

# SI231 - Matrix Computations, Fall 2020-21

## Solution of Homework Set #1

Prof. Yue Qiu and Prof. Ziping Zhao

### I. UNDERSTANDING RANK, RANGE SPACE AND NULL SPACE

**Problem 1.** (4 points  $\times$  5) This problem is graded by Zhicheng Wang (wangzhch1@).

- 1) For matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , prove that  $\mathbb{R}^n = \mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A}^T)$ <sup>1</sup>.

**Hint:**  $\dim(\mathcal{N}(\mathbf{A})) + \dim(\mathcal{R}(\mathbf{A}^T)) = n$ .

- 2) For matrices  $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{m \times n}$ , prove that  $\text{rank}(\mathbf{A} + \mathbf{B}) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$ .

- 3) For matrices  $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times p}$ , prove that  $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$  and  $\text{rank}(\mathbf{AB}) = n$  only when  $\mathbf{A}$  has full-column rank and  $\mathbf{B}$  has full-row rank.

- 4) For matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\mathbf{B} \in \mathbb{R}^{m \times p}$ , prove that  $\mathcal{R}(\mathbf{A}|\mathbf{B}) = \mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})$ <sup>2</sup>

- 5) For matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\mathbf{B} \in \mathbb{R}^{m \times p}$ , prove that

$$\text{rank}(\mathbf{A}|\mathbf{B}) = \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) - \dim(\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})).$$

**Hint:** Recall the result in (c).

**Solution:**

- 1) For any  $\mathbf{x} \in \mathcal{R}(\mathbf{A}^T) \cap \mathcal{N}(\mathbf{A})$ , we have that  $\mathbf{Ax} = \mathbf{0}$  and  $\exists \mathbf{c} \in \mathbb{R}^m$ , s.t.  $\mathbf{A}^T \mathbf{c} = \mathbf{x}$ . Therefore we have  $\mathbf{AA}^T \mathbf{c} = \mathbf{0}$  and consequently,

$$(\mathbf{A}^T \mathbf{c})^T \cdot (\mathbf{A}^T \mathbf{c}) = \mathbf{c}^T \mathbf{AA}^T \mathbf{c} = \mathbf{0},$$

i.e.,  $\mathcal{R}(\mathbf{A}^T) \cap \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$ <sup>2'</sup>. Let  $\mathcal{S} = \mathcal{R}(\mathbf{A}^T) \oplus \mathcal{N}(\mathbf{A})$ , since  $\dim(\mathcal{S}) = \dim(\mathcal{N}(\mathbf{A})) + \dim(\mathcal{R}(\mathbf{A}^T)) - \dim(\mathcal{R}(\mathbf{A}^T) \cap \mathcal{N}(\mathbf{A})) = n - 0 = n$ , we must have  $\mathcal{S} = \mathbb{R}^n$ . To sum up,  $\mathbb{R}^n = \mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A}^T)$ <sup>2'</sup>.

- 2)

$$\begin{aligned} \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) &= \text{rank} \left( \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \right) = \text{rank} \left( \begin{bmatrix} \mathbf{A} & \mathbf{A} + \mathbf{B} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \right) \\ &\geq \text{rank} \left( \begin{bmatrix} \mathbf{A} + \mathbf{B} \\ \mathbf{B} \end{bmatrix} \right) \geq \text{rank}(\mathbf{A} + \mathbf{B}) \text{ 4'}. \end{aligned}$$

$\dim(\mathbf{A})$  is wrong.  $\mathcal{R}(\mathbf{A} + \mathbf{B}) \subset \mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})$

- 3) Let  $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p], \mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p]$ .

- Let  $\mathbf{AB} = [\mathbf{Ab}_1, \mathbf{Ab}_2, \dots, \mathbf{Ab}_p]$ .  $\forall j \in [p]$ ,  $\mathbf{Ab}_j$  is a linear combination of  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ , so  $\mathbf{Ab}_j \in \text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ , i.e.,  $\{\mathbf{Ab}_1, \mathbf{Ab}_2, \dots, \mathbf{Ab}_p\} \subset \text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ ,  $\text{span}(\mathbf{Ab}_1, \mathbf{Ab}_2, \dots, \mathbf{Ab}_p) \subset \text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ , we then get  $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A})$ <sup>1'</sup>.

<sup>1</sup>Let  $\mathcal{S}_1$  and  $\mathcal{S}_2$  be two subspaces of  $\mathbb{R}^n$ , if  $\mathcal{S}_1 \cap \mathcal{S}_2 = \{\mathbf{0}\}$  and  $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^n$ , we define the **direct sum**  $\mathbb{R}^n = \mathcal{S}_1 \oplus \mathcal{S}_2$ .

<sup>2</sup>Here  $\mathbf{A}|\mathbf{B}$  denotes a new matrix combined by  $\mathbf{A}$  and  $\mathbf{B}$ . For example,  $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ ,  $\mathbf{B} = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}$ , then  $\mathbf{A}|\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & b_{11} \\ a_{21} & a_{22} & b_{21} \end{bmatrix}$ .

- Since  $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B}^T \mathbf{A}^T)$ , applying the same strategy, we can get

$$\text{rank}(\mathbf{B}^T \mathbf{A}^T) \leq \text{rank}(\mathbf{B}^T) = \text{rank}(\mathbf{B}) \quad \textcolor{red}{1'}.$$

Therefore,  $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\} \quad \textcolor{red}{1'}$ .

Moreover, if  $\mathbf{A}$  has full-column rank,  $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B})$ , and if  $\mathbf{B}$  has full-row rank,  $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{A})$ , then the second claim is proved 1'.

$$4) \quad \forall \mathbf{y} \in \mathcal{R}(\mathbf{A}|\mathbf{B}), \exists \mathbf{x} \in \mathbb{R}^{n+p} \text{ such that } \mathbf{y} = (\mathbf{A}|\mathbf{B})\mathbf{x} = (\mathbf{A}|\mathbf{B}) \begin{bmatrix} \mathbf{x}_n \\ \mathbf{x}_p \end{bmatrix} = \mathbf{A}\mathbf{x}_n + \mathbf{B}\mathbf{x}_p, \quad \mathbf{x}_n \in \mathbb{R}^n \text{ and } \mathbf{x}_p \in \mathbb{R}^p.$$

We have  $\mathbf{A}\mathbf{x}_n \in \mathcal{R}(\mathbf{A})$  and  $\mathbf{B}\mathbf{x}_p \in \mathcal{R}(\mathbf{B})$ ,  $\mathcal{R}(\mathbf{A}|\mathbf{B}) \subset (\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})) \quad \textcolor{red}{2'}$ .

$$\forall \mathbf{y}_n \in \mathcal{R}(\mathbf{A}), \mathbf{y}_p \in \mathcal{R}(\mathbf{B}), \exists \mathbf{x}_n \in \mathbb{R}^n \text{ and } \mathbf{x}_p \in \mathbb{R}^p \text{ such that } \mathbf{y}_n = \mathbf{A}\mathbf{x}_n \text{ and } \mathbf{y}_p = \mathbf{B}\mathbf{x}_p. \text{ We have}$$

$$\mathbf{y}_n + \mathbf{y}_p = \mathbf{A}\mathbf{x}_n + \mathbf{B}\mathbf{x}_p = (\mathbf{A}|\mathbf{B}) \begin{bmatrix} \mathbf{x}_n \\ \mathbf{x}_p \end{bmatrix}, \quad (\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})) \subset \mathcal{R}(\mathbf{A}|\mathbf{B}) \quad \textcolor{red}{2'}.$$

Therefore,  $(\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})) = \mathcal{R}(\mathbf{A}|\mathbf{B})$ .

5)

$$\begin{aligned} \text{rank}(\mathbf{A}|\mathbf{B}) &= \dim(\mathcal{R}(\mathbf{A}|\mathbf{B})) = \dim(\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{B})) \\ &= \dim(\mathcal{R}(\mathbf{A})) + \dim(\mathcal{R}(\mathbf{B})) - \dim(\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})) \\ &= \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) - \dim(\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})) \quad \textcolor{red}{4'}. \end{aligned}$$

## II. UNDERSTANDING SPAN, SUBSPACE

**Problem 1.** (10 points) This problem is graded by **Lin Zhu (zhulin@)**.

For a set of vectors  $\mathcal{S} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ , prove that  $\text{span}(\mathcal{S})$  is the intersection of all subspaces that contain  $\mathcal{S}$ , i.e., prove that  $\text{span}(\mathcal{S}) = \mathcal{M}$  where  $\mathcal{M} := \cap_{\mathcal{S} \subseteq \mathcal{V}} \mathcal{V}$  is the intersection of all subspaces that contain  $\mathcal{S}$  and  $\mathcal{V}$  denotes the subspace containing  $\mathcal{S}$ .

**Hint:** Prove that  $\text{span}(\mathcal{S}) \subseteq \mathcal{M}$  and  $\mathcal{M} \subseteq \text{span}(\mathcal{S})$ .

**Solution:** The proof consists of two parts:

- (5 points) First we prove that  $\text{span}(\mathcal{S}) \subset \mathcal{M}$ . For any  $\mathbf{x} \in \text{span}(\mathcal{S})$ , then  $\mathbf{x}$  can be linearly represented by  $\mathcal{S}$ , i.e.,  $\mathbf{x} = \sum_i \alpha_i \mathbf{v}_i$ . For any subspace  $\mathcal{V}$  containing  $\mathcal{S}$ , we must have  $\mathbf{x} \in \mathcal{V}$  since subspace  $\mathcal{V}$  is closed under addition. Therefore, we have  $\mathbf{x} \in \cap_{\mathcal{S} \subseteq \mathcal{V}} \mathcal{V} = \mathcal{M}$ . To sum up,  $\forall \mathbf{x} \in \text{span}(\mathcal{S}) \Rightarrow \mathbf{x} \in \mathcal{M} \Rightarrow \text{span}(\mathcal{S}) \subset \mathcal{M}$ .
- (5 points) Then we try to prove that  $\mathcal{M} \subset \text{span}(\mathcal{S})$ . By definition,  $\mathcal{M}$  is contained in every subspace which contains  $\mathcal{S}$ . (The intersection of subspaces is also a subspace.) And since  $\text{span}(\mathcal{S})$  is also a subspace, then we have  $\mathcal{M} \subset \text{span}(\mathcal{S})$ .

Therefore  $\text{span}(\mathcal{S}) = \mathcal{M}$ .

### III. BASIS, DIMENSION AND PROJECTION

**Problem 1.** (2 points  $\times$  2) This problem is graded by **Zhihang Xu (xuzhh@)**.

Determine the dimension of each of the following vector spaces:

- 1) The space of polynomials having degree  $n$  or less;
- 2) The space of  $n \times n$  symmetric matrices.

**Solution.**

- 1) (2 points) Let  $\mathcal{V}$  be the space of polynomials having degree  $n$  or less, then any element  $f(x) \in \mathcal{V}$  can be expressed as

$$f(x) = a_0 + a_1x + \cdots + a_nx^n,$$

which can be a linear combination of  $\{1, x, \dots, x^n\}$ . Therefore the basis of  $\mathcal{V}$  can be  $\{1, x, \dots, x^n\}$  and the dimension of  $\mathcal{V}$  is  $n + 1$ .

- 2) (2 points) Let  $\mathcal{V}$  be the space of  $n \times n$  symmetric matrices and let  $\mathbf{I}^{(ij)}$  denotes the symmetric matrix which has value 1 on  $\mathbf{I}_{ij}^{(ij)}$  and  $\mathbf{I}_{ji}^{(ij)}$  and value 0 on elsewhere, then any element  $\mathbf{A} \in \mathcal{V}$  can be expressed as

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} = a_{11} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + \cdots \\ &\quad + a_{1n} \begin{bmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} + \cdots + a_{nn} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \sum_{m=1}^n a_{mm} \mathbf{I}^{(mm)} + \sum_{m=2}^n a_{1m} \mathbf{I}^{(1m)} + \sum_{m=3}^n a_{2m} \mathbf{I}^{(2m)} + \cdots + a_{n-1,n} \mathbf{I}^{(n-1,n)}, \end{aligned}$$

and therefore the dimension of  $\mathcal{V}$  is given by

$$\dim(\mathcal{V}) = n + (n-1) + \cdots + 1 = \frac{n(n+1)}{2}.$$

**Remarks:**

- Deviation for the solution is not required, giving the solutions  $n + 1$  and  $n(n + 1)/2$  directly earns you full points.
- If the final solution is not right, but you give a right deviation, you will get half of the points.

## Problem 2. Some Important linear transformations

(6 points + 8 points) This problem is graded by **Zhihang Xu (xuzhh@)**.

1) **Rotations.** A rotation matrix  $\mathbf{R} \in \mathbb{R}^{n \times n}$  is an orthogonal matrix ( $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ ) such that  $\det(\mathbf{R}) = 1$ .

- According to the above definition, find all rotation matrices in  $\mathbb{R}^{2 \times 2}$ .
- Geometrically, if  $\mathbf{R} \in \mathbb{R}^{2 \times 2}$ , then  $\mathbf{R}\mathbf{x}$  means we rotate the vector  $\mathbf{x} \in \mathbb{R}^2$  from some angle  $\theta \in [0, 2\pi]$  in anti-clockwise direction. For  $\mathbf{x} = [\cos(\pi/4), \sin(\pi/4)]^T$ , compute  $\mathbf{R}\mathbf{x}$ , where  $\mathbf{R}$  represents the matrix that rotating  $\mathbf{x}$  by  $7/12\pi$  in anti-clockwise direction.

**Hint:** draw a plot of  $\mathbf{x}$  and  $\mathbf{R}\mathbf{x}$ .

2) **Reflections.** Let  $\mathbf{u} \in \mathbb{R}^n$  be a unit vector,  $\|\mathbf{u}\|_2 = 1$ . For a given vector  $\mathbf{x} \in \mathbb{R}^n$  and a hyperplane  $\mathcal{H}_u = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{u}^T \mathbf{x} = 0\}$ . Let  $\mathbf{Q} = \mathbf{I} - \mathbf{u}\mathbf{u}^T$ . Then a vector  $\mathbf{y} \in \mathbb{R}^n$  is said to be a *reflection* of  $\mathbf{x}$  with respect to  $\mathcal{H}$  if their projections onto the hyperplane  $\mathcal{H}$  (denoted as  $\mathbf{Q}\mathbf{x}$  and  $\mathbf{Q}\mathbf{y}$  respectively) satisfy

$$\mathbf{Q}\mathbf{x} = \mathbf{Q}\mathbf{y}, \quad \|\mathbf{x} - \mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{y} - \mathbf{Q}\mathbf{y}\|_2.$$

See Figure.1 for visualization.

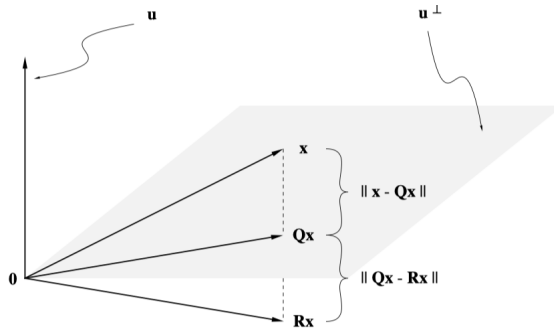


Figure 1. Reflection of  $\mathbf{x}$

A Householder matrix has the form  $\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$ . Prove that  $\mathbf{H}\mathbf{x}$  is a reflection of  $\mathbf{x}$  with respect to  $\mathcal{H}_u$ .

**Solution.**

1) (3 points)

**Solution #1.** In a two-dimensional plane, consider a vector  $\mathbf{x} = [x_1, x_2]^T$ , what a rotation matrix  $\mathbf{R}$  does to  $\mathbf{x}$  is to rotate  $\mathbf{x}$  by some angle  $\alpha$ . Suppose  $\mathbf{x} = (r, \theta)$  in polar coordinate system, and  $\mathbf{x}' := \mathbf{R}\mathbf{x} = [x'_1, x'_2]$  is the vector rotated by  $\alpha$  in anti-clockwise direction. Then by the definition of rotation, we have

$$\begin{cases} x'_1 = r \cos(\theta + \alpha), \\ x'_2 = r \sin(\theta + \alpha), \end{cases} \Rightarrow \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

and therefore the rotation matrix in anti-clockwise direction is given by

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}.$$

Similarly, the rotation matrix in clockwise direction can be derived in a similar manner,

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}.$$

- **Solution #2.** Let  $\mathbf{R} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , then by the definition of the rotation matrix, we have that

$$\begin{cases} \mathbf{R}\mathbf{R}^T = \mathbf{I} \\ \det(\mathbf{R}) = 1, \end{cases} \Rightarrow \begin{cases} ad - cb = 1, \\ a^2 + b^2 = 1, \\ c^2 + d^2 = 1, \\ ac + bd = 0. \end{cases} \quad (1)$$

Therefore, we can first assume that  $a = \cos \alpha$ ,  $b = \sin \alpha$  and  $c = \cos \beta$ ,  $d = \sin \beta$  for some  $\alpha$  and  $\beta$ .

Substituting  $a, b, c$ , and  $d$  into (1) gives

$$ad - bc = \sin(\beta - \alpha) = 1,$$

$$ac + bd = \cos(\beta - \alpha) = 0,$$

which implies  $\beta = \alpha + 2k\pi + \pi/2$ , where  $k$  is an integer, thus,

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}, \quad (2)$$

also, assume that  $a = \cos \alpha$ ,  $b = -\sin \alpha$  and  $c = \cos \beta$ ,  $d = \sin \beta$  for some  $\alpha$  and  $\beta$ , then analogously,  $\mathbf{R}$  is given by

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}. \quad (3)$$

And (2) and (3) represent the rotation matrix in clockwise direction and anti-clockwise direction respectively.

(3 points) Take  $\alpha = \frac{7\pi}{12}$ , we have

$$\mathbf{R}\mathbf{x} = \begin{bmatrix} \cos \frac{7\pi}{12} & -\sin \frac{7\pi}{12} \\ -\sin \frac{7\pi}{12} & \cos \frac{7\pi}{12} \end{bmatrix} \begin{bmatrix} \cos \frac{\pi}{4} \\ \sin \frac{\pi}{4} \end{bmatrix} = \begin{bmatrix} -\cos \frac{\pi}{6} \\ \sin \frac{\pi}{6} \end{bmatrix}.$$

2) (8 points) To prove that  $\mathbf{y} := \mathbf{H}\mathbf{x}$  is the reflection of  $\mathbf{x}$  with respect to  $\mathcal{H}_u$ .

- First, we prove that their projections are the same,

$$\begin{aligned} \mathbf{Q}\mathbf{y} &= \mathbf{Q}\mathbf{H}\mathbf{x} = (\mathbf{I} - \mathbf{u}\mathbf{u}^T)(\mathbf{I} - 2\mathbf{u}\mathbf{u}^T)\mathbf{x}, \\ &= (\mathbf{I} - 2\mathbf{u}\mathbf{u}^T - \mathbf{u}\mathbf{u}^T + 2\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T)\mathbf{x} \\ &= (\mathbf{I} - \mathbf{u}\mathbf{u}^T)\mathbf{x} = \mathbf{Q}\mathbf{x}. \end{aligned}$$

- Next, we prove that  $\|\mathbf{x} - \mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{y} - \mathbf{Q}\mathbf{y}\|_2$ ,

$$\|\mathbf{x} - \mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{x} - (\mathbf{I} - \mathbf{u}\mathbf{u}^T)\mathbf{x}\|_2 = \|\mathbf{u}\mathbf{u}^T\mathbf{x}\|_2,$$

and

$$\|\mathbf{y} - \mathbf{Q}\mathbf{y}\|_2 = \|\mathbf{H}\mathbf{x} - \mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{u}\mathbf{u}^T\mathbf{x}\|_2,$$

Therefore, by the definition of reflection, the proof completes.

**Remarks:** This problem consists of 2 sub-problems.

- 1) For the first sub-problem, first you are required to give all rotations matrices in  $\mathbb{R}^2$ , the derivation of rotation matrices and the conclusion are worthy of 2 points and 1 point respectively. Here, we give two ways to find all the rotation matrices, either way is accepted. Second, for a given  $\mathbf{x}$  and  $\theta$ , you are required to compute  $\mathbf{R}\mathbf{x}$ , this computation takes 3 points. However, since the computation of  $\mathbf{R}\mathbf{x}$  is completely based on the derivation of  $\mathbf{R}$ , you will get no points if  $\mathbf{R}$  you give is not correct.
- 2) For the second sub-problem, you are required to prove that  $\mathbf{H}\mathbf{x}$  is a reflection of  $\mathbf{x}$  with respect to  $\mathcal{H}_u$ . Directly prove  $\mathbf{Q}\mathbf{x} = \mathbf{Q}\mathbf{H}\mathbf{x}$  (4 points) and  $\|\mathbf{x} - \mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{H}\mathbf{x} - \mathbf{Q}\mathbf{H}\mathbf{x}\|_2$  (4 points) complete the proof.

## IV. DIRECT SUM

**Problem 1.**(10 points) This problem is graded by Song Mao (maosong@).

Let  $\mathcal{V}$  be a vector space, and  $\mathcal{B}$  be a basis for  $\mathcal{V}$ . Suppose that there exist subsets  $\mathcal{B}_1, \mathcal{B}_2$  of  $\mathcal{B}$ , such that  $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$  and  $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ . Then show that  $\mathcal{V} = \text{span}(\mathcal{B}_1) \oplus \text{span}(\mathcal{B}_2)$ .

**Solution.** Let  $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m, \mathbf{y}_1, \dots, \mathbf{y}_n\}$ ,  $\mathcal{B}_1 = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ , and  $\mathcal{B}_2 = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ . Let  $\mathcal{S}$  be the subspace spanned by  $\mathcal{B}_1$  and  $\mathcal{T}$  be the subspace spanned by  $\mathcal{B}_2$ .

- First, for any  $\mathbf{s} \in \mathcal{S}$ ,  $\mathbf{t} \in \mathcal{T}$ ,

$$\mathbf{s} + \mathbf{t} = \sum_{i=1}^m \alpha_i \mathbf{x}_i + \sum_{j=1}^n \beta_j \mathbf{y}_j,$$

which means that  $\mathbf{s} + \mathbf{t}$  can be represented by basis  $\mathcal{B}$ , therefore  $\mathbf{s} + \mathbf{t} \in \mathcal{V}$  and consequently  $\mathcal{S} + \mathcal{T} \subseteq \mathcal{V}$ . Also,  $\mathcal{V} \subseteq \mathcal{S} + \mathcal{T}$ . Therefore  $\mathcal{V} = \mathcal{S} + \mathcal{T}$ .

- Second, for  $\mathbf{z} \in \mathcal{S} \cap \mathcal{T}$ , there exists some  $\{\alpha_i\}_{i=1}^m$  and  $\{\beta_j\}_{j=1}^n$  then

$$\mathbf{z} = \sum_{i=1}^m \alpha_i \mathbf{x}_i = \sum_{j=1}^n \beta_j \mathbf{y}_j \Rightarrow \sum_{i=1}^m \alpha_i \mathbf{x}_i - \sum_{j=1}^n \beta_j \mathbf{y}_j = \mathbf{0}.$$

Since  $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$  is linearly independent, then  $\mathbf{z} = \mathbf{0}$  and it implies  $\mathcal{S} \cap \mathcal{T} = \{\mathbf{0}\}$ .

Consequently, by the definition of direct sum, we have  $\mathcal{V} = \text{span}(\mathcal{B}_1) \oplus \text{span}(\mathcal{B}_2)$ .

**Criterion:** Your grade depends on the Policy, and you may lose points because of Wrong cases:

1) **Policy.** You are required to provide details for the following two main parts of the proof to the problem.

- a) (5 points) Prove that  $\text{span}(\mathcal{B}_1) \cap \text{span}(\mathcal{B}_2) = \{\mathbf{0}\}$ .
- b) (5 points) Prove that  $\mathcal{V} = \text{span}(\mathcal{B}_1) + \text{span}(\mathcal{B}_2)$ . Precisely, prove that  $\mathcal{V} \subseteq \text{span}(\mathcal{B}_1) + \text{span}(\mathcal{B}_2)$  and  $\mathcal{V} \supseteq \text{span}(\mathcal{B}_1) + \text{span}(\mathcal{B}_2)$ .

You will only gain 5 points if you miss one part of the proof.

2) **Wrong cases.** You may lose points because of following mistakes:

- a) You will lose 2 points if your proof for  $\text{span}(\mathcal{B}_1) \cap \text{span}(\mathcal{B}_2) = \{\mathbf{0}\}$  is correct, but you write the result as  $\text{span}(\mathcal{B}_1) \cap \text{span}(\mathcal{B}_2) = \emptyset$ .
- b) When you write  $\dim(\mathcal{B})$ , you get 0 point for this part of proof, since  $\mathcal{B}$  is not a subspace, it is totally wrong to use  $\dim$  for a set. The correct form should be  $\dim(\text{span}(\mathcal{B}))$ .
- c) You may lose points if your proof for  $\mathcal{V} = \text{span}(\mathcal{B}_1) + \text{span}(\mathcal{B}_2)$  is too simple.
- d) The following claim:

$$\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset \implies \text{span}(\mathcal{B}) \cap \text{span}(\mathcal{B}) = \{\mathbf{0}\}$$

is wrong if you did not mention explicitly that  $\mathcal{B}_1 \cup \mathcal{B}_2$  is a list of linear independent vectors or a basis.

You will get 0 point for this part of proof.

- e) It's wrong to prove

$$\mathcal{V} \subseteq \text{span}(\mathcal{B}_1) \oplus \text{span}(\mathcal{B}_2)$$



and

$$\mathcal{V} \supseteq \text{span}(\mathcal{B}_1) \oplus \text{span}(\mathcal{B}_2)$$

since the direct sum may not be well-defined. You will get 0 point for this problem.

f) Though you gain 5 points if you prove the result that (no proof for the part a))

$$\mathcal{V} = \text{span}(\mathcal{B}_1) \oplus \text{span}(\mathcal{B}_2)$$

you should realize that *the direct sum is not well-defined* until you prove

$$\text{span}(\mathcal{B}_1) \cap \text{span}(\mathcal{B}_2) = \{0\}$$

**Problem 2.** (10 points) This problem is graded by Sihang Xu (xush@).

Let  $\mathcal{V}$  be a real vector space of dimension  $n$ . Let  $\mathcal{S}$  be a subspace of  $\mathcal{V}$  of dimension  $d \leq n$ . Prove that there exists a subspace  $\mathcal{T}$  of  $\mathcal{V}$  such that  $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$ .

**Solution.** We prove the claim by induction on  $d$ .

- When  $d = n - 0$ ,  $\mathcal{V} = \mathcal{S} = \mathcal{S} \oplus \{\mathbf{0}\}$ , here  $\mathcal{T} = \{\mathbf{0}\}$ .
- Assume that when  $d = n - k$ , the claim holds.
- Next, when  $d = n - (k + 1)$ ,  $\mathcal{S} \subsetneq \mathcal{V}$ , we can choose one vector  $\mathbf{v} \in \mathcal{V} \setminus \mathcal{S}$ , while  $\mathcal{S} \oplus \text{span}\{\mathbf{v}\}$  is a subspace of  $\mathcal{V}$  of dimension  $d + 1$ , i.e.,  $\dim(\mathcal{S} \oplus \text{span}\{\mathbf{v}\}) = n - k$ . By the hypothesis, there is a subspace  $\mathcal{T}'$  s.t.  $\mathcal{V} = (\mathcal{S} \oplus \text{span}\{\mathbf{v}\}) \oplus \mathcal{T}' = \mathcal{S} \oplus (\text{span}\{\mathbf{v}\} \oplus \mathcal{T}')$ . Consequently, there exists a subspace  $\mathcal{T} = \text{span}\{\mathbf{v}\} \oplus \mathcal{T}'$  of  $\mathcal{V}$  s.t.  $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$ .

By the principle of mathematical induction, the claim is proved.

**Remarks:** If you show one of following, then you will get the full credit:

- Prove the claim by induction as solution.
- Show a basis of  $\mathcal{V}$  can be extended by a basis of  $\mathcal{S}$ , then using the result of problem 1.
- Let  $\mathcal{T} = \mathcal{S}^\perp$ . (Not recommended)

Common mistake:

- Some students assume  $\mathcal{B} = \{v_1, \dots, v_n\}$  is a basis of  $V$ , and  $\mathcal{B}_1 \subset \mathcal{B}$  is a basis of  $B$ . In this case, 0 point will given because this claim is wrong. Consider the following case:  $V = \mathbb{R}^2$ ,  $\mathcal{B} = \{(1, 0), (0, 1)\}$ ,  $\mathcal{S} = \{(x, y) \mid x + y = 1\}$ . In this case,  $\mathcal{S}$  cannot be span of any subset of  $\mathcal{B}$ .

## V. UNDERSTANDING THE MATRIX NORM

(7 points  $\times$  2) This problem is graded by Xinyue Zhang (zhangxy11@).

Matrix norm is induced by vector norm,

$$\|\mathbf{A}\|_p = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|\mathbf{Ax}\|_p, \quad \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{x} \in \mathbb{R}^{n \times 1},$$

prove that (a)

1) the matrix 1-norm

$$\begin{aligned} \|\mathbf{A}\|_1 &= \max_{\|\mathbf{x}\|_1=1} \|\mathbf{Ax}\|_1 = \max_j \sum_i^m |a_{ij}| \\ &= \text{the largest absolute column sum.} \end{aligned}$$

2) the matrix  $\infty$ -norm

$$\begin{aligned} \|\mathbf{A}\|_\infty &= \max_{\|\mathbf{x}\|_\infty=1} \|\mathbf{Ax}\|_\infty = \max_i \sum_j^n |a_{ij}| \\ &= \text{the largest absolute row sum.} \end{aligned}$$

Solution.

1) (7 points) For all  $\mathbf{x}$  with  $\|\mathbf{x}\|_1 = 1$ , the scalar triangle inequality yields

$$\begin{aligned} \|\mathbf{Ax}\|_1 &= \sum_i |\mathbf{A}_{i*}\mathbf{x}| = \sum_i \left| \sum_j a_{ij}x_j \right| \leq \sum_i \sum_j |a_{ij}||x_j| = \sum_j \left( |x_j| \sum_i |a_{ij}| \right) \\ &\leq \left( \sum_j |x_j| \right) \left( \max_j \sum_i |a_{ij}| \right) = \max_j \sum_i |a_{ij}|, \end{aligned}$$

where  $\mathbf{A}_{i*}$  denotes  $i$ -th row of matrix  $\mathbf{A}$ . 6' Equality can be attained because if  $\mathbf{A}_{*k}$  is the column with the largest absolute sum, set  $\mathbf{x} = \mathbf{e}_k$ , and note that  $\|\mathbf{e}_k\|_1 = 1$  and  $\|\mathbf{A}\mathbf{e}_k\|_1 = \|\mathbf{A}_{*k}\|_1 = \max_j \sum_i |a_{ij}|$ . 1'

2) (7 points) For all  $\mathbf{x}$  with  $\|\mathbf{x}\|_\infty = 1$ ,

$$\|\mathbf{Ax}\|_\infty = \max_i \left| \sum_j a_{ij}x_j \right| \leq \max_i \sum_j |a_{ij}||x_j| \leq \max_i \sum_j |a_{ij}|.$$

6' Equality can be attained because if  $\mathbf{A}_{k*}$  is the row with the largest absolute sum, and if  $\mathbf{x}$  is the vector such that

$$x_j = \begin{cases} 1 & \text{if } a_{kj} \geq 0, \\ -1 & \text{if } a_{kj} < 0, \end{cases} \quad \text{then } \begin{cases} |\mathbf{A}_{i*}\mathbf{x}| = \left| \sum_j a_{ij}x_j \right| \leq \sum_j |a_{ij}| \text{ for all } i, \\ |\mathbf{A}_{k*}\mathbf{x}| = \sum_j |a_{kj}| = \max_i \sum_j |a_{ij}|, \end{cases}$$

so  $\|\mathbf{x}\|_\infty = 1$  and  $\|\mathbf{Ax}\|_\infty = \max_i |\mathbf{A}_{i*}\mathbf{x}| = \max_i \sum_j |a_{ij}|$ . 1'

## VI. UNDERSTANDING THE HÖLDER INEQUALITY

(6 points  $\times$  3 + 10 points) This problem is graded by **Yijia Chang (changyj@)**.

Hölder inequality:

$$|\mathbf{x}^T \mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q,$$

for any  $p, q$  such that  $1/p + 1/q = 1$ ,  $p \geq 1$ . Derive this inequality by executing the following steps: (a)

- 1) Consider the function  $f(t) = (1 - \lambda) + \lambda t - t^\lambda$  for  $0 < \lambda < 1$ , establish the inequality

$$\alpha^\lambda \beta^{1-\lambda} \leq \lambda \alpha + (1 - \lambda) \beta,$$

for nonnegative real numbers  $\alpha$  and  $\beta$ .

- 2) Let  $\hat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|_p$  and  $\hat{\mathbf{y}} = \mathbf{y}/\|\mathbf{y}\|_q$ , and apply the inequality of part (a) to obtain

$$\sum_{i=1}^n |\hat{x}_i \hat{y}_i| \leq \frac{1}{p} \sum_{i=1}^n |\hat{x}_i|^p + \frac{1}{q} \sum_{i=1}^n |\hat{y}_i|^q = 1.$$

- 3) Deduce the Hölder inequality with the above results.  
4) (Bouns question) Prove the general form of triangle inequality

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p.$$

**Hint:** For  $p > 1$ , let  $q$  be the number such that  $1/q = 1 - 1/p$ . Verify that for scalars  $\alpha$  and  $\beta$ ,

$$|\alpha + \beta|^p = |\alpha + \beta| |\alpha + \beta|^{p/q} \leq |\alpha| |\alpha + \beta|^{p/q} + |\beta| |\alpha + \beta|^{p/q}$$

and make use of Hölder's inequality.

**Solution.**

- 1) (6 points) By taking the derivative,

$$f'(t) = \lambda - \lambda t^{\lambda-1},$$

then we have that  $f'(t) < 0$  for  $t < 1$  and  $f'(t) > 0$  for  $t > 1$ , therefore we can conclude that  $f(t) \geq f(1) = 0$ .

In the case of  $\beta > 0$ , setting  $t = \alpha/\beta$  provides the desired inequality. In the case of  $\beta = 0$ , the desired inequality is  $0 \leq \lambda \alpha$ , which still holds because  $\lambda > 0$  and  $\alpha \geq 0$ .

- 2) (6 points) It is easy to verify that the inequality in part (a) still holds in the case of  $\lambda = 0$  and  $\lambda = 1$ . Let

$$\alpha = |\hat{x}_i|^p, \quad \beta = |\hat{y}_i|^q, \quad \lambda = 1/p, \quad \text{and} \quad (1 - \lambda) = 1/q,$$

then we have

$$\begin{aligned} |\hat{x}_i|^{p \times 1/p} |\hat{y}_i|^{q \times 1/q} &\leq \frac{1}{p} |\hat{x}_i|^p + \frac{1}{q} |\hat{y}_i|^q \\ \Rightarrow \sum_{i=1}^n |\hat{x}_i \hat{y}_i| &\leq \frac{1}{p} \sum_{i=1}^n |\hat{x}_i|^p + \frac{1}{q} \sum_{i=1}^n |\hat{y}_i|^q = \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

- 3) (6 points) Consequently we have,

$$|\mathbf{x}^T \mathbf{y}| = \left| \sum_{i=1}^n x_i y_i \right| = \|\mathbf{x}\|_p \|\mathbf{y}\|_q \left| \sum_{i=1}^n \hat{x}_i \hat{y}_i \right| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q \sum_{i=1}^n |\hat{x}_i \hat{y}_i| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q.$$

- 4) (4 points) The inequality in the hint follows from the fact that  $p = 1 + p/q$  together with the scalar triangle inequality, and it implies that

$$\sum_{i=1}^n |x_i + y_i|^p = \sum_{i=1}^n |x_i + y_i| |x_i + y_i|^{p/q} \leq \sum_{i=1}^n |x_i| |x_i + y_i|^{p/q} + \sum_{i=1}^n |y_i| |x_i + y_i|^{p/q}.$$

- (4 points) By Hölder's inequality, we have that

$$\begin{aligned} \sum_{i=1}^n |x_i| |x_i + y_i|^{p/q} &\leq \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \left( \sum_{i=1}^n |x_i + y_i|^{p/q \times q} \right)^{1/q} \\ &= \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \left( \sum_{i=1}^n |x_i + y_i|^p \right)^{p-1/p} \\ &= \|\mathbf{x}\|_p \|\mathbf{x} + \mathbf{y}\|_p^{p-1}. \end{aligned}$$

Similarly, we have that

$$\sum_{i=1}^n |y_i| |x_i + y_i|^{p/q} \leq \|\mathbf{y}\|_p \|\mathbf{x} + \mathbf{y}\|_p^{p-1},$$

- (2 points) Therefore,

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