Rules:

- Unless otherwise stated, all answers must be mathematically justified.
- Partial answers will be graded.
- Your submission has to be uploaded to Gradescope. In Gradescope, indicate the page on which each problem is written.
- You can work in groups but each student must write his/her/their own solution based on his/her/their own understanding of the problem. Please list on your submission the students you work with for the homework (this will not affect your grade).
- Problems with a ($\star$) are extra credit, they will not (directly) contribute to your score of this homework. However, for every 4 extra credit questions successfully answered your lowest homework score get replaced by a perfect score.
- If you have any questions, feel free to ask them on Ed Discussion (so that everyone can benefit from the answer) or stop at the office hours.

Problem 4.1 (2.5 points). *Given $x \in \mathbb{R}^n$, define*

$$N(x) := \max\{|x_1|, |x_2|, ..., |x_n|\}.$$

(a) *Show that this defines a valid norm, i.e., verify that $N(\cdot)$ satisfies the three norm properties.*

1. *Triangle Inequality: $\forall x, y \in \mathbb{R}^n$, we have*

$$\begin{aligned} N(x+y) &= \max\{|x_1+y_1|, |x_2+y_2|, \cdots, |x_n+y_n|\} \\ &\leq \max\{|x_1|+|y_1|, |x_2|+|y_2|, \cdots, |x_n|+|y_n|\} \\ &= \max\{|x_1|, |x_2|, \cdots, |x_n|\} + \max\{|y_1|, |y_2|, \cdots, |y_n|\} \\ &= N(x) + N(y) \end{aligned}$$

2. *Homogenity: $\forall x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, we have*

$$\begin{aligned} N(\alpha x) &= \max\{|\alpha x_1|, |\alpha x_2|, \cdots, |\alpha x_n|\} \\ &= \max\{|\alpha||x_1|, |\alpha||x_2|, \cdots, |\alpha||x_n|\} \\ &= |\alpha| \max\{|x_1|, |x_2|, ..., |x_n|\} \\ &= |\alpha| N(x) \end{aligned}$$

3. *Positive Definiteness: $\forall x \in \mathbb{R}^n$, we have $N(x) \geq 0$ since $|x_i| \geq 0 \forall i$. When $N(x) = 0$ we have $\max\{|x_1|, |x_2|, ..., |x_n|\} = 0$ which means $|x_i| \leq |x_{max}| \leq 0$($|x_{max}| = \max\{|x_1|, |x_2|, ..., |x_n|\}$), which implies $|x_i| = 0$ and $x = 0$.*

(b) *Let $x = (10, 0)$ and $y = (9, 9)$. Using the norm $N(\cdot)$, which of x and y has bigger norm? Using the Euclidean norm, which of x and y has bigger norm?*

Using $N(\cdot)$, we have $N(x) = \max\{|10|, |0|\} = 10, N(y) = \max\{|9|, |9|\} = 9$, which implies that x has bigger L-infinity norm.
Using Euclidean norm, we have $\|x\| = 10$ and $\|y\| = 9\sqrt{2} > 10$, which implies y has bigger euclidean norm.

(c) *Show that $N(x) > 1$ implies the Euclidean norm of x is also greater than 1, i.e., $\|x\|_2 > 1$.*

$N(x) = \max\{|x_1|, |x_2|, \cdots, |x_n|\} > 1$, which implies that $|x_{max}| > 1$
(where $x_{max} = \max\{|x_1|, |x_2|, \cdots, |x_n|\}$).
Moreover, $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{\sum_{i \neq i_{max}} |x_i|^2 + x_{max}^2} \geq \sqrt{x_{max}^2} > 1$.

Problem 4.2 (2 points). *Decide whether each of the following functions $N(\cdot)$ is a valid norm. If yes, you do not have to justify your answer. If no, then provide an explicit counterexample and state which norm property is violated.*

(a) $x \in \mathbb{R}^2$. $N(x) := x^\top A x$, where $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$.

It is not a valid norm. And it violates the homogenity property, we check it: $\forall \alpha \in \mathbb{R}$ and $x \in \mathbb{R}^2$ we should have $N(\alpha x) = (\alpha x)^\top A(\alpha x) = \alpha^2 x^\top A x$ if $N(\cdot)$ is indeed a valid norm, but $\alpha^2 \neq \alpha$ in general, so this is not a valid norm.

(b) $x \in \mathbb{R}^2$. $N(x) := x^\top A x$, where $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$.

It is not a valid norm. And it violates the positive definiteness condition. $\exists \vec{x} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$ such that $\begin{bmatrix} -2 & -3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ -3 \end{bmatrix} = 8 - 9 = -1 < 0$.

(c) $x \in \mathbb{R}^3$. $N(x) := 2|x_1| + \max\{|x_2|, |x_3|\}$.

It is a valid norm and we check the three properties:

1. *Triangle Inequality: $\forall x, y \in \mathbb{R}^3$, we have*

$$\begin{aligned} N(x+y) &= 2|x_1 + y_1| + \max\{|x_2 + y_2|, |x_3 + y_3|\} \\ &\leq 2|x_1| + 2|y_1| + \max\{|x_2| + |y_2|, |x_3| + |y_3|\} \\ &= 2|x_1| + 2|y_1| + \max\{|x_2|, |x_3|\} + \max\{|y_2|, |y_3|\} \\ &= 2|x_1| + \max\{|x_2|, |x_3|\} + 2|y_1| + \max\{|y_2|, |y_3|\} \\ &= N(x) + N(y) \end{aligned}$$

2. *Homogenity: $\forall \alpha \in \mathbb{R}$ and $\forall x \in \mathbb{R}^3$, we have*

$$\begin{aligned} N(\alpha x) &= 2|\alpha x_1| + \max\{|\alpha x_2|, |\alpha x_3|\} \\ &= |\alpha| 2|x_1| + |\alpha| \max\{|x_2|, |x_3|\} \\ &= |\alpha| N(x) \end{aligned}$$

3. *Positive Definiteness: We know that $|x_i| \geq 0, \forall x_i$, so $N(x) \geq 0, \forall x \in \mathbb{R}^3$. When $N(x) = 0$, we know that $2|x_1| + \max\{|x_2|, |x_3|\} = 0$ and $|x_i| \geq 0$, which implies $x = 0$.*

. Thus we know that $N(x)$ satisfies all three properties.

(d) $x \in \mathbb{R}^n$. $N(x) := (\prod_{i=1}^n |x_i|)^{1/n} + \max_{i\in[n]}\{|x_i|\}$.

It is not a valid norm. And we can check it as follows:

$$\begin{aligned} N(\alpha x) &= \left(\prod_{i=1}^n |\alpha x_i| \right)^{\frac{1}{n}} + \max_{i\in[n]} |\alpha x_i| \\ &= |\alpha|^{\frac{1}{n}} \left(\prod_{i=1}^n |x_i| \right)^{\frac{1}{n}} + |\alpha| \cdot \max_{i\in[n]} |x_i| \end{aligned}$$

, which is not equal to $|\alpha| N(x) \forall \alpha$. Thus we can conclude that it is not a valid norm.

Problem 4.3 (1 point)**.** *Show that for any vector $x \in \mathbb{R}^n$,*

$$\left(\sum_{k=1}^{n} x_k\right)^2 \leq n \sum_{k=1}^{n} x_k^2.$$

Hint: use Cauchy-Schwarz.

Pick two vectors $\vec{1}_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^\top$ and $\vec{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ and we use the Cauchy-Schwarz Inequality to get $|\vec{1}_n^\top \vec{x}| \leq \|\vec{1}_n\|_2 \|\vec{x}\|_2$, squaring on both sides we can get exactly what we want:

$$\left(\sum_{k=1}^{n} x_k\right)^2 \leq n \sum_{k=1}^{n} x_k^2.$$

.

Problem 4.4 (1.5 points)**.** *Consider the set V of random variables (on a probability space Ω) with a finite second moment (i.e., for $X \in V$, $\mathbb{E}[X^2] < \infty$), with the inner product $\langle X, Y\rangle = \mathbb{E}[XY]$. Prove the Cauchy-Schwarz inequality:*

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}.$$

Hint: given constants $a, b \geq 0$, evaluate the expression $\mathbb{E}[(aX - bY)^2]$. Then find appropriate values of a, b to prove the statement.

We evaluate the expression in the hint as follows:

$$\begin{aligned}
\mathbb{E}[(aX - bY)^2] &= \mathbb{E}[a^2X^2 - 2abXY + b^2Y^2] \\
&= a^2\mathbb{E}[X^2] - 2ab\mathbb{E}[XY] + b^2\mathbb{E}[Y^2] \\
&= \mathbb{E}[X^2]a^2 - 2\mathbb{E}[XY]b \cdot a + \mathbb{E}[Y^2]b^2 \\
&\geq 0
\end{aligned}$$

This hold for all $a, b \in \mathbb{R}$, then we have $4\mathbb{E}^2[XY] \times b^2 - 4\mathbb{E}[X^2]\mathbb{E}[Y^2]b^2 \leq 0$. Since $b^2 \geq 0$, we have $\mathbb{E}^2[XY] \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$. If $\mathbb{E}[X^2] = 0$, then $\mathbb{E}[XY] = 0$ where the inequality still holds. So we take the square root on both sides to get $|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$.
Another perspective is to first choose the appropriate a and b, here we choose $a = |\mathbb{E}[XY]| \geq 0$ and $b = \mathbb{E}[X^2] \geq 0$ and fit them into $\mathbb{E}[(aX - bY)^2]$, using the lienarity of expectation we could get:

$$\begin{aligned}
\mathbb{E}[(|\mathbb{E}[XY]|X - \mathbb{E}[X^2]Y)^2] &= \mathbb{E}[\mathbb{E}^2[XY]X^2 - 2\mathbb{E}[XY]\mathbb{E}[X^2]XY + \mathbb{E}^2[X^2]Y^2]] \\
&= \mathbb{E}^2[X^2]\mathbb{E}[Y^2] - \mathbb{E}^2[XY]\mathbb{E}[X^2] \\
&\geq 0
\end{aligned}$$

and we could get $\mathbb{E}[X^2]\mathbb{E}[Y^2] \geq \mathbb{E}^2[XY]$, taking the square root we get the inequality of interest.

Problem 4.5 (3 points). *Let S be a subspace of $\mathbb{R}^n$ with $x \mapsto P_S(x)$ the orthogonal projector of all $x \in \mathbb{R}^n$. We will use the Euclidean dot product as our inner product $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$. Show that for any $x \in \mathbb{R}^n$,*

(a) $\langle x, y\rangle = \langle P_S(x), y\rangle$, *for any vector y in S.*

Suppose the orthonormal basis for subspace S is $(v_1, \cdots, v_k)$ and V if we organize them by columns. We know that the projection of x on to the subspace S is $P_S(x) = V(V^\top V)^{-1}V^\top \vec{x} = VV^\top \vec{x}$, then $\langle P_S(\vec{x}), \vec{y}\rangle = \langle VV^\top \vec{x}, \vec{y}\rangle = \vec{x}^\top VV^\top \vec{y}$. On the other hand, $VV^\top \vec{y}$ is the projection of $\vec{y}$ onto the subspace S, which is just $\vec{y}$. In other words, $VV^\top \vec{y} = \vec{y}$(Short proof here, since $\vec{y} \in S$, we have $\vec{y} = V\vec{\alpha}$ for some $\vec{\alpha} \in \mathbb{R}^k$, thus $VV^\top \vec{y} = VV^\top V\vec{\alpha} = V\vec{\alpha} = \vec{y}$). Thus $\vec{x}^\top \vec{y} = \langle P_S(\vec{x}), \vec{y}\rangle$.

(b) $x - P_S(x)$ *is orthogonal to S,*

We just need to prove that $\vec{x} - VV^\top \vec{x} \perp \vec{y}, \forall \vec{y} \in S$. We evaluate the inner product and could get $(\vec{x} - VV^\top \vec{x})^\top \vec{y} = \vec{x}^\top \vec{y} - VV^\top \vec{x}^\top \vec{y} = 0$ by part (a).

(c) $\|P_S(x)\| \le \|x\|$, *and the equality holds when $x \in S$.*

By the Pythagorean theorem we have $\|P_S(\vec{x})\|_2^2 + \|\vec{x} - P_S(\vec{x})\|_2^2 = \|\vec{x}\|_2^2$(since $\vec{x} - P_S(\vec{x}) \perp P_S(\vec{x})$). Then since $\|\vec{x} - P_S(\vec{x})\|_2^2 \ge 0$, we have$\|P_S(x)\|_2^2 \le \|x\|_2^2$, then taking the square root on both sides and we have $\|P_S(x)\| \le \|x\|$ as desired.

Hint: express y in terms of the orthonormal basis of S: $(v_1, v_2, ..., v_k)$.

Problem 4.6 ($\star$)**.** *Given $x \in \mathbb{R}^n$ and $p \in (1, \infty)$, define the ℓ_p-norm as*

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

Prove that the ℓ_p-norm satisfies the triangle inequality, that is, for any $x, y \in \mathbb{R}^n$, show that

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

You may use the following inequality, known as Hölder's inequality, without proof: for any $x, y \in \mathbb{R}^n$, and constants $p, q \in \mathbb{R}_+$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_p \|y\|_q.$$

We first apply holder's inequality as follows:

$$\begin{aligned}
\|x+y\|_p^p &= \sum_{i=1}^{n} |x_i + y_i|^p && \text{(Definition)} \\
&= \sum_{i=1}^{n} |x_i + y_i||x_i + y_i|^{p-1} && \text{(Arithmetic Operations)} \\
&\leq \sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1} && \text{(Triangle Inequality)} \\
&\leq (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}} (\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q})^{\frac{1}{q}} + (\sum_{i=1}^{n} |y_i|^p)^{\frac{1}{p}} (\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q})^{\frac{1}{q}} \\
& && \text{(Holder's Inequality)} \\
&= (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}} (\sum_{i=1}^{n} |x_i + y_i|^p)^{\frac{p-1}{p}} + (\sum_{i=1}^{n} |y_i|^p)^{\frac{1}{p}} (\sum_{i=1}^{n} |x_i + y_i|^p)^{\frac{p-1}{p}} \\
& && \text{(Utilize the condition for holder's inequality, 1/p+1/q=1)} \\
&= \|x\|_p \|x+y\|_p^{p-1} + \|y\|_p \|x+y\|_p^{p-1} && \text{(By definition)} \\
&= (\|x\|_p + \|y\|_p)\|x+y\|_p^{p-1} && \text{(Arithmetic Operations)}
\end{aligned}$$

, thus we have $\|x+y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x+y\|_p^{p-1}$, canceling on both sides we have $\|x+y\|_p \leq \|x\|_p + \|y\|_p$.
If $\|x+y\|_p = 0$ we have $x + y = \vec{0}$ by definition of the norm and in this case $\|x+y\|_p \leq \|x\|_p + \|y\|_p$ definitely holds.