Solutions for Homework 1 Foundations of Computational Math 1 Fall 2017

Problem 1.1

This problem considers three basic vector norms: $\|.\|_1$,$\|.\|_2$,$\|.\|_\infty$.

1.1.a Prove that $\|.\|_1$ is a vector norm.

1.1.b Prove that $\|.\|_\infty$ is a vector norm.

1.1.c Consider $\|.\|_2$.

i Show that $\|.\|_2$ is definite.

ii Show that $\|.\|_2$ is homogeneous.

iii Show that for $\|.\|_2$ the triangle inequality follows from the Cauchy inequality $|x^H y| \leq \|x\|_2 \|y\|_2$.

iv Assume you have two vectors x and y such that $\|x\|_2 = \|y\|_2 = 1$ and $x^H y = |x^H y|$, prove the Cauchy inequality holds for x and y.

v Assume you have two arbitrary vectors $\tilde{x}$ and $\tilde{y}$. Show that there exists x and y that satisfy the conditions of part (iv) and $\tilde{x} = \alpha x$ and $\tilde{y} = \beta y$ where α and β are scalars.

vi Show the Cauchy inequality holds for two arbitrary vectors $\tilde{x}$ and $\tilde{y}$.

Solution: We consider each norm in turn.

One Norm: If $x \neq 0$ then there exists an element $\xi_j \neq 0$. Therefore,

$$\begin{aligned} \|x\|_1 &= \Sigma_i |\xi_i| \\ &\geq |\xi_j| \\ &\geq 0 \end{aligned}$$

and $x = 0$ implies that $\|x\|_1 = 0$. Since all terms in the sum are nonnegative the only way the sum can be 0 is if all terms are 0 which implies all $\xi_i = 0$. It follows that $\|x\|_1 = 0$ implies $x = 0$. Therefore $\|x\|_1$ is definite.

We have

$$\begin{aligned} \|\alpha x\|_1 &= \Sigma_i |\alpha \xi_i| \\ &= \Sigma_i |\alpha||\xi_i| \\ &= |\alpha| \Sigma_i |\xi_i| \\ &= |\alpha| \|x\|_1 \end{aligned}$$

and therefore $\|x\|_1$ is homogeneous.

We have, given the triangle inequality for magnitude on $\mathbb{R}$ and $\mathbb{C}$,

$$\begin{aligned}\|x+y\|_1 &= \Sigma_i|\xi_i+\eta_i| \\ &\leq \Sigma_i(|\xi_i|+|\eta_i|) \\ &= \Sigma_i|\xi_i|+\Sigma_i|\eta_i| \\ &= \|x\|_1+\|y\|_1\end{aligned}$$

and therefore $\|x\|_1$ satisfies the triangle inequality.

Max Norm: If $x \neq 0$ then there exists an element $\xi_j \neq 0$. Therefore,

$$\begin{aligned}\|x\|_\infty &= \max_i|\xi_i| \\ &\geq |\xi_j| \\ &\geq 0\end{aligned}$$

Therefore $\|x\|_\infty$ is definite.

We have

$$\begin{aligned}\|\alpha x\|_\infty &= \max_i|\alpha\xi_i| \\ &= \max_i(|\alpha||\xi_i|) \\ &= |\alpha|\max_i|\xi_i| \\ &= |\alpha|\|x\|_\infty\end{aligned}$$

and therefore $\|x\|_\infty$ is homogeneous.

We have

$$\begin{aligned}\|x+y\|_\infty &= \max_i|\xi_i+\eta_i| \\ &\leq \max_i(|\xi_i|+|\eta_i|) \\ &\leq \max_i|\xi_i|+\max_i|\eta_i| \\ &= \|x\|_\infty+\|y\|_\infty\end{aligned}$$

and therefore $\|x\|_\infty$ satisfies the triangle inequality.

Two Norm: If $x \neq 0$ then there exists an element $\xi_j \neq 0$. Therefore,

$$\begin{aligned}\|x\|_2^2 &= \Sigma_i|\xi_i|^2 \\ &\geq |\xi_j|^2 \\ &> 0\end{aligned}$$

Therefore $\|x\|_2$ is definite.

We have

$$\begin{aligned}\|\alpha x\|_2^2 &= \Sigma_i|\alpha\xi_i|^2\\ &= \Sigma_i(|\alpha|^2|\xi_i|^2)\\ &= |\alpha|^2\Sigma_i|\xi_i|^2\\ &= |\alpha|^2\|x\|_2^2\end{aligned}$$

and therefore $\|x\|_2$ is homogeneous.

The triangle inequality follows from the Cauchy inequality for the two norm:

$$|x^Hy| \leq \|x\|_2\|y\|_2$$

as follows

$$\begin{aligned}\|x+y\|_2^2 &= x^Hx + y^Hy + 2\mathcal{R}e(x^Hy)\\ &= \|x\|_2^2 + \|y\|_2^2 + 2\mathcal{R}e(x^Hy)\\ &\leq \|x\|_2^2 + \|y\|_2^2 + 2|x^Hy|\\ &\leq \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2\|y\|_2\\ &= (\|x\|_2 + \|y\|_2)^2.\end{aligned}$$

So the true problem is to prove the Cauchy inequality. To do so assume that we have two vectors x and y such that $\|x\|_2 = \|y\|_2 = 1$ and $x^Hy = |x^Hy|$. For any two such vectors we have

$$\begin{aligned}\|x-y\|_2^2 &= (x-y)^H(x-y)\\ &= 2 - 2x^Hy = 2(1-|x^Hy|)\\ &\geq 0\end{aligned}$$

Therefore $|x^Hy| \leq 1 = \|x\|_2\|y\|_2$.

To generalize to any two nonzero vectors $\tilde{x}$ and $\tilde{y}$ note that there must exist complex scalars α and β such that $\tilde{x} = \alpha x$ and $\tilde{y} = \beta y$ where x and y satisfy the conditions above (see Lemma below). We have

$$\begin{aligned}|\tilde{x}^H\tilde{y}| &= |\alpha x^Hy\beta|\\ &= |\alpha\beta||x^Hy|\\ &\leq |\alpha\beta|\|x\|_2\|y\|_2\\ &= |\alpha|\|x\|_2|\beta|\|y\|_2\\ &= \|\alpha x\|_2\|\beta y\|_2\\ &= \|\tilde{x}\|_2\|\tilde{y}\|_2\end{aligned}$$

Lemma. *Let $\tilde{x} \in \mathbb{C}^n$ and $\tilde{y} \in \mathbb{C}^n$ be such that $\tilde{x} \neq 0$ and $\tilde{y} \neq 0$. There exists $\alpha \in \mathbb{C}$, $\beta \in \mathbb{C}$, $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ such that*

$$\begin{aligned} \|x\|_2 =& \|y\|_2 = 1 & \qquad |x^H y| =& x^H y \\ \tilde{x} =& \alpha x & \qquad \tilde{y} =& \beta y \end{aligned}$$

Proof. Suppose $\tilde{x}^H \tilde{y} = \gamma e^{i\phi}$ where $\gamma \in \mathbb{R}$ and $\gamma > 0$. Let $\phi_1 \in \mathbb{R}$ and $\phi_2 \in \mathbb{R}$ be such that $\phi = \phi_1 + \phi_2$. The scalars α and β can be set as follows:

$$\alpha = \|\tilde{x}\| e^{-i\phi_1} \qquad\qquad \beta = \|\tilde{y}\| e^{i\phi_2}$$

and note that $|\alpha| = \|\tilde{x}\|$ and $|\beta| = \|\tilde{y}\|$.

Taking $\tilde{x} = \alpha x$ and $\tilde{y} = \beta y$ implies that

$$\|x\|_2^2 = x^H x = \frac{1}{|\alpha|^2} \tilde{x}^H \tilde{x} = \frac{1}{\|\tilde{x}\|_2^2} \tilde{x}^H \tilde{x} = 1$$

$$\|y\|_2^2 = y^H y = \frac{1}{|\beta|^2} \tilde{y}^H \tilde{y} = \frac{1}{\|\tilde{y}\|_2^2} \tilde{y}^H \tilde{y} = 1$$

So $\|x\|_2 = \|y\|_2 = 1$ and we have that

$$\gamma e^{i\phi} = \tilde{x}^H \tilde{y} = \bar{\alpha} \beta x^H y = e^{i\phi_1} e^{i\phi_2} \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y = e^{i(\phi_1 + \phi_2)} \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y = e^{i\phi} \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y$$

$$\therefore x^H y = \frac{\gamma}{\|\tilde{y}\|_2 \|\tilde{x}\|_2}$$

Since γ, $\|\tilde{y}\|_2$ and $\|\tilde{x}\|_2$ are real and positive it follows that $x^H y$ is also real and positive. Therefore, $x^H y = |x^H y|$ as desired. $\square$

Problem 1.2

Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a linear function, i.e.,

$$F(\alpha x + \beta y) = \alpha F(x) + \beta F(y)$$

.

1.2.a. Suppose you are given a routine that returns $F(x)$ given any $x \in \mathbb{R}^n$. How would you use this routine to determine a matrix $A \in \mathbb{R}^{m \times n}$ such that $F(x) = Ax$ for all $x \in \mathbb{R}^n$?

1.2.b. Show A is unique.

Solution: For the standard basis $e_1, \cdots, e_n$ of $\mathbb{R}^n$ we have uniquely $\forall x \in \mathbb{R}^n$

$$x = \sum_{i=1}^{n} e_i \xi_i.$$

By linearity we have

$$F(x) = F(\sum_{i=1}^{n} e_i \xi_i) = \sum_{i=1}^{n} F(e_i)\xi_i = \sum_{i=1}^{n} a_i \xi_i = Ax$$

where $F(e_i) = a_i = Ae_i$ determines the i-th column of A. So evaluating F on the n standard basis vectors yields A.

Suppose we are given two matrices A and B that define $F(\cdot)$. We have by definition

$$\begin{gathered} \forall x \in \mathbb{R}^n \;\; y = Ax = Bx \\ \therefore Ae_i = Be_i \;\; 1 \le i \le n \\ \therefore A = B \end{gathered}$$

Problem 1.3

Let $y \in \mathbb{R}^m$ and $\|y\|$ be any vector norm defined on $\mathbb{R}^m$. Let $x \in \mathbb{R}^n$ and A be an $m \times n$ matrix with $m > n$.

1.3.a. Show that the function $f(x) = \|Ax\|$ is a vector norm on $\mathbb{R}^n$ if and only if A has full column rank, i.e., $rank(A) = n$.

1.3.b. Suppose we choose $f(x)$ from part (1.3.a) to be $f(x) = \|Ax\|_2$. What condition on A guarantees that $f(x) = \|x\|_2$ for any vector $x \in \mathbb{R}^n$?

Solution:

This question essentially asks when can we embed the vector space $\mathbb{R}^n$ in $\mathbb{R}^m$ in order to define a norm on $\mathbb{R}^n$.

$\rightarrow$

Let $y \in \mathbb{R}^m$ and $g(y) = \|y\|$ be a vector norm on $\mathbb{R}^m$. We know by the definiteness of norms that $g(y) = 0$ only when $y = 0$. So, since $y = Ax$, we must consider what x can lead to $y = 0$. By assumption, $f(x)$ is a vector norm for $\mathbb{R}^n$. We know by the definiteness of norms that $f(x) = 0$ only when $x = 0$.

Now assume that $rank(A) < n$. This means that there exists an $x \neq 0$ such that $Ax = 0$. Therefore we have $\exists x \neq 0$ such that $f(x) = f(Ax) = f(0) = 0$. This contradicts the assumption that $f(x)$ is a vector norm on $\mathbb{R}^n$. Therefore, $rank(A) = n$ must hold if $f(x)$ is a vector norm on $\mathbb{R}^n$.

$\leftarrow$

Now suppose $rank(A) = n$. Since $g(y)$ is a norm on $\mathbb{R}^n$ we need only check that $f(x)$ satisfies the properties of a norm by writing it in terms of $g(y)$.

Since A is full rank, $x \neq 0 \to y = Ax \neq 0$. Therefore,

$$\begin{aligned} f(x) &= g(Ax) \\ &= g(y) \\ &= \|y\| \\ &\neq 0 \end{aligned}$$

and f is definite.

Let $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

$$\begin{array}{rclcl} f(\alpha x) &=& g(A(\alpha x)) &=& g(\alpha Ax) \\ &=& g(\alpha y) &=& \|\alpha y\| \\ &=& |\alpha|\|y\| &=& |\alpha| g(y) \\ &=& |\alpha| g(Ax) &=& |\alpha| f(x) \end{array}$$

Therefore $f(x)$ is homogeneous.

Let $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^n$.

$$\begin{array}{rclcl} f(x_1 + x_2) &=& g(A(x_1 + x_2)) &=& g(Ax_1 + Ax_2) \\ &=& g(y_1 + y_2) &=& \|y_1 + y_2\| \\ &\leq& \|y_1\| + \|y_2\| &=& g(y_1) + g(y_2) \\ &=& g(Ax_1) + g(Ax_2) &=& f(x_1) + f(x_2) \end{array}$$

For the second part of the question, if the matrix $A \in \mathbb{R}^{m \times n}$ is an isometry, i.e., it has orthonormal columns, then $A^T A = I_n$. We therefore have

$$\begin{aligned} f(x)^2 &= \|Ax\|_2^2 \\ &= x^T A^T A x \\ &= x^T x \\ &= \|x\|_2^2 \end{aligned}$$

and $f(x) = \|x\|_2$ on $\mathbb{R}^n$ as desired.

Problem 1.4

Theorem 1. *If $\mathcal{V}$ is a real vector space with a norm $\|v\|$ that satisfies the parallelogram law*

$$\forall x, \quad y \in \mathcal{V}, \quad \|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{1}$$

then the function

$$f(x, y) = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2$$

is an inner product on $\mathcal{V}$ and $f(x, x) = \|x\|^2$.

This problem proves this theorem by a series of lemmas. Prove each of the following lemmas and then prove the theorem.

Lemma 2. $\forall x \in \mathcal{V}$

$$f(x,x) = \|x\|^2$$

Lemma 3. $\forall x,\ \ y \in \mathcal{V}$ *$f(x,x)$ is definite and $f(x,y) = f(y,x)$, i.e., (f is symmetric)*

Lemma 4. *The following two "cosine laws" hold $\forall x,\ \ y \in \mathcal{V}$:*

$$2f(x,y) = \|x+y\|^2 - \|x\|^2 - \|y\|^2 \tag{2}$$

$$2f(x,y) = -\|x-y\|^2 + \|x\|^2 + \|y\|^2 \tag{3}$$

Lemma 5. $\forall x,\ \ y \in \mathcal{V}$:

$$|f(x,y)| \le \|x\|\|y\| \tag{4}$$

$$f(x,y) = \gamma\|x\|\|y\|, \quad sign(\gamma) = sign(f(x,y)), \quad 0 \le |\gamma| \le 1 \tag{5}$$

Lemma 6. $\forall x,\ \ y,\ \ z \in \mathcal{V}$:

$$f(x+z,y) = f(x,y) + f(z,y)$$

Lemma 7. $\forall x,\ \ y \in \mathcal{V},\ \ \alpha \in \mathbb{R}$

$$f(\alpha x,y) = \alpha f(x,y)$$

Solution:

Proof of Lemma 2.

This follows directly from the definition of f:

$$f(x,x) = \frac{1}{4}\|x+x\|^2 = \frac{1}{4}\|2x\|^2 = \|x\|^2$$

So if f is an inner product it induces the norm.

Proof of Lemma 3.

The definiteness of $f(x,x)$ follows directly from the definiteness of the norm $\|x\|$ and Lemma 2. The symmetry of f follows from its definition and the definition of a norm

$$f(x,y) = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 = \frac{1}{4}\|y+x\|^2 - \frac{1}{4}\|y-x\|^2 = f(y,x)$$

Proof of Lemma 4.

By definition and (1)

$$\begin{aligned} 2f(x,y) = \frac{1}{2}\|x+y\|^2 - \frac{1}{2}\|x-y\|^2 &= \frac{1}{2}\|x+y\|^2 + \frac{1}{2}\left[\|x+y\|^2 - 2\|x\|^2 - 2\|y\|^2\right] \\ &= \|x+y\|^2 - \|x\|^2 - \|y\|^2 \end{aligned}$$

proving (2).

By definition and (1)

$$2f(x,y) = \frac{1}{2}\|x+y\|^2 - \frac{1}{2}\|x-y\|^2 = -\frac{1}{2}\|x-y\|^2 + \frac{1}{2}\left[2\|x\|^2 + 2\|y\|^2 - \|x-y\|^2\right]$$
$$= -\|x-y\|^2 + \|x\|^2 + \|y\|^2$$

proving (3).

Proof of Lemma 5.

Assume $f(x,y) \geq 0$. By the triangle inequality property of the norm $\|v\|$,

$$\|x+y\| \leq \|x\| + \|y\|$$
$$\|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$

Using this and (2) yields,

$$0 \leq 2f(x,y) + \|x\|^2 + \|y\|^2 = \|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$
$$0 \leq f(x,y) \leq \|x\|\|y\|$$
$$\exists 0 \leq \gamma \leq 1, \text{ such that } f(x,y) = \gamma\|x\|\|y\|$$

Repeating for $f(x,y) < 0$ yields

$$f(x,y) = \gamma\|x\|\|y\|, \ \ 0 \leq |\gamma| \leq 1, \ \ sign(\gamma) = sign(f(x.y)).$$

Proof of Lemma 6.

By definition, (2) and (3),

$$4f(x,y) = 2\|x+y\|^2 - 2\|x\|^2 - 2\|y\|^2$$
$$4f(z,y) = -2\|z-y\|^2 + 2\|z\|^2 + 2\|y\|^2$$
$$4(f(x,y) + f(z,y)) = 2\|x+y\|^2 - 2\|z-y\|^2 - 2\|x\|^2 + 2\|z\|^2. \tag{6}$$

By definition and (1) (twice),

$$4f(x+z,y) = \|x+z+y\|^2 - \|x+z-y\|^2$$
$$= \left[-\|x+y-z\|^2 + 2\|x+y\|^2 + 2\|z\|^2\right] - \|x+z-y\|^2$$
$$= 2\|x+y\|^2 + 2\|z\|^2 - \|x+(y-z)\|^2 - \|x-(y-z)\|^2$$
$$= 2\|x+y\|^2 + 2\|z\|^2 - \|x-(y-z)\|^2 + \|x-(y-z)\|^2 - 2\|x\|^2 - 2\|y-z\|^2$$
$$= 2\|x+y\|^2 + 2\|z\|^2 - 2\|x\|^2 - 2\|z-y\|^2$$
$$= 4(f(x,y) + f(z,y))$$

with the last equality following from (6).

Proof of Lemma 7.

The result is trivially true for $\alpha = 0$. To prove the general result is not so simple. This is a standard exercise in vector spaces and various forms of the solution can be found in textbooks or on reputable mathematics websites. Here the solution in the exercises of Matrix Analysis by Horn and Johnson (Cambridge Press) is followed with some slight modifications.

Note first that it follows trivially from the definition that

$$-f(x,y) = f(-x,y) = f(x,-y). \tag{7}$$

Now consider $k \in \mathbb{Z}$ with $k > 0$. By applying Lemma 6 repeatedly as needed we have

$$kf(x,y) = f(x,y) + \cdots + f(x,y) = f(x + \cdots + x, y) = f(kx, y)$$

$$kf(x,y) = f(x,y) + \cdots + f(x,y) = f(x, y + \cdots + y) = f(x, ky)$$

and therefore

$$kf(x,y) = f(kx,y) = f(x,ky). \tag{8}$$

Combining (7) and a modified repetition of the derivation of (8) extends the result to $k \in \mathbb{Z}$.

The next step is to extend the result to rational numbers $r = j/k$ where $j,\ k \in \mathbb{Z}$:

$$j\ f(x,y) = j\ f(k\ k^{-1}\ x, y) = k\ j\ f(k^{-1}\ x, y) = k\ f(j\ k^{-1}\ x, y) = k\ f(r\ x, y)$$

$$\therefore\ \ j\ f(x,y) = k\ f(r\ x, y)$$

$$k \neq 0 \rightarrow\ \ \frac{j}{k}\ f(x,y) = r\ f(x,y) = f(r\ x, y).$$

Note that the rational number k^{-1} is not moved out of the arguments to f so there is nothing circular about the proof. It is also easy to see that a similar argument yields $r\ f(x,y) = f(x, r\ y)$.

To extend scaling by a real number α a limiting argument is used. This takes several forms in the literature. Let $\alpha \in \mathbb{R}$, $r \in \mathbb{Q}$, $x,\ y \in \mathbb{R}^n$.Using the scaling result for $r \in \mathbb{Q}$, Lemma 6, Lemma 5, scaling of norms, and the triangle inequality for absolute values on $\mathbb{R}$ yields

$$\begin{aligned}|f(\alpha\ x, y) - \alpha\ f(x,y)| &= |f(\alpha\ x, y) - \alpha\ f(x,y) + r\ f(x,y) - f(r\ x, y)| \\ &= |f((\alpha - r)\ x, y) + (r - \alpha)\ f(x,y)| \\ &\leq \|(\alpha - r)\ x\|\ \|y\| + |(r - \alpha)|\ \|x\|\ \|y\| \\ &\leq 2\ |(\alpha - r)|\ \|x\|\ \|y\|.\end{aligned}$$

Since $|\alpha - r|$ can be made arbitrarily small, it follows that

$$\alpha\ f(x,y) = f(\alpha\ x, y)$$

The argument is easily modified to add the additional equality

$$\alpha\ f(x,y) = f(\alpha\ x, y) = f(x, \alpha\ y).$$

Proof of Theorem 1.

By Lemma 3 $f(x,y)$ is definite.

By Lemma 7 $f(\alpha x,y)=\alpha f(x,y)$. By Lemma 6 $f(x+z,y)=f(x,y)+f(z,y)$. It follows that

$$f(\alpha x+\beta z,y)=\alpha f(x,y)+\beta f(z,y)$$

i.e., f is linear.

Therefore, $\langle x,y\rangle=f(x,y)$ is an inner product that, by Lemma 2, induces $\|x\|$.

Problem 1.5

1.5.a Let $A\in\mathbb{R}^{n\times n}$ and $B\in\mathbb{R}^{n\times n}$ be nonsingular matrices. Show $(AB)^{-1}=B^{-1}A^{-1}$.

1.5.b Suppose $A\in\mathbb{R}^{m\times n}$ with $m>n$ and let $M\in\mathbb{R}^{n\times n}$ be a nonsingular square matrix. Show that $\mathcal{R}(A)=\mathcal{R}(AM)$ where $\mathcal{R}(\dot{})$ denotes the range of a matrix.

Solution:

By assumption, the $n\times n$ nonsingular matrices A^{-1} and B^{-1} exist, are unique and $AA^{-1}=BB^{-1}=A^{-1}A=B^{-1}B=I$. Let $G=B^{-1}A^{-1}$.

We have

$$\begin{aligned}
(AB)G &= ABB^{-1}A^{-1}\\
&= A(BB^{-1})A^{-1}\\
&= AIA^{-1}\\
&= AA^{-1}\\
&= I\\
G(AB) &= B^{-1}A^{-1}AB\\
&= B^{-1}(A^{-1}A)B\\
&= B^{-1}IB\\
&= B^{-1}B\\
&= I
\end{aligned}$$

To see that G is unique, suppose there is another matrix $Q\neq G$ such that $ABQ=I$. We have

$$Q=IQ=(GAB)Q=G(ABQ)=G$$

Now suppose that $Q\neq G$ such that $QAB=I$. We have

$$Q=QI=Q(ABG)=(QAB)G=G$$

(Strictly speaking you need only prove one of these to show uniqueness.)

Assuming M is nonsingular, we must show $\mathcal{R}(A)\subseteq\mathcal{R}(AM)$ and $\mathcal{R}(A)\supseteq\mathcal{R}(AM)$. We have $y\in\mathcal{R}(A)\rightarrow\exists x\in\mathbb{R}^n$ such that $y=Ax$. M nonsingular implies that $\forall x\in\mathbb{R}^n\ \ \exists c\in\mathbb{R}^n$ such that $x=Mc$. Therefore, $y=AMc$ and $y\in\mathcal{R}(AM)$.

We have $y \in \mathcal{R}(AM) \to \exists x \in \mathbb{R}^n$ such that $y = AMx$. Also $M \in \mathbb{R}^{n \times n} \to b = Mx \in \mathbb{R}^n$. Therefore, $y = AMx = Ab \to y \in \mathcal{R}(A)$.

Problem 1.6

Consider the matrix

$$L = \begin{pmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & \lambda_{22} & 0 & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & 0 \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & \lambda_{44} \end{pmatrix}$$

Suppose that $\lambda_{11} \neq 0$, $\lambda_{33} \neq 0$, $\lambda_{44} \neq 0$ but $\lambda_{22} = 0$.

1.6.a. Show that L is singular.

1.6.b. Determine a basis for the nullspace $\mathcal{N}(L)$.

Solution: To show that L is singular and find the nullspace we must determine the structure of the $x \neq 0$ such that $Lx = 0$. Imposing the λ_{ii} constraints we have, $\lambda_{11} \neq 0 \to \xi_1 = 0$ and therefore,

$$\begin{pmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & 0 \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & \lambda_{44} \end{pmatrix} \begin{pmatrix} 0 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

where ξ_2 is arbitrary and

$$\begin{pmatrix} \lambda_{33} & 0 \\ \lambda_{43} & \lambda_{44} \end{pmatrix} \begin{pmatrix} \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} -\lambda_{32} \\ -\lambda_{42} \end{pmatrix} \xi_2.$$

Since $\lambda_{33} \neq 0$ and $\lambda_{44} \neq 0$ the 2×2 matrix is nonsingular and therefore, ξ_3 and ξ_4 are uniquely determined given a particular value for ξ_2, i.e., there are no further degrees of freedom in the null space vectors. It follows that the dimension of $\mathcal{N}(L)$ is 1. So

$$\mathcal{N}(L) = span \left[\begin{pmatrix} 0 \\ 1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \right] \quad \text{where} \quad \begin{pmatrix} \lambda_{33} & 0 \\ \lambda_{43} & \lambda_{44} \end{pmatrix} \begin{pmatrix} \nu_2 \\ \nu_3 \end{pmatrix} = \begin{pmatrix} -\lambda_{32} \\ -\lambda_{42} \end{pmatrix}$$

Problem 1.7

Suppose $A \in \mathbb{C}^{m \times n}$ and let the matrix B be **any submatrix of** A. Show that $\|B\|_p \leq \|A\|_p$.
Solution:

We prove this in three steps. First let $x \in \mathbb{C}^n$, $x_1 \in \mathbb{C}^{n_1}$, $A_{11} \in \mathbb{C}^{m_1 \times n_1}$ and partition A and x as

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Define $\mathcal{B}_n = \{x \in \mathbb{C}^n \mid \|x\|_p = 1\}$ and $\mathcal{B}_{n_1} = \{x \in \mathbb{C}^n \mid x^H = \begin{pmatrix} x_1^H & 0 \end{pmatrix},\ \|x_1\|_p = 1\}$.

First note that since $\mathcal{B}_{n_1} \subset \mathcal{B}_n$, we have $\forall x \in \mathbb{C}^n$,

$$\|A\|_p = \max_{x \in \mathcal{B}_n} \|Ax\|_p \geq \max_{x \in \mathcal{B}_{n_1}} \|Ax\|_p = \max_{\|x_1\|_p = 1} \|A_1 x_1\|_p = \|A_1\|_p.$$

Second note that $\forall y \in \mathbb{C}^m$, we have

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \to \|y\|_p^p = \|y_1\|_p^p + \|y_2\|_p^p.$$

It follows that $\forall x_1 \in \mathbb{C}^{n_1}$,

$$\|A_1\|_p = \max_{\|x_1\|_p = 1} \|A_1 x_1\|_p = \max_{\|x_1\|_p = 1} \left\| \begin{pmatrix} A_{11} x_1 \\ A_{21} x_1 \end{pmatrix} \right\|_p \geq \max_{\|x_1\|_p = 1} \|A_{11} x_1\|_p = \|A_{11}\|_p.$$

We therefore have

$$\|A\|_p \geq \|A_1\|_p \geq \|A_{11}\|_p.$$

Finally, this can be applied to any $m_1 \times n_1$ submatrix since any submatrix can be placed in the block A_{11} by applying row and column permutations to A, i.e.,

$$P_r A P_c = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

and permutations do not affect the p-norm.

Problem 1.8

Suppose that $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ and let $E = uv^T$.

1.8.a Show that $\|E\|_F = \|E\|_2 = \|u\|_2 \|v\|_2$.

1.8.b Show that $\|E\|_\infty = \|u\|_\infty \|v\|_1$.

Solution: Denote $e_i^T u = \mu_i$ and $e_i^T v = \nu_i$. We have $\|E\|_F = \|u\|_2 \|v\|_2$ since

$$\begin{aligned} \|E\|_F^2 &= \sum_{i=1}^n \|Ee_i\|_2^2 \\ &= \sum_{i=1}^n \|u\|_2^2 |\nu_i|^2 \\ &= \|u\|_2^2 \sum_{i=1}^n |\nu_i|^2 \\ &= \|u\|_2^2 \|v\|_2^2 \end{aligned}$$

We have $\|E\|_2 = \|u\|_2\|v\|_2$ since for $\|x\|_2 = 1$ it follows that

$$\begin{aligned}
\|Ex\|_2^2 &= x^T E^T E x \\
&= x^T v u^T u v^T x \\
&= \|u\|_2^2 (v^T x)^2 \\
\|Ex\|_2 &= \|u\|_2 |v^T x| \\
&= \|u\|_2 \|v\|_2 \|x\|_2 |\cos\theta| \\
&= \|u\|_2 \|v\|_2 |\cos\theta|
\end{aligned}$$

This is maximized when x is colinear with v and $\cos\theta = \pm 1$.

It follows that $\|E\|_F = \|E\|_2$ by transitivity.

We have

$$\begin{aligned}
\|E\|_\infty &= \max_{1\le i\le m} \|e_i^H E\|_1 \\
&= \max_{1\le i\le m} \|\mu_i v^H\|_1 \\
&= \max_{1\le i\le m} \left(|\mu_i| \|v^H\|_1 \right) \\
&= \|v^H\|_1 \left(\max_{1\le i\le m} |\mu_i| \right) \\
&= \|v^H\|_1 \|u\|_\infty \\
&= \|u\|_\infty \|v\|_1
\end{aligned}$$

Problem 1.9

Let $\mathcal{S}_1 \subset \mathbb{R}^n$ and $\mathcal{S}_2 \subset \mathbb{R}^n$ be two subspaces of $\mathbb{R}^n$.

1.9.a. Suppose $x_1 \in \mathcal{S}_1$, $x_1 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. $x_2 \in \mathcal{S}_2$, and $x_2 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. Show that x_1 and x_2 are linearly independent.

Solution:

Proof by contradiction. Assume they are dependent, i.e., $x_1 = \alpha x_2$. $x_1 = \alpha x_2 \rightarrow x_1 \in \mathcal{S}_2$ which is a contradiction. $\therefore x_1,\ x_2$ are linearly independent.

1.9.b. Suppose $x_1 \in \mathcal{S}_1$, $x_1 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. $x_2 \in \mathcal{S}_2$, and $x_2 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. Also, suppose that $x_3 \in \mathcal{S}_1 \cap \mathcal{S}_2$ and $x_3 \neq 0$, i.e., the intersection is not empty. Show that x_1, x_2 and x_3 are linearly independent.

Solution:

Proof by contradiction. Assume they are dependent, i.e., $x_3 = \alpha_1 x_1 + \alpha_2 x_2$. Three cases possible:

1. $\alpha_1 \neq 0$ and $\alpha_2 = 0 \to x_3 \in \mathcal{S}_1 \ \ and \ \ x_3 \notin \mathcal{S}_2$. Contradiction.
2. $\alpha_2 \neq 0$ and $\alpha_1 = 0 \to x_3 \in \mathcal{S}_2 \ \ and \ \ x_3 \notin \mathcal{S}_1$. Contradiction.
3. $\alpha_1 \neq 0$ and $\alpha_2 \neq 0 \to x_2 = \frac{\alpha_1}{\alpha_2} x_1 - \frac{1}{\alpha_2} x_3 \in \mathcal{S}_1$. Contradiction.

$\therefore x_1, \ x_2, \ x_3$ are linearly independent.

Problem 1.10

Suppose $A \in \mathbb{C}^{m \times n}$. Is the matrix norm $\|A\|_1$ induced by the two vector norms $\|x\|_1$ and $\|y\|_1$ for $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^m$?

Solution: The matrix norm is induced by the two vector norms.

The norm is defined as

$$\|A\|_1 = \max_{1 \leq i \leq n} \|Ae_i\|_1.$$

Consider the definition of the induced norm

$$\|A\| = \max_{\|x\|_1 = 1} \|Ax\|_1 = \max_{\|x\|_1 = 1} \|\sum_{i=1}^{n} \xi_i A e_i\|_1 \leq \max_{\|x\|_1 = 1} \left(\sum_{i=1}^{n} |\xi_i| \|Ae_i\|_1 \right) \leq \max_{1 \leq i \leq n} \|Ae_i\|_1$$

by the triangle inequality and since $|\xi_1| + \cdots + |\xi_n| = 1$. Since the bound on the right is achievable

$$\|A\| = \|A\|_1.$$

Problem 1.11

Consider the definition of the matrix norm $\|A\| = \max_{i,j} |\alpha_{i,j}|$ where $e_i^T A e_j = \alpha_{i,j}$.

1.11.a. Show that this defines a matrix norm.

1.11.b. Show that the matrix norm is not consistent.

Solution:

Recall that for the complex numbers the absolute value, $|\alpha|$ function is a norm and therefore satisfies all three required properties:

$$\begin{array}{rcl} \alpha \neq 0 & \to & |\alpha| > 0 \\ |\alpha\beta| & = & |\alpha||\beta| \\ |\alpha + \beta| & \leq & |\alpha| + |\beta| \end{array}$$

So we can check each of the properties for the proposed matrix norm.

The proposed norm is definite since $A \neq 0 \to \exists p, q$ such that $\alpha_{pq} \neq 0$ and therefore

$$\|A\| = \max_{ij} |\alpha_{ij}| \geq |\alpha_{pq}| > 0$$

The proposed norm is homogeneous. We have

$$\begin{aligned} \|\beta A\| &= \max_{ij} |\beta \alpha_{ij}| \\ &= \max_{ij} |\beta||\alpha_{ij}| \\ &= |\beta| \max_{ij} |\alpha_{ij}| \\ &= \beta \|A\| \end{aligned}$$

The proposed norm satisfies the triangle inequality.

$$\begin{aligned} \|A + B\| &= \max_{ij} |\alpha_{ij} + \beta_{ij}| \\ &\leq \max_{ij}(|\alpha_{ij}| + |\beta_{ij}|) \\ &\leq \max_{ij}(|\alpha_{ij}|) + \max_{ij}(|\beta_{ij}|) \\ &= \|A\| + \|B\| \end{aligned}$$

The proposed norm however, **is not consistent**. To show this we need only produce a single counterexample. Consider

$$\begin{aligned} A &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ \|AA\| &= \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \\ &= 2 \\ \|A\|\|A\| &= 1 \end{aligned}$$