IE 411: Introduction to Nonlinear Optimization

Fall 2022 - Homework Assignment 1
Solutions

Question 1. Show that $\|\cdot\|_p$ for $p = \frac{1}{2}$ which is given by

$$\|x\|_{\frac{1}{2}} := \left(\sum_{i=1}^{n} \sqrt{|x_i|}\right)^2$$

is not a norm. (**Hint:** It is sufficient to find a counterexample.)

Solution

Let $n = 2$, with $x = (1,4)^\mathsf{T}$ and $y = (1,9)^\mathsf{T}$. We have $\|x\|_{\frac{1}{2}} = 9$, $\|y\|_{\frac{1}{2}} = 16$ and $\|x+y\|_{\frac{1}{2}} = 15 + 2\sqrt{26}$. We have $\|x+y\|_{\frac{1}{2}} > \|x\|_{\frac{1}{2}} + \|y\|_{\frac{1}{2}}$ violating the triangle inequality.

Question 2. In this question, you will prove the following statement step by step.

"Let $\|\cdot\|$ be the Euclidean (ℓ_2) norm. For all $x, y \in \mathbb{R}^n$, we have

$$|x^\mathsf{T} y| \leq \|x\| \, \|y\| \, . \tag{1}$$

Moreover, the equality holds if and only if $x = ky$ for some $k \in \mathbb{R}$."

The inequality given by (1) is called the **Cauchy-Schwarz inequality.**

a) Show the statement for $x = 0 \in \mathbb{R}^n$.

For the remaining parts, assume that $x \neq 0$.

b) Show that the following equality holds for all $x, y \in \mathbb{R}^n$, $x \neq 0$:

$$\frac{1}{\|x\|^2} \left\| \|x\|^2 .y - (x^\mathsf{T} y).x \right\|^2 = \|x\|^2 \|y\|^2 - |x^\mathsf{T} y|^2 \tag{2}$$

c) Using equality (2), show that inequality (1) holds.

d) Assume that $\|x\|\,\|y\| = |x^\mathsf{T} y|$ holds. Using equality (2), show that $y = k.x$ for some $k \in \mathbb{R}$. (Write the value of k in terms of x, y.)

Solution

a) Let $x = 0 \in \mathbb{R}^n$. For any $y \in \mathbb{R}^n$, we have $x^\mathsf{T} y = 0$ and $|x^\mathsf{T} y| = 0$. Using the non-negativity property of a norm we have $\|x\| = 0$, so $\|x\|\,\|y\| = 0$. Hence $|x^\mathsf{T} y| \leq \|x\|\,\|y\|$.

b) Let $x \neq 0$. Then,

$$\begin{aligned}
\frac{1}{\|x\|^2}\left\|\|x\|^2 y - (x^\mathsf{T} y)x\right\|^2 &= \left(\frac{\left\|\|x\|^2 y - (x^\mathsf{T} y)x\right\|}{\|x\|}\right)^2 \\
&= \left(\left\|\|x\|\, y - (x^\mathsf{T} y)\frac{x}{\|x\|}\right\|\right)^2 && \text{(pos. hom.)} \\
&= \left(\|x\|\, y - (x^\mathsf{T} y)\frac{x}{\|x\|}\right)^\mathsf{T}\left(\|x\|\, y - (x^\mathsf{T} y)\frac{x}{\|x\|}\right) \\
&= \|x\|^2\, y^\mathsf{T} y - 2(x^\mathsf{T} y)^2 + (x^\mathsf{T} y)^2 \frac{x^\mathsf{T} x}{\|x\|^2} \\
&= \|x\|^2\,\|y\|^2 - (x^\mathsf{T} y)^2.
\end{aligned}$$

c) Let $x \neq 0$. Then

$$\begin{aligned}
& \frac{1}{\|x\|^2}\left\|\|x\|^2 y - (x^\mathsf{T} y)x\right\|^2 \geq 0 && \text{(non-negativity)} \\
\Rightarrow\quad & \frac{1}{\|x\|^2}\left\|\|x\|^2 y - (x^\mathsf{T} y)x\right\|^2 + (x^\mathsf{T} y)^2 \geq (x^\mathsf{T} y)^2 \\
\Rightarrow\quad & \|x\|^2\,\|y\|^2 \geq (x^T y)^2 && \text{(Previous problem.)} \\
\Rightarrow\quad & \sqrt{\|x\|^2\,\|y\|^2} \geq \sqrt{(x^\mathsf{T} y)^2} && \text{(Square root is monotonic.)} \\
\Rightarrow\quad & \|x\|^2\,\|y\|^2 \geq |x^\mathsf{T} y|.
\end{aligned}$$

d) If $x = 0 \in \mathbb{R}^n$, then $\|x\|\,\|y\| = |x^\mathsf{T} y|$ holds and we can say that $y = kx$

for $k = 0$. Let $x \neq 0$. If $\|x\| \|y\| = |x^\top y|$ holds then we have

$$\begin{aligned}
&\frac{1}{\|x\|^2} \left\| \|x\|^2 y - (x^\top y)x \right\|^2 = 0 && \text{(Previous problem.)} \\
\Rightarrow \quad & \left\| \|x\|^2 y - (x^\top y)x \right\|^2 = 0 && \text{(non-negativity)} \\
\Rightarrow \quad & \|x\|^2 y - (x^\top y)x = 0 && \text{(non-negativity)} \\
\Rightarrow \quad & y = \frac{x^\top y}{\|x\|^2} x.
\end{aligned}$$

We have shown that under the given assumptions y and x are linearly dependent with $k = \frac{x^\top y}{\|x\|^2}$.

Question 3. Let $T \in \mathbb{R}^{2\times 2}$ be a linear operator defined such that for any $x = (x_1, x_2)^\top \in \mathbb{R}^2$, we have $Tx = (x_2, x_1)^\top$. Find all eigenvalues and eigenvectors of T.

Solution

The linear operator T defined in the problem can be found as $T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. We have to check pairs (ν, λ) satisfying $T\nu = \lambda\nu$. This equation yields the following system, $\nu_2 = \lambda\nu_1$ and $\nu_1 = \lambda\nu_2$. This implies $\nu_2 = \lambda^2\nu_2$. We have three possible cases,

- $\nu_2 = 0$, this will force $\nu_1 = 0$ which is not possible for an eigenvector.
- $\lambda = 1$, with the eigenvector $\nu = (1, 1)^\top$.
- $\lambda = -1$, with the eigenvector $\nu = (-1, 1)^\top$.

Since we found two eigenvalues, we are done with the search. Eigenvalues are $1, -1$ with eigenvectors $(1, 1)^\top, (-1, 1)^\top$ respectively.

Question 4. Let $A \in \mathbb{R}^{n\times n}$ be the matrix of all 1's, that is,

$$A = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

Find all the eigenvalues of A. (**Hint:** Instead of writing the characteristic equation, you may write the definition of eigenvalue and eigenvector to compute the eigenvalues. You may consider $n = 2, n = 3$ cases separately to see a pattern.)

Solution

Definition of the eigenvalues and eigenvectors gives us $A\nu = \lambda\nu$ for each pair (ν, λ). Using the given matrix A, we see that $\sum_{i=1}^{n} \nu_i = \lambda\nu_j$ for any $j = 1, \dots, n$. This implies, $n \sum_{i=1}^{n} \nu_i = \lambda \sum_{i=1}^{n} \nu_i$ and this equality yields two possible options that are

- $\lambda = n$, with the eigenvector $\nu = (1, \dots, 1)^\mathsf{T}$ since we have $\sum_{i=1}^{n} \nu_i = n\nu_j$ for any $j = 1, \dots, n$.
- $\lambda \neq n$, then we should have $\sum_{i=1}^{n} \nu_i = 0$. Since we cannot have $\nu_j = 0$ for all j, we have $\lambda = 0$ in this case.

Since we checked all possible cases and observed that eigenvalue $\lambda = 0$ has geometric multiplicity 1, we can say that eigenvalues of matrix A are $(n, \underbrace{0, \dots, 0}_{n-1})$.

Question 5. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined as $f(x, y) = x^2 + y^2 + 2x - 3y$. Find a global minimum point of f over the the unit ball $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

Solution

Let $a = \begin{bmatrix} x \\ y \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$. Then for any $(x, y) \in S$ we have $0 \leq \|a\| \leq 1$ and

$$f(x, y) = a^\mathsf{T} a + a^\mathsf{T} b \geq \|a\|^2 - \|a\| \, \|b\| \geq \|a\|^2 - \sqrt{13} \, \|a\| \, ,$$

where we use the Cauchy-Schwarz inequality.

To find a lower bound we minimize $\|a\|^2 - \sqrt{13} \, \|a\|$ over the region $0 \leq \|a\| \leq 1$ and obtain $1 - \sqrt{13}$ as a lower bound attained at $\|a\| = 1$. In order to find these solution we used the fact that $\|a\|^2 - \sqrt{13} \, \|a\|$ is a parabola and the minimum should be at one of the end points $\{0, 1\}$ or the parabola vertex $\frac{\sqrt{13}}{2}$.

Now we know that $f(x,y) \geq 1-\sqrt{13}$. If we pick the pair $(\hat{x},\hat{y}) = \left(\frac{-2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\right)$, we find a global minimizer since $f(\hat{x},\hat{y}) = 1-\sqrt{13}$. This pair can be found by solving the system $\begin{cases} 2x-3y=-\sqrt{13} \\ x^2+y^2=1 \end{cases}$, since we have found the lower bound at $\|a\|=1$. A candidate local optimum can be also found by using optimality conditions or inspection.

F.O.C. Approach

We check the gradient of f, that is $\nabla f(x,y) = \begin{bmatrix} 2x+2 \\ 2y-3 \end{bmatrix}$. This function has a zero at the pair $(\hat{x},\hat{y}) = (-1, \frac{3}{2})$. Unfortunately, this stationary point is out of the feasible region, so the optimal solution should be on the boundary. In that case we have $x^2+y^2=1$ and we may switch to polar coordinates using $x=\cos\theta$ and $y=\sin\theta$. In that case we solve the equivalent problem minimizing $\tilde{f}(\theta) = 1+2\cos\theta - 3\sin\theta$ over $[0,2\pi]$. Now, we compute the derivative $\frac{d\tilde{f}(\theta)}{d\theta} = -2\sin\theta - 3\cos\theta$ and it has stationary points of the form $\theta_1 = \pi - \tan^{-1}(\frac{3}{2})$ and $\theta_2 = 2\pi - \tan^{-1}(\frac{3}{2})$. Then we switch back to the Cartesian coordinate system to obtain pairs,

$$(x_1,y_1) = \left(\frac{-2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\right), \quad (x_2,y_2) = \left(\frac{2}{\sqrt{13}}, \frac{-3}{\sqrt{13}}\right)$$

with objective values $f(x_1,y_1) = 1-\sqrt{13}$ and $f(x_2,y_2) = 1+\sqrt{13}$. So we found the same global solution as above.

Question 6. Find the global minimum and maximum points of the function $f(x,y) = 2x-3y$ over the set $S = \{(x,y) : 2x^2+5y^2 \leq 1\}$.

Solution

Let $a=\sqrt{2}x$ and $b=\sqrt{5}y$. Then $S := \{(a,b) : a^2+b^2 \leq 1\}$ and $f(a,b) = \sqrt{2}a - \frac{3}{\sqrt{5}}b$. Using Cauchy-Schwarz inequality we have $-\frac{19}{5} \leq f(a,b) \leq \frac{19}{5}$. By solving the system $\begin{cases} \sqrt{2}a - \frac{3}{\sqrt{5}}b = 19/5 \\ a^2+b^2=1 \end{cases}$ we can find the pair $(\overline{a},\overline{b}) = \left(\sqrt{\frac{10}{19}}, -\sqrt{\frac{9}{19}}\right)$ to be a global maximizer. Using the linearity of the $f(a,b)$ with a similar argument we obtain $(\underline{a},\underline{b}) = \left(-\sqrt{\frac{10}{19}}, \sqrt{\frac{9}{19}}\right)$ to be a global minimizer. Necessary linear transformations can be done in order to find a global minimum $(\underline{x}, \underline{y})$ and a global maximum $(\overline{x}, \overline{y})$.

Question 7. For each of the following functions, determine whether it is coercive or not:

a. $f(x_1, x_2) = 2x_1^2 - 8x_1x_2 + x_2^2$.

b. $f(x_1, x_2) = 4x_1^2 + 2x_1x_2 + 2x_2^2$.

c. $f(x_1, x_2) = x_1^4 + x_2^4$.

d. $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$.

Solution

a. $f(x_1, x_2) = 2x_1^2 - 8x_1x_2 + x_2^2$.
Not coercive. Counter example:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a \\ a \end{pmatrix} \in \mathbb{R}^2$$

Clearly as $a \to \infty$, $\|x\| \to \infty$, and we have

$$\lim_{a\to\infty} f(a, a) = \lim_{a\to\infty} 2a^2 - 8a^2 + a^2 \to -\infty.$$

b. $f(x_1, x_2) = 4x_1^2 + 2x_1x_2 + 2x_2^2$.
Coercive. We have

$$\begin{aligned} f(x_1, x_2) &= 4x_1^2 + 2x_1x_2 + 2x_2^2 \\ &= 3x_1^2 + x_2^2 + (x_1 + x_2)^2 \\ &\geq x_1^2 + x_2^2 = \|x\|^2. \end{aligned}$$

As $\|x\| \to +\infty$, $\|x\|^2 \to +\infty$, and forces $f(x_1, x_2)$ to move toward $+\infty$. Therefore, $f(x_1, x_2)$ is coercive.

c. $f(x_1, x_2) = x_1^4 + x_2^4$.
Coercive. We have Arithmetic Geometric Mean Inequality $\Rightarrow x + y \geq$

$2\sqrt{xy}$

$$\begin{aligned} f(x_1, x_2) &= x_1^4 + x_2^4 \\ &\geq 2x_1x_2 = ((x_1 + x_2)^2) - x_1^4 - x_2^4 \end{aligned}$$

Then we modify the inequalies and we have

$$\begin{aligned} f(x_1, x_2) &= x_1^4 + x_2^4 \\ &\geq \frac{(x_1 + x_2)^2}{2} = \|x\|^4 \end{aligned}$$

As $\|x\| \to +\infty$, $\|x\|^4 \to +\infty$, and forces $f(x_1, x_2)$ to move toward $+\infty$. Therefore, $f(x_1, x_2)$ is coercive.

d. $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$.
Not coercive. Counter example:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -a \end{pmatrix}$$

Let $a \to \infty$, then $\|(0, 0, -a)\| \to \infty$

$$\lim_{a\to\infty} f(0, 0, -a) = \lim_{a\to\infty} (-a^3) \to -\infty.$$