\sum_{j=1}^m |b_{ij}|^q\Big)^{\frac{1}{q}},\end{aligned}$$

while if $q = \infty$,

$$\frac{\|B\|_{p,q}}{2} \leqslant \|B\boldsymbol{x}\|_\infty = \max_{1\leqslant i\leqslant n}\Big|\sum_{j=1}^m b_{ij}x_j\Big| \leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^m |b_{ij}| \leqslant \sum_{i=1}^n\sum_{j=1}^m |b_{ij}|\,.$$

In either cases, we conclude that

$$\|B\|_{p,q} \leqslant f(|b_{11}|, |b_{12}|, \cdots, |b_{nm}|) \tag{$\diamond$}$$

for some function f of nm variables satisfying that $f(\boldsymbol{y}) \to 0$ as $\boldsymbol{y} \to 0$.

($\Rightarrow$) Using ($\star$), we find that for each $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$,

$$0 \leqslant \big|a_{ij}^{(k)}\big| \leqslant \|A_k\|_{p,q}\,.$$

Since $\lim\limits_{k\to\infty} \|A_k\|_{p,q} = 0$, by the Sandwich Lemma we conclude that

$$\lim_{k\to\infty}\big|a_{ij}^{(k)}\big| = 0 \qquad \forall\, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\,.$$

($\Leftarrow$) Suppose that $\lim\limits_{k\to\infty} \big|a_{ij}^{(k)}\big| = 0$ for all $1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m$. Then ($\diamond$) implies that

$$0 \leqslant \|A_k\|_{p,q} \leqslant f\big(\big|a_{11}^{(k)}\big|, \big|a_{12}^{(k)}\big|, \cdots, \big|a_{nm}^{(k)}\big|\big) \qquad (\diamond)$$

for some function f of nm variables satisfying that $f(\boldsymbol{y}) \to 0$ as $\boldsymbol{y} \to 0$. Therefore, the Sandwich Lemma implies that $\lim\limits_{k\to\infty} \|A_k\|_{p,q} = 0$. □

Problem 7. Let $\mathcal{M}_{n\times m}(\mathbb{F})$ be the collection of all $n \times m$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define $\|\cdot\|_F : \mathcal{M}_{n\times m}(\mathbb{F}) \to \mathbb{R}$ by

$$\|A\|_F \equiv \Big(\sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|^2\Big)^{\frac{1}{2}}.$$

1. Show that $\|A\|_F^2 = \mathrm{tr}(A^\dagger A)$, where $A^\dagger$ is the conjugate transpose of A, and $\mathrm{tr}(M)$ is the trace of square matrix M.

2. Show that $\|\cdot\|_F$ is a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$ (for all $n, m \in \mathbb{N}$). This norm is called the Frobenius norm of matrices.

3. Show that $\|AB\|_F \leqslant \|A\|_F\|B\|_F$ whenever $A \in \mathcal{M}_{n\times m}(\mathbb{F})$ and $B \in \mathcal{M}_{m\times p}(\mathbb{F})$.

Hint: 3. Let $A = \big[\boldsymbol{a}_1 \vdots \boldsymbol{a}_2 \vdots \cdots \vdots \boldsymbol{a}_m\big]$ and $B = \big[\boldsymbol{b}_1 \vdots \boldsymbol{b}_2 \vdots \cdots \vdots \boldsymbol{b}_m\big]^{\mathrm{T}}$; that is, $\boldsymbol{a}_k$ is the k-th column of A and $\boldsymbol{b}_\ell$ is the ℓ-th row of B. Then $AB = \sum\limits_{k=1}^{m} \boldsymbol{a}_k\boldsymbol{b}_k$. First show that $\|\boldsymbol{a}_k\boldsymbol{b}_k^{\mathrm{T}}\|_F = \|\boldsymbol{a}_k\|_2\|\boldsymbol{b}_k\|_2$ and use the triangle inequality to conclude the desired equality.

Proof. 1. Note that if $C = AB$ and $A = [a_{ij}]$, $B = [b_{ij}]$ and $C = [c_{ij}]$, then

$$c_{ij} = \sum_k a_{ik}b_{kj}\,. \qquad (0.1)$$

Therefore, if $B = A^\dagger A$, where $A = [a_{ij}] \in \mathcal{M}_{n\times m}(\mathbb{F})$ and $B = [b_{ij}] \in \mathcal{M}_{m\times m}(\mathbb{F})$, then the (i,k)-entry of $A^\dagger$ is $\overline{a_{ki}}$ so that

$$b_{ij} = \sum_{k=1}^{n} \overline{a_{ki}}a_{kj}\,;$$

thus

$$\mathrm{tr}(A^\dagger A) = \sum_{i=1}^{m} b_{ii} = \sum_{i=1}^{m}\sum_{k=1}^{n} \overline{a_{ki}}a_{ki} = \sum_{i=1}^{m}\sum_{k=1}^{n} |a_{ki}|^2 = \|A\|_F^2\,.$$

2. Clearly $\|\cdot\|_F$ satisfies properties (a)-(c) in Definition **??**, so it suffices to show the triangle inequality. Let $A = [a_{ij}]$ and $B = [b_{ij}]$. Define two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}^{nm}$ by

$$\boldsymbol{u} = (a_{11}, a_{12}, \cdots, a_{1m}, a_{21}, \cdots, a_{2m}, a_{31}, \cdots, a_{3m}, \cdots, a_{n1}, \cdots, a_{nm})$$

and

$$\boldsymbol{v} = (b_{11}, b_{12}, \cdots, b_{1m}, b_{21}, \cdots, b_{2m}, b_{31}, \cdots, b_{3m}, \cdots, b_{n1}, \cdots, b_{nm})\,.$$

Using the triangle inequality for the norm $\|\cdot\|_{\mathbb{F}^{nm}}$, we obtain that

$$\begin{aligned}\|A+B\|_F &= \Big(\sum_{i=1}^n\sum_{j=1}^m |a_{ij}+b_{ij}|^2\Big)^{\frac{1}{2}} = \|\boldsymbol{u}+\boldsymbol{v}\|_{\mathbb{F}^{nm}} \leqslant \|\boldsymbol{u}\|_{\mathbb{F}^{nm}} + \|\boldsymbol{v}\|_{\mathbb{F}^{nm}} \\ &= \Big(\sum_{i=1}^n\sum_{j=1}^m |a_{ij}|^2\Big)^{\frac{1}{2}} + \Big(\sum_{i=1}^n\sum_{j=1}^m |b_{ij}|^2\Big)^{\frac{1}{2}} = \|A\|_F + \|B\|_F\end{aligned}$$

so that the triangle inequality for $\|\cdot\|_F$ is established.

3. Let $\boldsymbol{a}_i$ and $\boldsymbol{b}_j$ denote the i-th column of A and j-th row of B, respectively. Then (0.1) implies that

$$AB = \boldsymbol{a}_1\boldsymbol{b}_1 + \boldsymbol{a}_2\boldsymbol{b}_2 + \cdots + \boldsymbol{a}_m\boldsymbol{b}_m\,. \tag{0.2}$$

Note that for column vector $\boldsymbol{a} = (a_1, \cdots, a_n)^{\mathrm{T}} \in \mathbb{F}^n$ and row vector $\boldsymbol{b} = (b_1, \cdots, b_p) \in \mathbb{F}^p$,

$$\|\boldsymbol{a}\boldsymbol{b}\|_F^2 = \sum_{i=1}^n\sum_{j=1}^p |a_ib_j|^2 = \Big(\sum_{i=1}^n |a_i|^2\Big)\Big(\sum_{j=1}^p |b_j|^2\Big) = \|\boldsymbol{a}\|_2^2\|\boldsymbol{b}\|_2^2\,;$$

thus (0.2) and the triangle inequality imply that

$$\|AB\|_F \leqslant \sum_{k=1}^m \|\boldsymbol{a}_k\boldsymbol{b}_k\|_F \leqslant \sum_{k=1}^m \|\boldsymbol{a}_k\|_2\|\boldsymbol{b}_k\|_2\,.$$

The Cauchy-Schwarz inequality further shows that

$$\|AB\|_F^2 \leqslant \Big(\sum_{k=1}^m \|\boldsymbol{a}_k\|_2\|\boldsymbol{b}_k\|_2\Big)^2 \leqslant \Big(\sum_{k=1}^m \|\boldsymbol{a}_k\|_2^2\Big)\Big(\sum_{k=1}^m \|\boldsymbol{b}_k\|_2^2\Big) = \|A\|_F^2\|B\|_F^2\,;$$

thus $\|AB\|_F \leqslant \|A\|_F\|B\|_F$. □

Problem 8. Let $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$. Show that

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle = 0 \qquad \text{if and only if} \qquad \|\boldsymbol{y}\| \leqslant \|\lambda\boldsymbol{x}+\boldsymbol{y}\| \quad \forall\,\lambda\in\mathbb{F}\,,$$

where $\|\cdot\|$ is the norm induced by the inner product $\langle\cdot,\cdot\rangle$.

Proof. Let $\alpha = \langle\boldsymbol{x},\boldsymbol{y}\rangle$. Then

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

"⇒" If $\alpha = 0$, then $\|\lambda\boldsymbol{x}+\boldsymbol{y}\|^2 = |\lambda|^2\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 \geqslant \|\boldsymbol{y}\|^2$ for all $\lambda\in\mathbb{F}$.

"⇐" W.L.O.G. we can assume that $\boldsymbol{x} \neq \boldsymbol{0}$. Letting $\lambda = -\dfrac{\bar{\alpha}}{\|\boldsymbol{x}\|^2}$, we find that

$$\|\boldsymbol{y}\|^2 \leqslant \|\lambda\boldsymbol{x}+\boldsymbol{y}\|^2 = \frac{|\alpha|^2}{\|\boldsymbol{x}\|^4}\|\boldsymbol{x}\|^2 + 2\frac{-|\alpha|^2}{\|\boldsymbol{x}\|^2} + \|\boldsymbol{y}\|^2 = \|\boldsymbol{y}\|^2 - \frac{|\alpha|^2}{\|\boldsymbol{x}\|^2}\,;$$

thus $\alpha = 0$. □

Problem 9. Let $(\mathcal{V}, +, \cdot, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{R}$, and define $\|\boldsymbol{v}\| = \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{1/2}$ for all $\boldsymbol{v} \in \mathcal{V}$. Show that

1. $2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2 = \|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2$ (parallelogram law).
2. $\big|\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big| \leqslant \|\boldsymbol{x}+\boldsymbol{y}\|\|\boldsymbol{x}-\boldsymbol{y}\| \leqslant \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.
3. $4\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2$ (polarization identity).

Can the p-norm $\|\cdot\|_p$ on $\mathbb{R}^n$ be induced from any inner product (on $\mathbb{R}^n$) for $p \neq 2$?

Proof. Note that if $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$, by Proposition 2.25 of the lecture note we have

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2 &= \langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 + \langle \boldsymbol{y}, \boldsymbol{x} \rangle + \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 , \\
\|\boldsymbol{x}-\boldsymbol{y}\|^2 &= \langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 - \langle \boldsymbol{y}, \boldsymbol{x} \rangle - \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 .
\end{aligned}$$

Since $\mathcal{V}$ is a vector space over $\mathbb{R}$, (e) of the definition of inner products implies that $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$; thus

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 \qquad \text{and} \qquad \|\boldsymbol{x}-\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 . \tag{0.3}$$

1. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$ be given. Then (0.3) implies that

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2\big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big) .$$

2. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$ be given. Then (0.3) implies that

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2\|\boldsymbol{x}-\boldsymbol{y}\|^2 &= \big(\|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2\big)\big(\|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2\big) \\
&= \big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big)^2 - 4\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big|^2 \leqslant \big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big)^2 ;
\end{aligned}$$

thus $\|\boldsymbol{x}+\boldsymbol{y}\|\|\boldsymbol{x}-\boldsymbol{y}\| \leqslant \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.

On the other hand, the Cauchy-Schwarz inequality implies that

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2\|\boldsymbol{x}-\boldsymbol{y}\|^2 &= \big(\|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2\big)\big(\|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2\big) \\
&= \big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big)^2 - 4\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big|^2 \geqslant \big(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\big)^2 - 4\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 \\
&= \|\boldsymbol{x}\|^4 + 2\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 + \|\boldsymbol{y}\|^4 - 4\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 \\
&= \|\boldsymbol{x}\|^4 - 2\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 + \|\boldsymbol{y}\|^4 = \big(\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big)^2 \geqslant 0 ;
\end{aligned}$$

thus $\|\boldsymbol{x}+\boldsymbol{y}\|\|\boldsymbol{x}-\boldsymbol{y}\| \geqslant \big|\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big|$.

3. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$ be given. Then (0.3) implies that

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2\langle \boldsymbol{y}, \boldsymbol{x} \rangle + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 4\langle \boldsymbol{x}, \boldsymbol{y} \rangle .$$

Suppose that $\|\cdot\|_p$ is induced by an inner production $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^n$. Then 1 implies that

$$2\|\boldsymbol{x}\|_p^2 + 2\|\boldsymbol{y}\|_p^2 = \|\boldsymbol{x}+\boldsymbol{y}\|_p^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_p^2 \qquad \forall\, \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n .$$

Let $\boldsymbol{x} = \mathbf{e}_1$ and $\boldsymbol{y} = \mathbf{e}_2$. Then $\|\boldsymbol{x}\|_p = \|\boldsymbol{y}\|_p = 1$ and $\|\boldsymbol{x}+\boldsymbol{y}\|_p = \|\boldsymbol{x}-\boldsymbol{y}\|_p = 2^{\frac{1}{p}}$ so that

$$4 = 2^{\frac{2}{p}} + 2^{\frac{2}{p}}$$

which holds only for $p = 2$. Therefore, if $p \neq 2$, then $\|\cdot\|_p$ is not induced by an inner product on $\mathbb{R}^n$. □

Problem 10. Let $(\mathcal{V},\langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{C}$. Show the polarization identity

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle = \frac{1}{4}\Big(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\|\boldsymbol{x}-i\boldsymbol{y}\|^2\Big) \qquad \forall\, \boldsymbol{x},\boldsymbol{y} \in \mathcal{V} .$$

Proof. Let $\boldsymbol{x},\boldsymbol{y} \in \mathcal{V}$ be given. Then

$$\begin{aligned}
&\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\|\boldsymbol{x}-i\boldsymbol{y}\|^2 \\
&\quad = \langle\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y}\rangle - \langle\boldsymbol{x}-\boldsymbol{y},\boldsymbol{x}-\boldsymbol{y}\rangle + i\langle\boldsymbol{x}+i\boldsymbol{y},\boldsymbol{x}+i\boldsymbol{y}\rangle - i\langle\boldsymbol{x}-i\boldsymbol{y},\boldsymbol{x}-i\boldsymbol{y}\rangle \\
&\quad = 2\big(\langle\boldsymbol{x},\boldsymbol{y}\rangle + \langle\boldsymbol{y},\boldsymbol{x}\rangle\big) + 2i\big(\langle\boldsymbol{x},i\boldsymbol{y}\rangle + \langle i\boldsymbol{y},\boldsymbol{x}\rangle\big) .
\end{aligned}$$

By Proposition 2.25 of the lecture note, we conclude that

$$i\big(\langle\boldsymbol{x},i\boldsymbol{y}\rangle + \langle i\boldsymbol{y},\boldsymbol{x}\rangle\big) = \langle\boldsymbol{x},\boldsymbol{y}\rangle - \langle\boldsymbol{y},\boldsymbol{x}\rangle ;$$

thus

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 - i\|\boldsymbol{x}+i\boldsymbol{y}\|^2 + i\|\boldsymbol{x}-i\boldsymbol{y}\|^2 = 4\langle\boldsymbol{x},\boldsymbol{y}\rangle .$$ □

Problem 11. Let (M,d) be a metric space. Define $\rho : M \times M \to \mathbb{R}$ by

$$\rho(x,y) = \frac{d(x,y)}{1+d(x,y)} .$$

Show that (M,ρ) is also a metric space.

Proof. By the fact that d is a metric, we find that $\rho(x,y) \geqslant 0$ and $\rho(x,y) = \rho(y,z)$ for all $x,y \in M$. Moreover,

$$\rho(x,y) = 0 \quad \text{if and only if} \quad d(x,y) = 0 \quad \text{if and only if} \quad x = y .$$

Therefore, if suffices to shows the triangle inequality. Let $x,y,z \in M$ be given. Then

$$\begin{aligned}
\big(1+d(x,z)\big)\big(\rho(x,y)+\rho(y,z)\big) &= \big(1+d(x,z)\big)\Big(\frac{d(x,y)}{1+d(x,y)} + \frac{d(y,z)}{1+d(y,z)}\Big) \\
&= \frac{d(x,y)\big(1+d(y,z)\big)\big(1+d(x,z)\big) + d(y,z)\big(1+d(x,y)\big)\big(1+d(x,z)\big)}{\big(1+d(x,y)\big)\big(1+d(y,z)\big)} \\
&\geqslant \frac{d(x,z) + d(x,y)d(x,z) + d(y,z)d(x,z) + d(x,y)d(y,z)d(x,z)}{1+d(x,y)+d(y,z)+d(x,y)d(y,z)} \\
&= d(x,z)\frac{1+d(x,y)+d(y,z)+d(x,y)d(y,z)}{1+d(x,y)+d(y,z)+d(x,y)d(y,z)} = d(x,z) ;
\end{aligned}$$

thus $\rho(x,y)+\rho(y,z) \geqslant \dfrac{d(x,z)}{1+d(x,z)} = \rho(x,z)$. □

Problem 12. Let $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$d(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} |x_1 - y_1| & \text{if } x_2 = y_2\,, \\ |x_1 - y_1| + |x_2 - y_2| + 1 & \text{if } x_2 \neq y_2\,, \end{cases} \quad \text{where } \boldsymbol{x} = (x_1, x_2) \text{ and } \boldsymbol{y} = (y_1, y_2).$$

Show that d is a metric on $\mathbb{R}^2$.

Proof. Let $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{y} = (y_1, y_2)$ and $\boldsymbol{z} = (z_1, z_2)$ in $\mathbb{R}^2$.

1. Clearly $d(\boldsymbol{x}, \boldsymbol{y}) \geqslant 0$.
2. $d(\boldsymbol{x}, \boldsymbol{y}) = 0 \Leftrightarrow (x_2 = y_2) \wedge |x_1 - y_1| = 0 \Leftrightarrow (x_2 = y_2) \wedge (x_1 = y_1) \Leftrightarrow \boldsymbol{x} = \boldsymbol{y}$.
3. (a) The case $x_2 = y_2$: In this case $d(\boldsymbol{x}, \boldsymbol{y}) = |x_1 - y_1|$ and $d(\boldsymbol{y}, \boldsymbol{x}) = |y_1 - x_1|$; thus if $x_2 = y_2$ then $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$.

   (b) The case $x_2 \neq y_2$: In this case

   $$d(\boldsymbol{x}, \boldsymbol{y}) = |x_1 - y_1| + |x_2 - y_2| + 1 \quad \text{and} \quad d(\boldsymbol{y}, \boldsymbol{x}) = |y_1 - x_1| + |y_2 - x_2| + 1\,;$$

   thus if $x_2 \neq y_2$ then $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$.

   In either cases, we have $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$.
4. (a) The case $x_2 = y_2$: In this case

   $$d(\boldsymbol{x}, \boldsymbol{y}) = |x_1 - y_1| \leqslant |x_1 - z_1| + |z_1 - y_1| \leqslant d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})\,.$$

   (b) The case $x_2 \neq y_2$: In this case z_2 is different from at least one of the second component x_2, y_2. W.L.O.G. we assume that $z_2 \neq x_2$. Then

   $$\begin{aligned} d(\boldsymbol{x}, \boldsymbol{y}) &= |x_1 - y_1| + |x_2 - y_2| + 1 \\ &\leqslant |x_1 - z_1| + |z_1 - y_1| + |x_2 - z_2| + |z_2 - y_2| + 1 \\ &= d(\boldsymbol{x}, \boldsymbol{z}) + |z_1 - y_1| + |z_2 - y_2| \leqslant d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})\,. \end{aligned}$$

   In either cases, d satisfies the triangle inequality. □