SI231 - Matrix Computations, Fall 2020-21

Homework Set #1

Name: Major: EE

Student No.: E-mail:

I. UNDERSTANDING RANK, RANGE SPACE AND NULL SPACE

1) **Solution:**

a) 为了证明 $\mathcal{R}(A^T) \oplus \mathcal{N}(A) = \mathbb{R}^n$, 等价于证明 $\mathcal{R}(A^T) \bigcap \mathcal{N}(A) = 0 \; and \; \mathcal{R}(A^T) + \mathcal{N}(A) = \mathbb{R}^n$，

又因为$A^T \in \mathbb{R}^{m\times n}$, $\mathcal{R}(A^T) = \{y \in \mathbb{R}^n \mid y = Ax, x \in \mathbb{R}^n\}$, $\mathcal{N}(A) = \{y \in \mathbb{R}^n \mid Ay = 0\}$,

假设存在$\mathbf{y} \neq \mathbf{0}$,由$\mathcal{N}(A) = \{y \in \mathbb{R}^n \mid Ay = 0\}$得$rank(A) = n$,则只存在唯一x满足$y = Ax$,所以假设不成立,

又易知$\mathcal{R}(A^T)$和$\mathcal{N}(A)$均包含$\mathbf{0}$,所以$\mathcal{R}(A^T) \bigcap \mathcal{N}(A) = \mathbf{0}$

设$rank(A) = k$,则$rank(\mathcal{R}(A^T)) = k, dim(\mathcal{R}(A^T)) = k, dim(\mathcal{N}(A)) = n - k$,

所以，$dim(\mathcal{R}(A^T)) + dim(\mathcal{N}(A)) = n$,

即$\mathcal{R}(A^T) + \mathcal{N}(A) = \mathbb{R}^n$.

综上,$\mathcal{R}(A^T) \oplus \mathcal{N}(A) = \mathbb{R}^n$

b) The set is For matrices $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{m\times n}$, prove that $rank(\mathbf{A} + \mathbf{B}) \leq rank(\mathbf{A}) + rank(\mathbf{B})$

设$rank(\mathbf{A}) = p, rank(\mathbf{B}) = q$,将$\mathbf{A}, \mathbf{B}$按列分块得

$\mathbf{A} = [\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \cdots, \boldsymbol{\alpha_s}]$,$\mathbf{B} = [\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_s}]$

于是 $[\mathbf{A}, \mathbf{B}] = [\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \cdots, \boldsymbol{\alpha_s}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_s}]$,

$\mathbf{A} + \mathbf{B} = [\boldsymbol{\alpha_1} + \boldsymbol{\beta_1}, \boldsymbol{\alpha_2} + \boldsymbol{\beta_2}, \boldsymbol{\alpha_3} + \boldsymbol{\beta_3}, \cdots, \boldsymbol{\alpha_s} + \boldsymbol{\beta_s}]$,

因$\boldsymbol{\alpha_i} + \boldsymbol{\beta_i} (i = 1, 2, \cdots, s)$均可由向量组$\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \cdots, \boldsymbol{\alpha_s}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_s}$线性表出，故

$rank(\mathbf{A} + \mathbf{B}) \leq rank([\mathbf{A}, \mathbf{B}])$

又设$\mathbf{A}, \mathbf{B}$的列空间的极大线性无关组分别为$\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \cdots, \boldsymbol{\alpha_p}$和$\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_p}$，将$\mathbf{A}$的极大线性无关组$\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \cdots, \boldsymbol{\alpha_p}$扩充成$[\mathbf{A}, \mathbf{B}]$的极大线性无关组，

设为$\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \cdots, \boldsymbol{\alpha_p}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_w}$，显然$w \leq q$，故有

$rank([\mathbf{A}, \mathbf{B}]) = p + w \leq p + q = rank(\mathbf{A}) + rank(\mathbf{B})$

故$rank(\mathbf{A} + \mathbf{B}) \leq rank([\mathbf{A}, \mathbf{B}]) \leq rank(\mathbf{A}) + rank(\mathbf{B})$

c) 将$\mathbf{A}, \mathbf{AB}$按行分块为 $\begin{bmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \\ \vdots \\ \boldsymbol{\beta_n} \end{bmatrix}$, $\mathbf{AB} = \mathbf{C} = \begin{bmatrix} \boldsymbol{\gamma_1} \\ \boldsymbol{\gamma_2} \\ \vdots \\ \boldsymbol{\gamma_m} \end{bmatrix}$，于是

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \\ \vdots \\ \boldsymbol{\beta_n} \end{bmatrix} = \begin{bmatrix} a_{11}\boldsymbol{\beta_1} + a_{12}\boldsymbol{\beta_2} + \cdots + a_{1n}\boldsymbol{\beta_n} \\ a_{21}\boldsymbol{\beta_1} + a_{22}\boldsymbol{\beta_2} + \cdots + a_{2n}\boldsymbol{\beta_n} \\ \cdots \\ a_{m1}\boldsymbol{\beta_1} + a_{m2}\boldsymbol{\beta_2} + \cdots + a_{mn}\boldsymbol{\beta_n} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma_1} \\ \boldsymbol{\gamma_2} \\ \vdots \\ \boldsymbol{\gamma_m} \end{bmatrix}$$

所以$\mathbf{AB}$的行向量$\boldsymbol{\gamma_m}(i = 1, 2, 3, \cdots, m)$均可由$\mathbf{B}$的行向量线性表出，故

$rank(\mathbf{AB}) \leq rank(\mathbf{B})$.

同理可证$rank(\mathbf{AB}) \leq rank(\mathbf{A})$，故有$rank(\mathbf{AB}) \leq min\{rank(\mathbf{A}), rank(\mathbf{B})\}$.

如果$rank(\mathbf{AB}) = n$

则$n \leq rank(\mathbf{A}) \leq min(m, n), n \leq rank(\mathbf{B}) \leq min(n, q)$

所以$\mathbf{A}$列满秩，$\mathbf{B}$行满秩

d) 设$\mathbf{A}$的列空间为$\mathcal{R}(A) = \{\alpha \in \mathbb{R}^n \mid \alpha = Ax_1, x_1 \in \mathbb{R}^n\}$，设$\mathbf{B}$的列空间为$\mathcal{R}(B) = \{\beta \in \mathbb{R}^n \mid \beta = Ax_2, x_2 \in \mathbb{R}^p\}$，设$\mathbf{AB}$的列空间为$\mathcal{R}([A \mid B]) = \{\gamma \in \mathbb{R}^m \mid \gamma = [A \mid B]\, x_3, x_3 \in \mathbb{R}^{n+p}\}$，

易知，$\mathcal{R}([A \mid B])$可和$\mathcal{R}(A) + \mathcal{R}(B)$相互线性表示，

所以，$\mathcal{R}([A \mid B]) = \mathcal{R}(A) + \mathcal{R}(B)$

e) 设$\mathbf{A}$的列空间为$\mathcal{R}(A) = \{\alpha \in \mathbb{R}^n \mid \alpha = Ax_1, x_1 \in \mathbb{R}^n\}$，设$\mathbf{B}$的列空间为$\mathcal{R}(B) = \{\beta \in \mathbb{R}^n \mid \beta = Ax_2, x_2 \in \mathbb{R}^p\}$，设$\mathbf{AB}$的列空间为$\mathcal{R}([A \mid B]) = \{\gamma \in \mathbb{R}^m \mid \gamma = [A \mid B]\, x_3, x_3 \in \mathbb{R}^{n+p}\}$，

易知，$\mathcal{R}([A \mid B])$可和$\mathcal{R}(A) + \mathcal{R}(B)$相互线性表示，

所以，$\mathcal{R}([A \mid B]) = \mathcal{R}(A) + \mathcal{R}(B)$

II. UNDERSTANDING SPAN, SUBSPACE

1) **Solution:**

   a) We hvve to prove that $span(S) \subseteq \mathcal{M}$ and $\mathcal{M} \subseteq span(S)$

   if $\mathcal{V}$ is a subspace and $span(\mathcal{S}) \subseteq \mathcal{V}$ then $span(\mathcal{S}) \subseteq \mathcal{M}$.

   If $\mathbf{x} \in span(\mathcal{S})$ and $\mathcal{V}$ is any subspce containing $\mathcal{S}$, then $\mathcal{V}$ contains $\mathbf{x}$ (because $\mathbf{x}$ is a linear combinatinn of elements of $\mathcal{S}$). Hence x belongs to the intersection of all such $\mathcal{V}$,which is $\mathcal{M}$. Thus $span(\mathcal{S}) \subseteq \mathcal{S}$. $\mathcal{M} \subseteq span(\mathcal{S})$ follows from the fact that $span(\mathcal{S})$ is itself one of the subspaces containing $\mathcal{S}$

III. BASIS, DIMENSION AND PROJECTION

1) **Problem 1. Solution:**

   a) The dimension of the space of polynomials having degree n is $n + 1$

   b) The dimension of the space of $n \times n$ symmetric matrices is $n(n+1)/2$

2) **Problem 2. Solution:**

   a) rotation matrix in $\mathbb{R}^{2\times2}$ is $\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$ or $\begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}$

b) To rotate x by $\frac{7\pi}{12}$ in anti-clockwise direction,the $\mathbf{R} = \begin{bmatrix} cos\frac{7\pi}{12} & -sin\frac{7\pi}{12} \\ sin\frac{7\pi}{12} & cos\frac{7\pi}{12} \end{bmatrix}$,so $\mathbf{Rx} = \begin{bmatrix} cos\frac{7\pi}{12} & -sin\frac{7\pi}{12} \\ sin\frac{7\pi}{12} & cos\frac{7\pi}{12} \end{bmatrix} \begin{bmatrix} cos\frac{\pi}{4} \\ sin\frac{\pi}{4} \end{bmatrix} =$

c) $\mathbf{Hx} = (1 - 2\mathbf{uu})^T\mathbf{x} = (1 - \mathbf{uu})^T\mathbf{x} - \mathbf{uu}^T\mathbf{x} = \mathbf{Qx} - (\mathbf{x} - \mathbf{Qx}) = \mathbf{Rx}$

so, $\mathbf{Hx}$ is a reflection of x with respect to $\mathcal{H}_u$

IV. DIRECT SUM

1) **Problem 1. Solution:**

a) suppose the column maximal linearly independent subset of $\mathcal{V}$ is $\mathcal{B} = \{\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_n}\}$,

because $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ and $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$,

so,I can suppose $\mathcal{B}_1 = \{\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \cdots, \boldsymbol{\beta_s}\}$ and $\mathcal{B}_2 = \{\boldsymbol{\beta_{s+1}}, \boldsymbol{\beta_{s+2}}, \boldsymbol{\beta_{s+3}}, \cdots, \boldsymbol{\beta_n}\}$

Then $dim(\mathcal{B}_1) + dim(\mathcal{B}_2) = n = dim(\mathcal{B}) = dim(\mathcal{V})$

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

2) **Problem 2. Solution:**

a) suppose the column maximal linearly independent subset of $\mathcal{V}$ is $\{\boldsymbol{\nu_1}, \boldsymbol{\nu_2}, \boldsymbol{\nu_3}, \cdots, \boldsymbol{\nu_n}\}$and $\mathcal{S} = \{\boldsymbol{\nu_1}, \boldsymbol{\nu_2}, \boldsymbol{\nu_3}, \cdots, \boldsymbol{\nu_d}\}$ $(d < n)$

$\mathcal{T} = \{\boldsymbol{\nu_{d+1}}, \boldsymbol{\nu_{d+2}}, \boldsymbol{\nu_{d+3}}, \cdots, \boldsymbol{\nu_n}\}$

Then $dim(\mathcal{S}) + dim(\mathcal{T}) = n = dim(\mathcal{V})$

so $\mathcal{V} = span(\mathcal{S}) \oplus span(\mathcal{T})$

V. UNDERSTANDING THE MATRIX NORM

1) **Solution:**

a) supppose $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ and we have $\sum_{i=1}^{n} x_i = 1$

so $\mathbf{Ax} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$

Then $\|\mathbf{A}\|_1 = \underset{\|\mathbf{x}\|_1=1}{max} \|\mathbf{Ax}\|_1 = max\{(\sum_{i=1}^{m} a_{i1})x_1 + (\sum_{i=1}^{m} a_{i2})x_2 + \cdots + (\sum_{i=1}^{m} a_{in})x_n\}$

and we know $\sum_{i=1}^{n} x_i = 1$ to make the formula above be maximum,we should find the maximum $\sum_{i=1}^{m} a_{ij}$,then let $x_j = 1$ and $\underset{i \neq j}{x_i} = 0$

At last $\|\mathbf{A}\|_1 = \underset{\|\mathbf{x}\|_1=1}{max} \|\mathbf{Ax}\|_1 = the\ largest\ absolute\ column\ sum$

b) supppose $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ and we have $\underset{i=1,2,\cdots,n}{max} \mid x_i \mid = 1$

so $\mathbf{Ax} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix}$

Then $\|\mathbf{A}\|_\infty = \underset{\|\mathbf{x}\|_\infty=1}{max} \|\mathbf{Ax}\|_\infty = \underset{i}{max}\{\sum_{j=1}^{n} a_{ij}x_j\}$

and we know $\underset{i=1,2,\cdots,n}{max} \mid x_i \mid = 1$ to make the formula above be maximum,

we should let $x_1 = x_2 = \cdots = x_n = 1$

so the maximum $\underset{i}{max}\{\sum_{j=1}^{n} a_{ij}x_j\} = \underset{i}{max}\{\sum_{j=1}^{n} a_{ij}\} = the\ largest\ absolute\ column\ sum$

VI. UNDERSTANDING THE HÖLDER INEQUALITY

1) **Solution:**

a) first if $\beta = 0$ $\alpha^\lambda \beta^{1-\lambda} \leq \lambda\alpha + (1-\lambda)\beta \;\Leftrightarrow\; \lambda\alpha \geq 0$,this is clearly established

then if $\beta > 0$ $\alpha^\lambda \beta^{1-\lambda} \leq \lambda\alpha + (1-\lambda)\beta \;\Leftrightarrow\; \alpha^\lambda \beta^{-\lambda} \leq \lambda\frac{\alpha}{\beta} + (1-\lambda) \;\Leftrightarrow\; (\frac{\alpha}{\beta})^\lambda \leq \lambda\frac{\alpha}{\beta} + (1-\lambda) \;\Leftrightarrow$

$t^\lambda \leq \lambda t + (1-\lambda)$ where $t = \frac{\alpha}{\beta} \geq 0$

so the question is equal to prove $f(t) \geq 0$

Derivative of a function $f(t)$ is $f^{'}(t) = \lambda - \lambda t^{\lambda-1},\ 0 < \lambda < 1$

it's easy to find when $t \in [0, 1), f^{'}(t) \geq 0$ and when $t > 1, f^{'}(t) \leq 0$

we also find $f(1) = 0$,so $f(t) \geq 0$

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

b) first we prove when $x > 0, y > 0, p > 0, q > 0, \frac{1}{p} + \frac{1}{q} = 1$,then $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$

the above is equal to prove $ln(xy) \leq ln(\frac{x^p}{p} + \frac{y^q}{q})$

because $x > 0$,suppose $f(x) = ln(x) \Rightarrow f^{''}(x) = -\frac{1}{x^2} < 0 \Rightarrow f(x\Psi$'s image is raised,then use concavity and convexity definition,we have $f[\lambda x_1 + (1-\lambda)y_1] \geq \lambda f(x_1) + (1-\lambda)f(y_1)$.

in the above formula,command $\lambda = \frac{a}{p}, x_1 = x^p, y_1 = y^q$,then $1 - \lambda = 1 - \frac{1}{p} = \frac{1}{q}$,so we get

$ln(\frac{x^p}{p} + \frac{y^q}{q}) \geq \frac{1}{p}f(x^p) + \frac{1}{q}f(y^q) = ln(xy)$

so we get $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$

so $\sum_{i=1}^{n} \mid \hat{x}_i\hat{y}_i \mid \leq \sum_{i=1}^{n}(\frac{1}{p} \mid \hat{x}_i \mid^p + \frac{1}{q} \mid \hat{y}_i \mid^q) = \frac{1}{p}\sum_{i=1}^{n} \mid \hat{x}_i \mid^p + \frac{1}{q}\sum_{i=1}^{n} \mid \hat{y}_i \mid^q$

and bucasue $\hat{x}_i = \frac{x_i}{(\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}}, \hat{y}_i = \frac{y_i}{(\sum_{i=1}^{n}|y_i|^q)^{\frac{1}{q}}}$

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

c) with the above results,because $\mathbf{x} = \|\mathbf{x}\|_p \hat{\mathbf{x}}, \quad \mathbf{y} = \|\mathbf{y}\|_q \hat{\mathbf{y}}$

so $\mid \mathbf{x^T y} \mid = \|\mathbf{x}\|_p \|\mathbf{y}\|_q \mid \hat{\mathbf{x}}^T \hat{\mathbf{y}} \mid = \|\mathbf{x}\|_p \|\mathbf{y}\|_q \sum_{i=1}^n \mid \hat{x}_i \hat{y}_i \mid \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q (\frac{1}{p} \sum_{i=1}^n \mid \hat{x}_i \mid^p + \frac{1}{q} \sum_{i=1}^n \mid \hat{y}_i \mid^q) = \|\mathbf{x}\|_p \|\mathbf{y}\|_q$

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

d) Let $\mathbf{u} = (u_1, u_1, \cdots, u_n)$ with $u_i = \mid x_i + y_i \mid^{p-1}$.Since $q(p-1) = p$ and $\frac{p}{q} = p - 1$,we find $\|\mathbf{u}\|_q = (\sum_{i=1}^n |x_i + y_i|^{q(p-1)})^{\frac{1}{q}} = (\sum_{i=1}^n |x_i + y_i|^p)^{\frac{1}{q}} = \|\mathbf{x} + \mathbf{y}\|_p^{\frac{p}{q}} = \|\mathbf{x} + \mathbf{y}\|_p^{p-1}$

Using this and the Holder inequality we obtain

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

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

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