Problem 1
[30 marks] Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, $R(.)$ denotes the range of the matrix, $N(.)$ denotes the null space of a given amtrix, $dim(.)$ denotes the dimension of a vector space, then prove the following:

(a) $dim[R(\mathbf{AB})] \leq dim[R(\mathbf{A})]$

(b) If the matrix $\mathbf{B}$ is non-singular then $dim[R(\mathbf{AB})] = dim[R(\mathbf{A})]$.

(c) $dim[N(AB)] \leq dim[N(\mathbf{A})] + dim[N(\mathbf{B})]$

(d) $dim[R(\mathbf{A})] + dim[N(\mathbf{A})] = n$

(e) $rank(\mathbf{A}) + rank(\mathbf{B}) - n \leq rank(\mathbf{AB}) \leq min(rank(\mathbf{A}), rank(\mathbf{B}))$

(f) Given a vector $\mathbf{u} \in \mathbb{R}^n$, $rank(\mathbf{uu}^T)$ is 1.

(g) Row rank always equals column rank.

Solution:

(a) Rank of a matrix $\mathbf{M}$ is the same as dimension of the range $R(\mathbf{M})$ of the matrix. Therefore

$$rank(\mathbf{AB}) = dim[R(\mathbf{AB})]$$
$$rank(\mathbf{A}) = dim[R(\mathbf{A})]$$

$\therefore$ the problem boils down to prove that $rank(\mathbf{A}) \leq rank(\mathbf{AB})$. Also, if $\mathbf{V}$ is a subset of vector space $\mathbf{W}$, one can write:

$$dim[\mathbf{V}] \leq dim[\mathbf{W}]$$

Range of a matrix $\mathbf{M}$ is the span of its column vectors. That means any vector $\mathbf{y}$ belonging to $R(\mathbf{M})$ can be written as a linear combination of the column vectors of $\mathbf{M}$.
If $\mathbf{y}$ is a vector belonging to $R(\mathbf{AB})$, then $\mathbf{y}$ can be written as $\mathbf{y} = (\mathbf{AB})\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^p$. Take

$$\mathbf{z} = \mathbf{Bx}, \ \mathbf{z} \in \mathbb{R}^{(n)}$$

Then, we have $\mathbf{y} = \mathbf{A(Bx)} = \mathbf{Az}$.
But any vector $\mathbf{y'} \in R(\mathbf{A})$ also means $\mathbf{y'} \in \mathbb{R}^{n \times 1}$. That means, $\mathbf{y} \in R(\mathbf{A})$ as well.
If $\mathbf{y} \in R(\mathbf{AB}) \implies \mathbf{y} \in R(\mathbf{A})$, it means that $R(\mathbf{AB}) \subset R(\mathbf{A})$.
Then it follows from above that

$$rank(\mathbf{AB}) \leq rank(\mathbf{A})$$
$$dim[R(\mathbf{AB})] \leq dim[R(\mathbf{A})]$$

(b) From **(a)**, the following result was obtained:

$$rank(\mathbf{AB}) \leq rank(\mathbf{A})$$

Given the matrix $\mathbf{B}$ is non-singular. Thus it is invertible and $\mathbf{B}^{-1}$ exists.
Replacing $\mathbf{AB}$ and $\mathbf{A}$ in the above equation with $\mathbf{(AB)B}^{-1}$ and $\mathbf{(AB)}$ respectively, one can write:

$$rank(\mathbf{(AB)B}^{-1}) \leq rank(\mathbf{(AB)})$$

But $\mathbf{(AB)B}^{-1} = \mathbf{A(BB}^{-1}\mathbf{)} = \mathbf{AI} = \mathbf{A}$. Therefore

$$rank(\mathbf{A}) \leq rank(\mathbf{AB})$$

But in **(a)**, it was already proven that $rank((\mathbf{AB})) \leq rank(\mathbf{A})$. This implies,

$$
\begin{aligned}
rank(\mathbf{AB}) &= rank(\mathbf{A}) \\
dim[R(\mathbf{AB})] &= dim[R(\mathbf{A})]
\end{aligned}
$$

(c) Using the definitions of $R(\mathbf{A}), R(\mathbf{B}), N(\mathbf{A}), N(\mathbf{B})$:

$$
\begin{aligned}
R(\mathbf{A}) &= \{\mathbf{y} \mid \mathbf{Ax} = \mathbf{y},\ \mathbf{A} \in \mathbb{R}^{(m\times n)},\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{y} \in \mathbb{R}^m\} \\
R(\mathbf{B}) &= \{\mathbf{y} \mid \mathbf{Bx} = \mathbf{y},\ \mathbf{B} \in \mathbb{R}^{(m\times n)},\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{y} \in \mathbb{R}^m\} \\
N(\mathbf{A}) &= \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{0},\ \mathbf{A} \in \mathbb{R}^{(m\times n)},\ \mathbf{x} \in \mathbb{R}^n\} \\
N(\mathbf{B}) &= \{\mathbf{x} \mid \mathbf{Bx} = \mathbf{0},\ \mathbf{B} \in \mathbb{R}^{(n\times p)},\ \mathbf{x} \in \mathbb{R}^p\}
\end{aligned}
$$

Using rank-nullity theorem (proved in **1.c**), one may write:

$$
\begin{aligned}
&dim(R(A)) + dim(N(A)) = n \\
&dim(R(B)) + dim(N(B)) = p \\
&dim(R(AB)) + dim(N(AB)) = p
\end{aligned}
$$

as $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ and $AB \in \mathbb{R}^{m\times p}$. Then we have:

$$
\begin{aligned}
&dim(N(A)) = n - dim(R(A)) \\
&dim(N(B)) = p - dim(R(B)) \\
&dim(N(AB)) = p - dim(R(AB))
\end{aligned}
$$

Substituting this into the given equation,

$$
\begin{aligned}
&dim[N(AB)] \leq dim[N(A)] + dim[N(B)] \\
&\implies p - dim[R(AB)] \leq n - dim[R(A)] + p - dim[R(B)] \\
&\implies -dim[R(AB)] \leq -dim[R(A)] - dim[R(B)] + n \\
&\implies dim[R(A)] + dim[R(B)] - n \leq dim[R(AB)] \\
&\implies rank(A) + rank(B) - n \leq rank(AB)
\end{aligned}
$$

So the given problem boils down to proving this:

$$rank(A) + rank(B) - n \leq rank(AB)$$

But we know that $rank(AB) = rank(B) - dim(N(A) \cap R(B))$
Note that $N(A) \cap R(B) \subseteq N(A)$
So, $dim(N(A) \cap R(B)) \leq dim(N(A))$
$dim(N(A)) \cap R(B)) \leq n - rank(A)$
But $rank(AB) = rank(B) - dim(N(A) \cap R(B))$
$\implies rank(AB) \geq rank(A) + rank(B) - n$

Thus it follows that

$$\underline{\underline{dim[N(AB)] \leq dim[N(A)] + dim[N(B)]}}$$

(d)

(e) To prove:

$$rank(A) + rank(B) - n \leq rank(AB) \leq minrank(A), rank(B)$$

Each column of AB is a combination of the columns of A, which implies that $R(AB) \subseteq R(A)$.
Each row of AB is a combination the rows of B, which means rowspace(AB) $\subseteq$ rowspace(B), but the dimension of the rowspace = dimension of the column space = rank, so that rank(AB) $\leq$ rank(B). Therefore,

$$rank(AB) \leq minrank(A), rank(B)$$

To show that $rank(A) + rank(B) - n \leq rank(AB)$, let

$$r_B = rank(B)$$
$$r_A = rank(A)$$

where $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$.// Now, let $\{v_1, v_2, \ldots, v_{r_B}\}$ be a basis set of $R(B)$, and add $n - r_B$ linearly independent vectors $\{w_1, \ldots, w_{r_B}\}$ to this basis to span all of $\mathbb{R}^n$, $v_1, v_2, \ldots, v_n, w_1, \ldots, w_{n-r_B}$. Let

$$M = (v_1|v_2|\ldots|v_{r_B}) = (VW)$$

Suppose $x \in \mathbb{R}^n$, then $x = M\alpha$ for some $\alpha \in \mathbb{R}^n$.
The following are known:

$$R(A) = R(AM) = R([AV|AW])$$
$$R(AB) = R(AV)$$

Using first equation above, it can be observed that the number of linearly independent columns of A is less than or equal to the number of linearly indpendent columns of AV + the number of columns of AW, which means that:

$$rank(A) \leq rank(AV) + rank(AW)$$

Using second equation $R(AB) = R(AV)$, we can see that:

$$rank(AV) = rank(AB) \rightarrow rank(A) \leq rank(AB) + rank(AW)$$

yet, there are only $n - r_B$ columns of AW. Thus:

$$rank(AW) \leq n - r_B$$
$$rank(A) - rank(AB) \leq rank(AW) \leq n - r_B$$
$$r_A - (n - r_B) \leq r_{AB}$$

or $rank(A) + rank(B) - n \leq rank(AB)$

(f) Let the matrix formed from uu^T be M. Then $rank(M) = dim[R(M)]$. $R(M)$ is the same as column space of M. So all the possible linear combinations of column vectors of M falls into $R(M)$. But all the column vectors are scaled versions of u alone.

$$uu^T = \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_n \end{pmatrix} \times \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix}$$

$$= u1 \times \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_n \end{pmatrix} + u2 \times \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_n \end{pmatrix} + \cdots + un \times \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_n \end{pmatrix}$$

Since all the vectors in $R(A)$ can be spanned with just one vector alone, i.e., u, $dim[R(uu^T)] = 1$ or $rank(uu^T) = 1$.

(g) We write $A = (a_{ij})$ and let $\mathbf{A}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{mi} \end{bmatrix}$ be the i-th column vector of A for $i = 1, 2, \ldots, n$. Also let $\mathbf{B}_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \cdots \\ a_{in} \end{bmatrix}$ be the ith column vector of A^T for $i = 1, 2, \ldots, m$, that is B_i is the transpose of the ith row vector of A. Suppose that $rank(A) = dim(R(A^T)) = k$ and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a basis for the range $R(A^T)$. We write $\mathbf{v}_i = \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{in} \end{bmatrix}$ for $i = 1, 2, \ldots, k$. Then each column vector $\mathbf{B}_i$ of A^T is a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$. Thus we have

$$\begin{aligned}
\mathbf{B}_1 &= c_{11}\mathbf{v}_1 + \cdots + c_{1k}\mathbf{v}_k \\
\mathbf{B}_2 &= c_{21}\mathbf{v}_1 + \cdots + c_{2k}\mathbf{v}_k \\
&\vdots \\
\mathbf{B}_m &= c_{m1}\mathbf{v}_1 + \cdots + c_{mk}\mathbf{v}_k.
\end{aligned}$$

More explicitly we have,

$$\begin{bmatrix} a_{11} \\ a_{12} \\ \cdots \\ a_{1n} \end{bmatrix} = c_{11} \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix} + \cdots + c_{1k} \begin{bmatrix} v_{k1} \\ v_{k2} \\ \vdots \\ v_{kn} \end{bmatrix}$$

$$\begin{bmatrix} a_{21} \\ a_{22} \\ \cdots \\ a_{2n} \end{bmatrix} = c_{21} \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix} + \cdots + c_{2k} \begin{bmatrix} v_{k1} \\ v_{k2} \\ \vdots \\ v_{kn} \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} a_{m1} \\ a_{m2} \\ \cdots \\ a_{mn} \end{bmatrix} = c_{m1} \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix} + \cdots + c_{mk} \begin{bmatrix} v_{k1} \\ v_{k2} \\ \vdots \\ v_{kn} \end{bmatrix}$$

Now, we look at the i-th entries for the above vectors and we have

$$\begin{aligned}
a_{1i} &= c_{11}v_{1i} + \cdots + c_{1k}v_{ki} \\
a_{2i} &= c_{21}v_{1i} + \cdots + c_{2k}v_{ki} \\
&\vdots \\
a_{mi} &= c_{m1}v_{1i} + \cdots + c_{mk}v_{ki} \\
&.
\end{aligned}$$

We rewrite these as a vector equality and obtain

$$\begin{aligned}\mathbf{A}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{mi} \end{bmatrix} &= v_{1i} \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + \cdots + v_{ki} \begin{bmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{bmatrix} \\ &= v_{1i}\mathbf{c}_1 + \cdots + v_{ki}\mathbf{c}_k,\end{aligned}$$

where we put $c_j = \begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{mj} \end{bmatrix}$ for $j = 1, 2, \ldots, k$.

This shows that any column vector $\mathbf{A}_i$ of A is a linear combination of vectors $\mathbf{c}_1, \ldots, \mathbf{c}_k$. Therefore we have $R(A) = \{\mathbf{A}_1, \ldots, \mathbf{A}_n\} \subset \{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$. Since the dimension of a subspace is smaller than or equal to the dimension of a vector space containing it, we have $rank(A) = dim(R(A)) \leq dim(\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}) \leq k$. Hence we obtain $rank(A) \leq rank(A^T)$.

To achieve the opposite inequality, we repeat this argument using A^T, and we obtain

$rank(A^T) \leq rank((A^T)^T) = rank(A)$ since we have $(A^T)^T = A$.

Therefore, required equality is proved: $rank(A) = rank(A^T)$. or row rank equals column rank.

Problem 2
[10 marks] Suppose there always exists a set of real coefficients $c_1, c_2, c_3, \ldots, c_{10}$ for any set of real numbers $d_1, d_2, d_3, \ldots, d_{10}$

$$\sum_{j=1}^{10} c_j f_j(i) = d_i$$

for $i \in 1, 2, \ldots 10$, where $f_1, f_2, f_3, \ldots f_{10}$ are a set of functions defined on the interval [1, 10].

(a) Use the concepts discussed in class to show that $d_1, d_2, d_3, \ldots d_{10}$ determine $c_1, c_2, c_3, \ldots c_{10}$ uniquely.

(b) Let $\mathbf{A}$ be a 10×10 matrix representing the linear mapping from data $d_1, d_2, \ldots, d_10$ to coefficients $c_1, c_2, c_3, \ldots, c_{10}$. What is the i, jth entry of $\mathbf{A}^{-1}$?

Solution:

Problem 3
[15 marks] A matrix $\mathbf{S}$ is said to be symmetric if $\mathbf{S}^T = \mathbf{S}$ and skew-symmetric if $\mathbf{S}^T = -S$. Now verify the following:

(a) The matrix $\mathbf{Q} = \mathbf{(I - S)^{-1}(I + S)}$ is an orthogonal matrix for any skew-symmetric matrix $\mathbf{S}$.

(b) Note that a symmetric matrix $\mathbf{A} \in \mathbb{R}^{m\times m}$ can be decomposed as $\mathbf{QDQ}^T$ where Q is an orthogonal matrix and $\mathbf{D}$ is a diagonal matrix. Using this result, show that $\mathbf{u}^T\mathbf{Au} = 0 \ \forall \ \mathbf{u} \in \mathbb{R}^m$, if and only if $\mathbf{A} = 0$.

(c) Show that $\mathbf{u}^T\mathbf{Su} = 0 \ \forall \ \mathbf{u} \in \mathbb{R}^m$ if and only if $\mathbf{S}$ is a skew-symmetric matrix.

Solution:

(a)

$$
\begin{aligned}
Q &= (I-S)^{-1}(I+S) \\
Q^T &= ((I-S)^{-1}(I+S))^T \\
&= (I+S)^T((I-S)^{-1})^T \\
&= (I+S^T)((I-S)^T)^{-1} \\
&= (I-S)(I+S)^{-1} \\
Q^TQ &= (I-S)((I+S)^{-1}(I+S)((I-S)^{-1} \\
&= (I-S)I(I-S)^{-1} \\
&= (I-S)(I-S)^{-1} \\
&= I
\end{aligned}
$$

For invertible Q,

$$
\begin{aligned}
&Q^TQQ^{-1} = IQ^{-1} \\
&\underline{\underline{Q^T = Q^{-1}}}
\end{aligned}
$$

Thus we've shown that $Q^TQ = I$. Similarly it can be shown that $QQ^T = I$. That means Q is an orthogonal matrix for any skew-symmetric matrix S.

(b) If $A = 0$, then it follows trivially that $u^TAu = 0$.
If A be a symmetric matrix, then,

$$
\begin{aligned}
u^TAu &= u^TQDQ^Tu \\
&= (Q^Tu)^TD(Q^Tu) \\
&= \sum_{i=1}^{m} d_{i,i}(q_i^Tu)^2
\end{aligned}
$$

The above sum is zero only if all the $d_{i,i}$ were 0. Then $QDQ^T = 0$ and in turn, $A = 0$.

(c) To prove: $u^TSu = 0 \forall u \in \mathbb{R}^m$ if and only if S is a skew-symmetric matrix.
Let $u^TSu = 0 \forall u \in \mathbb{R}^m$.
Then,

$$
\begin{aligned}
&u^TSu = \sum_{i}^{m}\sum_{j}^{m} u_js_{i,j}u_i = 0 \forall u \\
&\implies u_1^2s_{1,1} + u_1s_{1,2}u_2 + \cdots = 0 \\
&\implies (u_1^2s_{1,1} + u_2^2s_{2,2} + u_3^2s_{3,3} + \cdots + u_m^2s_{m,m}) + (u_1u_2s_{1,2} + u_2u_1s_{2,1} + \dots) = 0
\end{aligned}
$$

It is given that $u^T S u = 0$. Inorder to obtain this, all $s_{i,i}$ must be zero as well as $s_{i,j} = -s_{j,i}$. Thus S is skew-symmetric.
Now consider that S is skew-symmetric. Then,
since $u^T S u$ is a scalar, $u^T S u = (u^T S u)^T$

$$\begin{aligned}
u^T S u &= (u^T S u)^T \\
&= u^T S^T (u^T)^T \\
&= u^T S^T u \\
&= -u^T S u \\
\implies u^T S u &= -u^T S u \\
\implies 2u^T S u &= 0 \\
u^T S u &= 0
\end{aligned}$$

Thus it is proved that S being skew-symmetric implies $u^T S u = 0$. This proves that $u^T S u = 0 \forall u \in \mathbb{R}^m$, if and only if S is a skew-symmetric matrix.

Problem 4
[**35 marks** If $\mathbf{x} \in \mathbb{R}^{m \times n}$, then show the following:

(a) $||\mathbf{x}||_\infty \leq ||\mathbf{x}||_2$

(b) $||\mathbf{x}||_2 \leq \sqrt{m}||\mathbf{x}||_\infty$

(c) $||\mathbf{A}||_\infty \leq \sqrt{n}||\mathbf{A}||_2$

(d) $||\mathbf{A}||_2 \leq \sqrt{m}||\mathbf{A}||_\infty$

(e) $||\mathbf{A}||_F \leq \sqrt{tr(\mathbf{A}^T\mathbf{A})}$

(f) $\frac{1}{\sqrt{m}}||\mathbf{A}||_1 \leq ||\mathbf{A}||_2 \leq \sqrt{n}||\mathbf{A}||_1$

(g) $||\mathbf{A}||_2 \leq \sqrt{||\mathbf{A}||_1||\mathbf{A}||_\infty}$

Give an example of a non-zero vector or matrix for which equality is achieved in the above inequalities.

Solution:

(a) Given,

$$||x||_\infty = max_{1\leq i\leq m}|x_i|$$
$$||x||_2 = (\sum_{i=1}^{m}|x_i|^2)^{\frac{1}{2}}$$
$$= \sqrt{x^Tx}$$

To prove: $||x||_\infty \leq ||x||_2$
Let $max_{1\leq i\leq m}||x_i|| = x_k$, where k is a natural number between 1 and m included.

$$|x_k| \leq \sum_{i=1}^{m}|x_i|^2)^{\frac{1}{2}}$$
$$(|x_k|^2) \leq \sum_{i=1}^{m}|x_i|^2)$$
$$|x_k|^2 \leq |x_1|^2 + |x_2|^2 + \cdots + |x_k|^2 + \ldots |x_m|^2$$
$$0 \leq \sum_{i=1, i\neq k}^{m}|x_i|^2$$

which is true if $x \in \mathbb{R}^m$ as $x_is \in \mathbb{R}$.
Thus the starting assumption that $||x||_\infty = max_{1\leq i\leq m}|x_i|$ is true.

(b) To prove:

$$||x||_2 \leq \sqrt{m}||x||_\infty$$

Let $max_{1\leq i\leq m}||x_i|| = x_k$, where k is a natural number between 1 and m included.
Assume the above equation is true.

$$||x||_2 \leq \sqrt{m}||x||_\infty$$
$$(\sum_{i=1}^{m}|x_i|^2)^{1/2} \leq \sqrt{m}|x_k|$$
$$\sum_{i=1}^{m}|x_i|^2 \leq m|x_k|^2$$

But since $x_k \geq x_i, 1 \leq i \leq m$,

$$\sum_{i=1}^{m} x_k^2 \geq \sum_{i=1}^{m} x_i^2$$

$$m x_k \geq \sum_{i=1}^{m} x_i^2$$

Thus it follows that

$$\underline{\underline{||x||_2 \leq \sqrt{m}||x||_\infty}}$$

(c)

(d)

(e) To prove: $||A_F|| = \sqrt{tr(A^T A)}$.

$$||A_F||^2 = \sum_{j=1}^{n} ||a_j||_2^2$$

$$||a_j||_2 = \sum_{j=1}^{n}(\sum_{i=1}^{n} |a_{ji}|^2)$$

$$= \sum_{i,j} a_{ij}^2$$

$$= \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}^T a_{ji})$$

$$= \sum_{i=1}^{n}(A^T A)_{ii}$$

$$= tr(A^T A)$$

$$\implies ||A_F|| = \sqrt{tr(A^T A)}$$

(f)

(g) To prove:

$$||A||_2 \leq \sqrt{||A||_1||A||_\infty}$$

If $z \neq 0$ is such that $A^T A z = \mu^2 z$ with $\mu = ||A||_2$,
then $\mu^2||z||_1 = ||A^T A z||_1 \leq ||A^T||_1||A||_1||z||_1 = ||A^T||_\infty||A||_1||z||_1$

(h)

Problem 5
[10 marks] Induced matrix norm is defined as $||\mathbf{A}||^{(m,n)} = max||\mathbf{Ax}||^m$, where $\mathbf{x} \in \mathbb{R}^n$ and is a unit vector. $||.||$ corresponds to p-norm ($1 \leq p < \infty$). For this exercise, let us consider p to be natural number.

Using MATLAB/ Octave/ Python programming environment, create a matrix using $\mathbf{A} = randn(100, 2)$. Subsequently, create random unit vectors $\mathbf{x}$ using $temp = randn(2, 1)$ and normalize $\mathbf{x}$ using $x = \frac{temp}{norm(temp)}$. Checking for multiple random vectors $\mathbf{x}$ (use a loop and check for about 1000 random vectors $\mathbf{x}$) using $norm_of_Ax = norm(\mathbf{Ax}, p)$ for $p = 1, 2, 3, 4, 5, 6, \infty$. What is the maximum value of p-norm for the vector $\mathbf{Ax}$? Now calculate p-norm of $\mathbf{A}$ using $norm_of_A = norm(\mathbf{A}, p)$ for $p = 1, 2, \infty$ within the same programming environment you used before. Verify the equality $||(A)||^{(m)}$ for $p = 1, 2, \infty$. Note that this equality is true for other values of p as well but you are restricting to $p = 1, 2, \infty$ in this exercise.

Solution:

Programming environment of choice was Python (3.9.6).

```python
from numpy.linalg import norm
from numpy.random import randn
import numpy
x = []
Ax_norms = [0,0,0,0,0,0]
A_norms = []
A = randn(100, 2)
for i in range(1000):
    temp = randn(2, 1)
    temp_normed = temp/norm(temp)
    x.append(temp_normed)
    # calculate matrix norms
    for p in range(6):
        norm_of_Ax = 0
        if p == 0:
            norm_of_Ax = norm(A.dot(x[i]), numpy.inf)
        else:
            norm_of_Ax = norm(A.dot(x[i]), p)
        if(Ax_norms[p] < norm_of_Ax):
            Ax_norms[p] = norm_of_Ax
    # calculate vector norms

for p in range(3):
    if p == 0:
        norm_of_A = norm(A, numpy.inf)
    else:
        norm_of_A = norm(A, p)
    A_norms.append(norm_of_A)

print("NORM \t\tA_norm Ax_norm")
for i in range(len(Ax_norms)):
    if i == 0:
        which_norm = "inf-norm"
    else:
        which_norm = "  " + str(i) + "-norm"
    print(which_norm + "\t" + str(round(A_norms[i],3)) + "\t" + str(round(Ax_norms[i],3)))
```

The following output was obtained:

```
    NORM  A_norm Ax_norm
inf-norm 4.843 3.635
  1-norm 93.035 93.035
  2-norm 11.358 11.358
```

Observation: 1-norms and 2-norms showed excellent matching with 1000 random unit vectors, but ∞-norms showed non-negligible difference. With larger sample size for unit vectors, favourable matching was noted for ∞-norm.