SI231 - Matrix Computations, 2021 Fall

Solution of Homework Set #1

Prof. Yue Qiu

Acknowledgements:

1) Deadline: **2021-10-12 23:59:59**
2) **Late Policy details** can be found on piazza.
3) Submit your homework in **Homework 1** on **Gradscope**. Entry Code: **2RY68R**. Make sure that you have correctly select pages for each problem. If not, you probably will get 0 point.
4) No handwritten homework is accepted. You need to write LaTeX. (If you have difficulty in using LaTeX, you are allowed to use **MS Word or Pages** for the first and the second homework to accommodate yourself.)
5) Use the given template and give your solution in English. Solution in Chinese is not allowed.
6) Your homework should be uploaded in the PDF format, and the naming format of the file is not specified.

I. Vector space and subspace

Problem 1. (6 points $\times$ 3)

1) Let $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ be two subspaces of a vector space $\boldsymbol{\mathcal{V}}$:
   a) Prove that the intersection $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{Y}}$ is also a subspace of $\boldsymbol{\mathcal{V}}$.
   b) Show that the union of $\boldsymbol{\mathcal{X}} \cup \boldsymbol{\mathcal{Y}}$ need not to be a subspace of $\boldsymbol{\mathcal{V}}$.
2) Prove or give a counterexample:
   a) If $\boldsymbol{\mathcal{U}}_1$, $\boldsymbol{\mathcal{U}}_2$, and $\boldsymbol{\mathcal{W}}$ are subspaces of $\boldsymbol{\mathcal{V}}$ such that $\boldsymbol{\mathcal{U}}_1 + \boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{U}}_2 + \boldsymbol{\mathcal{W}}$, then $\boldsymbol{\mathcal{U}}_1 = \boldsymbol{\mathcal{U}}_2$.
   b) If $\boldsymbol{\mathcal{U}}_1$, $\boldsymbol{\mathcal{U}}_2$, and $\boldsymbol{\mathcal{W}}$ are subspaces of $\boldsymbol{\mathcal{V}}$ such that $\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{U}}_1 \oplus \boldsymbol{\mathcal{W}}$ and $\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{U}}_2 \oplus \boldsymbol{\mathcal{W}}$, then $\boldsymbol{\mathcal{U}}_1 = \boldsymbol{\mathcal{U}}_2$.[1].
3) Let $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r\}$ and $\mathbf{V} = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r, \mathbf{v}\}$ be two sets of vectors from the same vector space, prove that $span(\mathbf{U}) = span(\mathbf{V})$ if and only if $\mathbf{v} \in span(\mathbf{U})$.

Solution:

1) a) Let $\mathbf{u}, \mathbf{v} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{Y}}$ then $\mathbf{u}, \mathbf{v} \in \boldsymbol{\mathcal{X}}$ and $\mathbf{u}, \mathbf{v} \in \boldsymbol{\mathcal{Y}}$. From the closure property for vector addition that $\mathbf{u}+\mathbf{v} \in \boldsymbol{\mathcal{X}}$ and $\mathbf{u}+\mathbf{v} \in \boldsymbol{\mathcal{Y}}$, we have $\mathbf{u}+\mathbf{v} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{Y}}$. (2 points) From the closure property for scalar multiplication that $\alpha\mathbf{u} \in \boldsymbol{\mathcal{X}}$ and $\alpha\mathbf{u} \in \boldsymbol{\mathcal{Y}}$ for any α, we have $\alpha\mathbf{u} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{Y}}$.(2 points)
   b) For example, the union of two lines through the Origin in $\mathbb{R}^2$ is not a subspace. (2 points)
2) a) Counterexample: If $\boldsymbol{\mathcal{U}}_1 = \{(z,z) \in \mathbb{R}^2 : x \in \mathbb{R}\}$,$\boldsymbol{\mathcal{U}}_2 = \{(0,y) \in \mathbb{R}^2 : x \in \mathbb{R}\}$ and $\boldsymbol{\mathcal{W}} = \{(x,0) \in \mathbb{R}^2 : z \in \mathbb{R}\}$, then $\boldsymbol{\mathcal{U}}_1 + \boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{U}}_2 + \boldsymbol{\mathcal{W}}$ but $\boldsymbol{\mathcal{U}}_1 \neq \boldsymbol{\mathcal{U}}_2$. (3 points)
   b) Counterexample: If $\boldsymbol{\mathcal{U}}_1 = \{(z,z) \in \mathbb{R}^2 : x \in \mathbb{R}\}$, $\boldsymbol{\mathcal{U}}_2 = \{(0,y) \in \mathbb{R}^2 : x \in \mathbb{R}\}$ and $\boldsymbol{\mathcal{W}} = \{(x,0) \in \mathbb{R}^2 : z \in \mathbb{R}\}$, implies $\boldsymbol{\mathcal{U}}_1 \cap \boldsymbol{\mathcal{W}} = \{\mathbf{0}\}$ and $\boldsymbol{\mathcal{U}}_2 \cap \boldsymbol{\mathcal{W}} = \{\mathbf{0}\}$ then $\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{U}}_1 \oplus \boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{U}}_2 \oplus \boldsymbol{\mathcal{W}}$ but $\boldsymbol{\mathcal{U}}_1 \neq \boldsymbol{\mathcal{U}}_2$. (3 points)

[1] 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** $\mathbb{R}^n = \mathcal{S}_1 \oplus \mathcal{S}_2$.

3) $\Rightarrow$ If $span(\mathbf{U}) = span(\mathbf{V})$ then $\mathbf{v}$ must be a linear combination of $\mathbf{u}_i$'s, then $\mathbf{v} \in span(\mathbf{U})$.(2 points)

$\Leftarrow$ First we can notice that $span(\mathbf{U}) \subseteq span(\mathbf{V})$. And given $\mathbf{v} \in span(\mathbf{U})$ there exists the linear combination $\mathbf{v} = \sum_{i=1}^{r} \alpha_i \mathbf{u}_i$. For any $\mathbf{y} \in span(\mathbf{V})$ we have

$$\begin{aligned}\mathbf{y} &= \sum_{i=1}^{r} \beta_i \mathbf{u}_i + \beta_{r+1} \mathbf{v} \\ &= \sum_{i=1}^{r} \beta_i \mathbf{u}_i + \beta_{r+1} \sum_{i=1}^{r} \alpha_i \mathbf{u}_i \\ &= \sum_{i=1}^{r} (\beta_i + \beta_{r+1} \alpha_i) \mathbf{u}_i.\end{aligned}$$

Then $\mathbf{y} \in span(\mathbf{U})$, therefore $span(\mathbf{U}) \supseteq span(\mathbf{V})$.(4 points)

II. Basis, dimension and rank

Problem 1. (5 points $\times$ 2) For matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, we have $\mathcal{V} = \{\mathbf{X} \in \mathbb{R}^{n\times n} | \mathbf{AX} = \mathbf{XA}\}$,

1) Prove that $\mathcal{V}$ is a linear subspace of the linear space $\mathbb{R}^{n\times n}$;
2) If $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$, please give a basis and the dimension of $\mathcal{V}$.

Solution:

1) For $\forall k \in \mathbb{R}^n$, and $\forall \mathbf{X}, \mathbf{Y} \in \mathcal{V}$, we have:

   $\mathbf{A0}_{n\times n} = \mathbf{0}_{n\times n}\mathbf{A} \Rightarrow \{\mathbf{0}_{n\times n}\} \in \mathcal{V}$. (1 points)

   $\mathbf{A}(k\mathbf{X}) = k\mathbf{AX} = k\mathbf{XA} = (k\mathbf{X})\mathbf{A} \Rightarrow k\mathbf{X} \in \mathcal{V}$. (2 points)

   $\mathbf{A}(\mathbf{X}+\mathbf{Y}) = \mathbf{AX} + \mathbf{AY} = \mathbf{XA} + \mathbf{YA} = (\mathbf{X}+\mathbf{Y})\mathbf{A} \Rightarrow \mathbf{X}+\mathbf{Y} \in \mathcal{V}$. (2 points)

   So we can conclude that $\mathcal{V}$ is the linear subspace of the linear space $\mathbb{R}^{n\times n}$.

2) For $\forall \mathbf{X} = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \in \mathcal{V}$:

$$\mathbf{AX} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = \begin{pmatrix} x_1 + x_3 & x_2 + x_4 \\ 2x_1 - x_3 & 2x_2 - x_4 \end{pmatrix}$$

$$\mathbf{XA} = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 & x_1 - x_2 \\ x_3 + 2x_4 & x_3 - x_4 \end{pmatrix}$$

$$\mathbf{AX} = \mathbf{XA} \Rightarrow \begin{pmatrix} x_1 + x_3 & x_2 + x_4 \\ 2x_1 - x_3 & 2x_2 - x_4 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 & x_1 - x_2 \\ x_3 + 2x_4 & x_3 - x_4 \end{pmatrix} \Rightarrow \begin{cases} x_1 = 2x_2 + x_4 \\ x_2 = x_2 \\ x_3 = 2x_2 \\ x_4 = x_4 \end{cases} \quad \text{(1 point)}$$

$$\Rightarrow \mathbf{X} = \begin{pmatrix} 2x_2 + x_4 & x_2 \\ 2x_2 & x_4 \end{pmatrix} = x_2 \begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So for $\forall \mathbf{X} \in \mathcal{V}$, $\mathbf{X} = \begin{pmatrix} 2x_2 + x_4 & x_2 \\ 2x_2 & x_4 \end{pmatrix} = x_2 \begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

And $\begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are linear independent.

Then $\left\{ \begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is a basis of $\boldsymbol{\mathcal{V}}$ (2 points) whose dimension is 2. (2 points)

Problem 2. (5 points)The linear space $\mathcal{S}$ contains the following polynomials: $f_1(t) = 1 + 4t - 2t^2 + t^3$, $f_2(t) = -1 + 9t - 3t^2 + 2t^3$, $f_3(t) = -5 + 6t + t^3$, $f_4(t) = 5 + 7t - 5t^2 + 2t^3$. Please give the rank of the quadruple $(f_1(t), f_2(t), f_3(t), f_4(t))$ and its maximal linearly independent set.

Solution: We can choose a set of basis in linear space $\mathcal{S}$: $1, t, t^2, t^3$. (1 point)

$$(f_1(t), f_2(t), f_3(t), f_4(t)) = \begin{bmatrix} 1 & t & t^2 & t^3 \end{bmatrix} \begin{bmatrix} 1 & -1 & -5 & 5 \\ 4 & 9 & 6 & 7 \\ -2 & -3 & 0 & -5 \\ 1 & 2 & 1 & 2 \end{bmatrix} (1 point)$$

.

$$\begin{bmatrix} 1 & -1 & -5 & 5 \\ 4 & 9 & 6 & 7 \\ -2 & -3 & 0 & -5 \\ 1 & 2 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & -5 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} (1 point)$$

So the rank of the $(f_1(t), f_2(t), f_3(t), f_4(t))$ is 2 (1 point) and its maximal linearly independent set is $(f_1(t), f_2(t))$ or $(f_1(t), f_3(t))$ or $(f_1(t), f_4(t))$ (1 point).

Problem 3.(5 points $\times$ 2) For any matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathcal{S}_1 = \{\mathbf{A} \in \mathbb{R}^{n\times n} | \mathbf{A}^T = \mathbf{A}\}$ and $\mathcal{S}_2 = \{\mathbf{A} \in \mathbb{R}^{n\times n} | \mathbf{A}^T = -\mathbf{A}\}$ are two subspaces of $\mathbb{R}^{n\times n}$,

1) Prove that $\mathbb{R}^{n\times n} = \mathcal{S}_1 \oplus \mathcal{S}_2$.
2) If n = 3, please give a basis of $\mathcal{S}_1$ and the dimension of $\mathcal{S}_2$.

Solution:

1) For $\forall$A$\in$ $\mathbb{R}^{n\times n}$, suppose $\mathbf{B} = \frac{A+A^T}{2}$, $\mathbf{C} = \frac{A-A^T}{2}$, then we can get $\mathbf{B} \in \mathcal{S}_1$, $\mathbf{C} \in \mathcal{S}_2$, and $\mathbf{A} = \mathbf{B} + \mathbf{C}$. $\mathbb{R}^{n\times n} \subseteq \mathcal{S}_1 + \mathcal{S}_2$. (2 points) Vice versa, $\mathcal{S}_1 + \mathcal{S}_2 \subseteq \mathbb{R}^{n\times n}$. (2 points) So, $\mathbb{R}^{n\times n} = \mathcal{S}_1 + \mathcal{S}_2$.
   If $\mathbf{A} \in \mathcal{S}_1 \cap \mathcal{S}_2$, $\mathbf{A}^T = \mathbf{A}$ and $\mathbf{A}^T = -\mathbf{A}$, so $\mathbf{A} = \mathbf{0}_{n\times n}$, i.e. $\mathcal{S}_1 \cap \mathcal{S}_2$ = $\{\mathbf{0}_{n\times n}\}$. (1 point)
   Finally, we can conclude: $\mathbb{R}^{n\times n} = \mathcal{S}_1 \oplus \mathcal{S}_2$.
2) When n = 3, if $\mathbf{A} \in \mathcal{S}_1$, then $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$ = $a_{11} \cdot \mathbf{E}_{11}$ + $a_{22} \cdot \mathbf{E}_{22}$ + $a_{33} \cdot \mathbf{E}_{33}$ + $a_{12} \cdot (\mathbf{E}_{12} + \mathbf{E}_{21})$ + $a_{13} \cdot (\mathbf{E}_{13} + \mathbf{E}_{31})$ + $a_{23} \cdot (\mathbf{E}_{23} + \mathbf{E}_{32})$. (1 point)
   This shows: matrix $\mathbf{A}$ can be expressed linearly with $\mathbf{E}_{11}, \mathbf{E}_{22}, \mathbf{E}_{33}, \mathbf{E}_{12} + \mathbf{E}_{21}, \mathbf{E}_{13} + \mathbf{E}_{31}, \mathbf{E}_{23} + \mathbf{E}_{32}$. Because they are linear independent, $\mathbf{E}_{11}, \mathbf{E}_{22}, \mathbf{E}_{33}, \mathbf{E}_{12} + \mathbf{E}_{21}, \mathbf{E}_{13} + \mathbf{E}_{31}, \mathbf{E}_{23} + \mathbf{E}_{32}$ are the basis of $\mathcal{S}_1$. (2 points)
   $dim(\mathcal{S}_2)$ = $\mathbb{R}^{3\times 3} - dim(\mathcal{S}_1) = 9 - 6 = 3$. (2 points)

III. Four fundamental subspaces

Problem 1.(2 points + 5 points) For an $n \times m$ real matrix $\mathbf{A}$.

1) Determine the relationship of $dim(\mathcal{R}(\mathbf{A}))$, $dim(\mathcal{N}(\mathbf{A}))$, and $rank(\mathbf{A})$.
2) Prove that $\mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A^T}) = \mathbb{R}^m$.

Solution:

1) $dim(\mathcal{R}(\mathbf{A})) = rank(\mathbf{A})$, and $dim(\mathcal{N}(\mathbf{A})) = \mathbf{m} - dim(\mathcal{R}(\mathbf{A}))$ Remarks:Each one is 1 point
2) $\mathcal{N}(\mathbf{A})$ and $\mathcal{R}(\mathbf{A^T})$ are two subspace of vector space $\mathbb{R}^\mathbf{m}$, then $dim(\mathcal{N}(\mathbf{A}) + \mathcal{R}(\mathbf{A^T})) = dim(\mathcal{N}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{A^T})) - dim(\mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{A^T}))$. Because $\mathcal{R}(\mathbf{A}) \cap \mathcal{N}(\mathbf{A^T}) = \{0\}$, so $dim(\mathcal{N}(\mathbf{A}) \cap \mathcal{R}(\mathbf{A^T})) = 0$, (2.5 point) so $dim(\mathcal{N}(\mathbf{A}) + \mathcal{R}(\mathbf{A^T})) = dim(\mathcal{N}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{A^T})) = \mathbf{m}$, so $\mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A^T}) = \mathbb{R}^m$.(2.5 point)

Problem 2.(3 points + 5 points $\times$ 3) Given matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times m}$ and $rank([\mathbf{A}, \mathbf{B}]) = n$.

1) Determine the relationship between $dim(\mathcal{N}(\mathbf{A^T}))$ and $dim(\mathcal{R}(\mathbf{B}))$.
2) If $\mathbf{A^T B} = 0$, determine the relationship between $\mathcal{N}(\mathbf{A^T})$ and $\mathcal{R}(\mathbf{B})$.
3) Please determine the rank of $\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}$.
4) Please determine the **Supremacy** and **Infimum** of the rank of $\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}$ using m or n.

   (All Matrix $\mathbf{A}$'s, and $\mathbf{B}$'s that satisfy the mentioned condition)

Solution:

1) $n = dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{N}(\mathbf{A^T}))$, and $n = rank([\mathbf{A}, \mathbf{B}]) = dim([\mathbf{A}, \mathbf{B}]) = dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{B})) - dim(\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})) <= dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{B}))$, Remarks: process 2 point so $dim(\mathcal{R}(\mathbf{B})) >= dim(\mathcal{N}(\mathbf{A^T}))$ Remarks: conclusion 1 point
2) Because $\mathbf{A^T B} = 0$, so $\mathcal{R}(\mathbf{B}) \subseteq \mathcal{N}(\mathbf{A^T})$. and so $dim(\mathcal{R}(\mathbf{B})) \leq dim(\mathcal{N}(\mathbf{A^T}))$; since as problem1 $dim(\mathcal{R}(\mathbf{B})) \geq dim(\mathcal{N}(\mathbf{A^T}))$, so $dim(\mathcal{R}(\mathbf{B})) = dim(\mathcal{N}(\mathbf{A^T}))$.Remarks: process 2 point . sum up, $\mathcal{R}(\mathbf{B}) = \mathcal{N}(\mathbf{A^T})$. conclusion 2 point .
3) If you consider the $\mathbf{A^T B} = 0$ condition, the answer is: $rank(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix})) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix})) + dim(\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix})) - dim(\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix}) \cap \mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix}))$, because $\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix}) \cap \mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix}) = \emptyset$, so $dim(\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix}) \cap \mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix})) = 0$, Remarks: process 2 point . sum up, $rank(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix})) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix})) + dim(\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix})) = dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{B})) = dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{N}(\mathbf{A^T})) = n$ (as above conditions, we have $dim(\mathcal{R}(\mathbf{B})) = dim(\mathcal{N}(\mathbf{A^T}))$). Remarks: conclusion 2.5 point .

   If you not consider the $\mathbf{A^T B} = 0$ condition, the answer is: $rank(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix})) =$

$dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix})) + dim(\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix})) = dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{B}))$, $dim(\mathcal{R}(\mathbf{A})) + dim(\mathcal{R}(\mathbf{B})) \geq n$,$dim(\mathcal{R}(\mathbf{A})) \leq min(m,n)$, $dim(\mathcal{R}(\mathbf{B})) \leq min(m,n)$. Remarks: process 2 point . sum up:$n \leq rank(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}) \leq min(2n, 2m)$ Remarks: conclusion 2.5 point .

4) If you consider the $\mathbf{A^T B} = 0$ condition, the answer is: $dim(\mathcal{R}(\mathbf{A})) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ 0 \end{pmatrix})) <= dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}))$, and $dim(\mathcal{R}(\mathbf{B})) = dim(\mathcal{R}(\begin{pmatrix} 0 \\ \mathbf{B} \end{pmatrix})) <= dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}))$, so $max(dim(\mathcal{R}(\mathbf{A})), dim(\mathcal{R}(\mathbf{B}))) <= dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}))$; and $n = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix})) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ \mathbf{B} & \mathbf{B} \end{pmatrix})) >= dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ \mathbf{B} & 0 \end{pmatrix})) = dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}))$. sum up, $\lceil n/2 \rceil <= max(dim(\mathcal{R}(\mathbf{A})), dim(\mathcal{R}(\mathbf{B}))) <= dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix})) <= min(dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}))$ $min(n, m)$. Please using some special cases to prove the advisability of the equal sign. If you not consider the $\mathbf{A^T B} = 0$ condition, the answer is: $\lceil n/2 \rceil <= max(dim(\mathcal{R}(\mathbf{A})), dim(\mathcal{R}(\mathbf{B}))) <= dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix})) <= min(dim(\mathcal{R}(\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix})), m) = min(2n, m)$. Please using some special cases to prove the advisability of the equal sign. Remarks:Supremacy 2.5 points and Infimum 2.5 points ,if your answer is right ,you will get 5 points

IV. Vector norm and matrix norm

Problem 1. (5 points $\times$ 3) The Frobenius norm of a $\mathbb{R}^{m\times n}$ matrix $\mathbf{A}$ defined as the square root of the sum of the absolute squares of its elements,

$$\|\mathbf{A}\|_F \equiv \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2},$$

it also equal to the square root of the matrix trace of $\mathbf{A}^T\mathbf{A}$, where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$,

$$\|\mathbf{A}\|_F = \sqrt{\mathbf{Tr}(\mathbf{A}^T\mathbf{A})}.$$

1) Show that Frobenius norm is a matrix norm.

   Hint: You may use the Cauchy-Schwarz inequality

$$\|\mathbf{AB}\|_F \leq \|\mathbf{A}\|_F\|\mathbf{B}\|_F$$

2) The spectral norm of a matrix $\mathbf{A}$ is the largest singular value of $\mathbf{A}$ (the square root of the largest eigenvalue of the matrix $\mathbf{A}^T\mathbf{A}$,

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{max}(\mathbf{A}^T\mathbf{A})}.$$

   Show that $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \leq \sqrt{n}\|\mathbf{A}\|_2$

3) Suppose $\mathbf{A} = \mathbf{x}\mathbf{y}^T$, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, show that

$$\|\mathbf{A}\|_F^2 = \|\mathbf{x}\|_2^2\|\mathbf{y}\|_2^2$$

Solution:

1) If $\mathbf{A} = 0$, it's obvious that $\|\mathbf{A}\|_F = 0$, if $\mathbf{A} \neq 0$, there exists a element $a_{ij} \neq 0$, one has

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} \geq \sqrt{a_{ij}} > 0.\text{(1 point)}$$

   Let $\alpha \in \mathbb{R}$ be arbitrarily chosen,

$$\|\alpha\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|\alpha a_{ij}|^2} = \sqrt{\alpha^2\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = |\alpha|\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = |\alpha|\|\mathbf{A}\|_F.\text{(2 points)}$$

   Let $\mathbf{B}$ be a $\mathbb{R}^{n\times m}$ matrix,

$$\|\mathbf{A}+\mathbf{B}\|_F^2 = \mathbf{Tr}((\mathbf{A}+\mathbf{B})^T(\mathbf{A}+\mathbf{B})) = \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 + \mathbf{Tr}(\mathbf{A}^T\mathbf{B}+\mathbf{B}^T\mathbf{A})\text{(2 points)}$$

   Using Cauchy-Schwarz inequality we have

$$\mathbf{Tr}(\mathbf{A}^T\mathbf{B}+\mathbf{B}^T\mathbf{A}) \leq 2\|\mathbf{A}\|_F\|\mathbf{B}\|_F$$

   and this gives

$$\|\mathbf{A}+\mathbf{B}\|_F^2 \leq \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 + 2\|\mathbf{A}\|_F\|\mathbf{B}\|_F = (\|\mathbf{A}\|_F + \|\mathbf{B}\|_F)^2$$

It is obvious that (Cauchy-Schwarz inequality)

$$\|\mathbf{AB}\|_F \le \|\mathbf{A}\|_F \|\mathbf{B}\|_F$$

2) Let $\sigma_1, \sigma_2, \cdots, \sigma_n$ be singular values of $\mathbf{A}$. Since

$$\sigma_1 \le \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2} \le \sqrt{\sigma_1^2 + \sigma_1^2 + \cdot + \sigma_1^2} = \sqrt{n}\sigma_1.$$

Remarks:Supremacy 2 points and Infimum 2 points ,if your answer is right ,you will get 5 points

3)

$$\|\mathbf{A}\|_F^2 = \sum_{i=1}^{n} \|y_i \mathbf{x}\|_2^2 = \sum_{i=1}^{n} y_i^2 \|\mathbf{x}\|_2^2 = \|\mathbf{x}\|_2^2 \sum_{i=1}^{n} y_i^2 = \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2 \text{(5 points)}$$

V. Projector and projection

Problem 1. (2points + 5 points $\times$ 3) A rotation matrix $\mathbf{R} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix ($\mathbf{R}\mathbf{R}^T = \mathbf{I_n}$).[2]

1) According to the above definition, find all rotation matrices in $\mathbb{R}^{2\times 2}$.
2) Let $\mathbf{R}_1$ and $\mathbf{R}_2$ be the rotation matrices in $\mathbb{R}^{2\times 2}$, if $\mathbf{R}_1$ is the rotation matrix that rotates vectors counterclockwise through α_1 and $\mathbf{R}_2$ is the rotation matrix that rotates vectors counterclockwise through α_2. Consider: is $\mathbf{R}_1\mathbf{R}_2$ the rotation matrix ? If the answer is "yes", what is the angle of rotation, or else explain why the answer is "no".
3) For arbitrarily rotation matrix $\mathbf{R} \in \mathbb{R}^{n\times n}$, if $\mathbf{S} = (\mathbf{R} - \mathbf{I_n})(\mathbf{R} + \mathbf{I_n})^{-1}$, show that $\mathbf{S}$ is a skew symmetric matrix ($\mathbf{S}^T = -\mathbf{S}$)
4) If $\mathbf{S} \in \mathbb{R}^{n\times n}$ is a skew symmetric matrix, show that $\mathbf{R} = (\mathbf{I_n} - \mathbf{S})^{-1}(\mathbf{I_n} + \mathbf{S})$ is a rotation matrix.

Solution:

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

   remarks : The rotation matrix has geometric significance with det(R) = 1. So for this question, if your answer is $\mathbf{R} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$ or $\mathbf{R} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$, $\mathbf{R} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$, you will get 2 points

2)
$$\mathbf{R_1R_2} = \begin{bmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \end{bmatrix}\begin{bmatrix} \cos\alpha_2 & -\sin\alpha_2 \\ \sin\alpha_2 & \cos\alpha_2 \end{bmatrix} = \begin{bmatrix} \cos(\alpha_1+\alpha_2) & -\sin(\alpha_1+\alpha_2) \\ \sin(\alpha_1+\alpha_2) & \cos(\alpha_1+\alpha_2) \end{bmatrix}$$

   So $\mathbf{R_1R_2}$ is a rotation matrix which rotates the plane through an angle of $\alpha_1+\alpha_2$.

   remarks :if you use the matrix $\mathbf{R} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$, the answer is $\alpha_1-\alpha_2$. If you don't write the derivative process, you will lose 2 points.

3)
$$\mathbf{S} = (\mathbf{R}-\mathbf{I})(\mathbf{R}+\mathbf{I})^{-1} = \mathbf{R}(\mathbf{R}+\mathbf{I})^{-1} - (\mathbf{R}+\mathbf{I})^{-1}$$

   and so

$$\begin{aligned} \mathbf{S}^T &= (\mathbf{R}^T+\mathbf{I})^{-1}\mathbf{R}^T - (\mathbf{R}^T+\mathbf{I})^{-1} \\ &= (\mathbf{R}^T+\mathbf{I})^{-1}\mathbf{R}^{-1} - (\mathbf{R}^T+\mathbf{R}\mathbf{R}^T)^{-1} \\ &= (\mathbf{R}(\mathbf{R}^T+\mathbf{I}))^{-1} - (\mathbf{R}^T+\mathbf{R}\mathbf{R}^T)^{-1} \text{ (5 points)} \\ &= (\mathbf{R}(\mathbf{R}^T+\mathbf{I}))^{-1} - ((\mathbf{I}+\mathbf{R})\mathbf{R}^{-1})^{-1} \\ &= (\mathbf{I}+\mathbf{R})^{-1} - \mathbf{R}(\mathbf{R}+\mathbf{I})^{-1} = -\mathbf{S} \end{aligned}$$

[2]$\mathbf{I_n}$ is the identity matrix of $n\times n$

4)

$$\begin{aligned}\mathbf{R}\mathbf{R}^T &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}+\mathbf{S})^T(\mathbf{I}-\mathbf{S})^{-T} \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})(\mathbf{I}+\mathbf{S})^{-1} \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}-\mathbf{S}-\mathbf{S^2})(\mathbf{I}+\mathbf{S})^{-1} \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S})(\mathbf{I}+\mathbf{S})(\mathbf{I}+\mathbf{S})^{-1} = \mathbf{I}\end{aligned}$$

(5 points)