

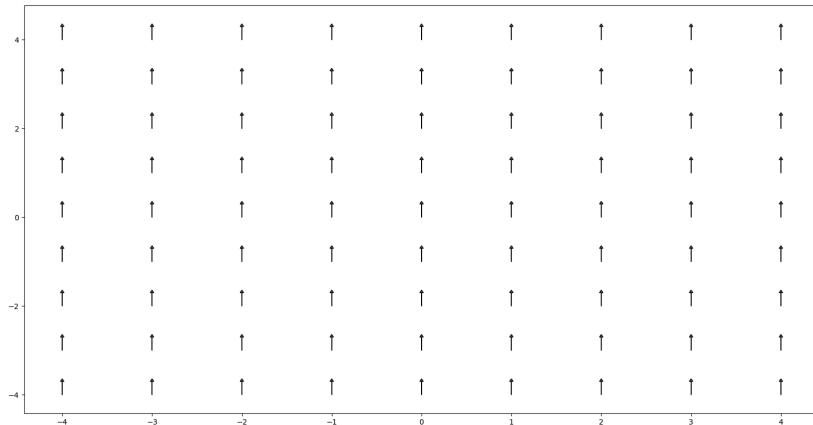
Solved selected problems of Vector Calculus, Linear Algebra and Differential Forms by Hubbard

Franco Zacco

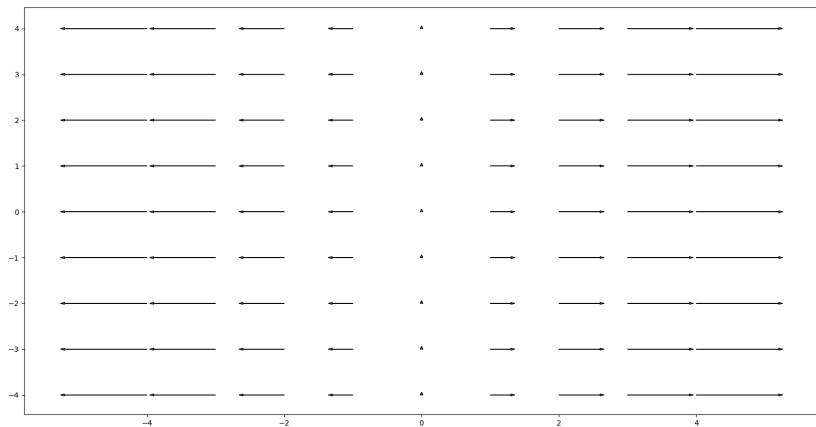
1.1 - Introducing the Actors: Points and Vectors

Proof. 1.1.6

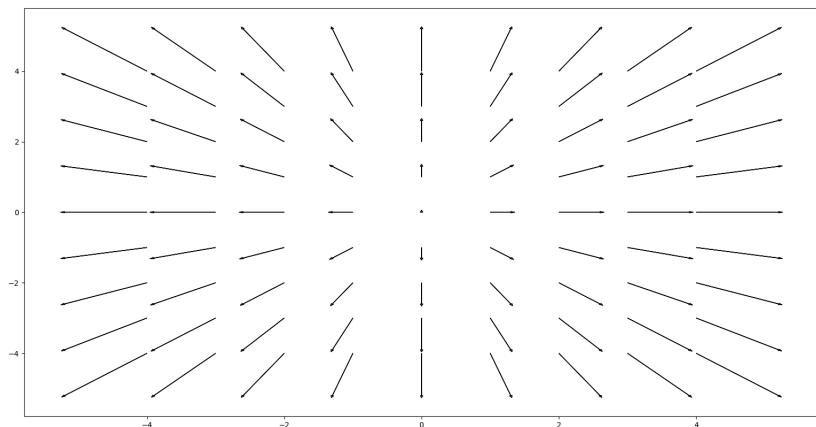
a. $\vec{F} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



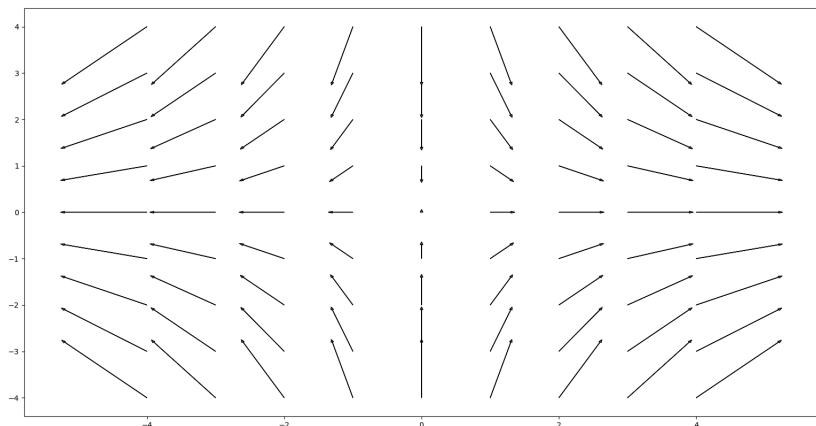
b. $\vec{F}(x) = \begin{bmatrix} x \\ 0 \end{bmatrix}$



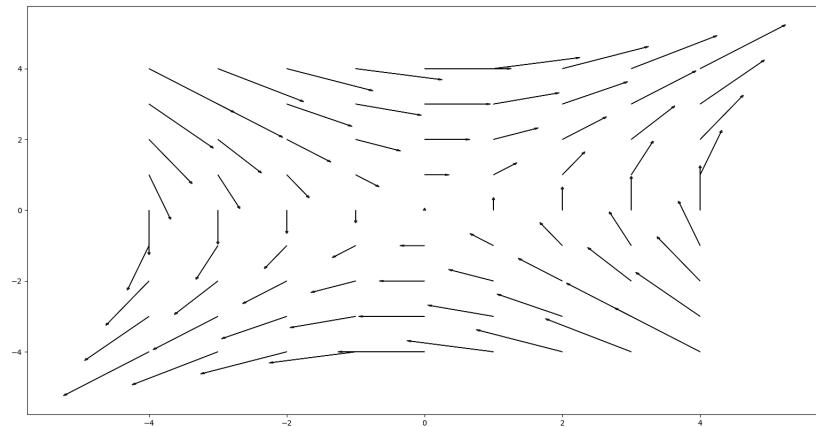
c. $\vec{F}(x) = \begin{bmatrix} x \\ y \end{bmatrix}$



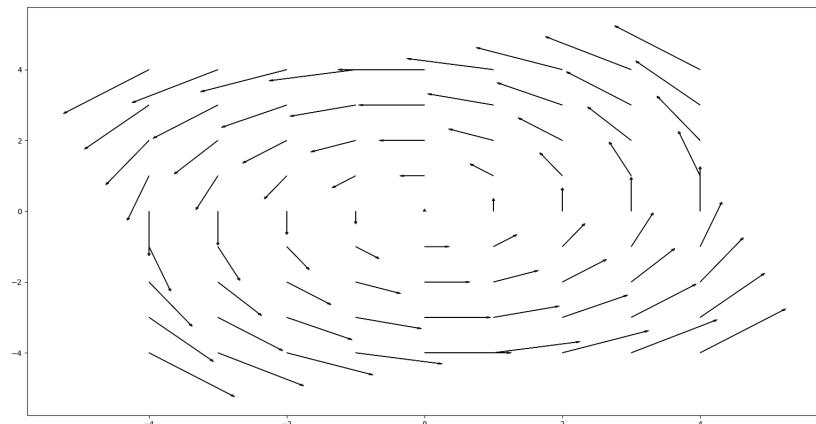
d. $\vec{F}(x) = \begin{bmatrix} x \\ -y \end{bmatrix}$



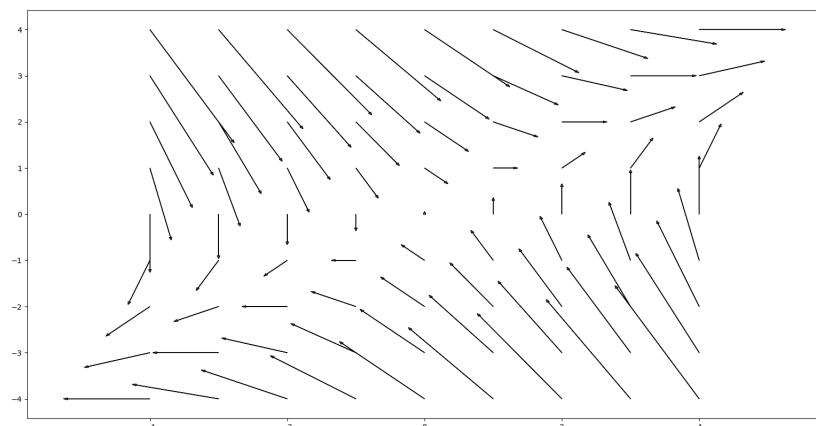
e. $\vec{F}(x) = \begin{bmatrix} y \\ x \end{bmatrix}$



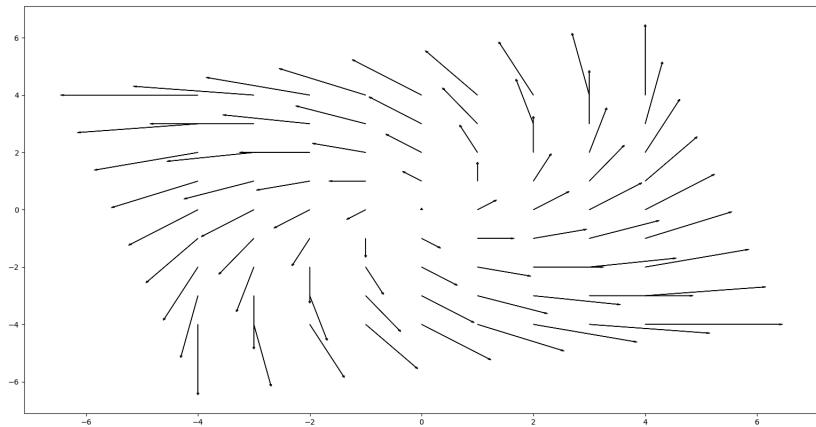
f. $\vec{F}(x) = \begin{bmatrix} -y \\ x \end{bmatrix}$



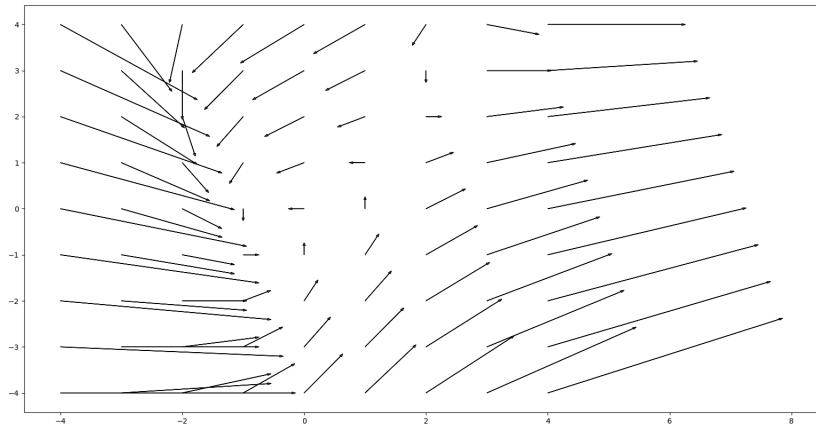
g. $\vec{F}(x) = \begin{bmatrix} y \\ x-y \end{bmatrix}$



h. $\vec{F}(x) = \begin{bmatrix} x - y \\ x + y \end{bmatrix}$



i. $\vec{F}(x) = \begin{bmatrix} x^2 - y - 1 \\ x - y \end{bmatrix}$



□

Proof. 1.1.8

- a. Assuming the pipe is in the direction of z -axis and we are in \mathbb{R}^3 space then the vector field describing the flow in the pipe must be

$$\vec{F} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} 0 \\ 0 \\ r^2 - x^2 - y^2 \end{bmatrix}$$

Where we used that $a^2 = x^2 + y^2$.

- b. Let us define a tube with a "torus" form, centered at 0, where the distance from the center to the tube axis is 1 (the unit circle is the axis of the tube), with a tube radius r then any point in the tube is given by

$$\begin{aligned} x &= (1 + r \cos \phi) \cos \theta \\ y &= (1 + r \cos \phi) \sin \theta \\ z &= r \sin \phi \end{aligned}$$

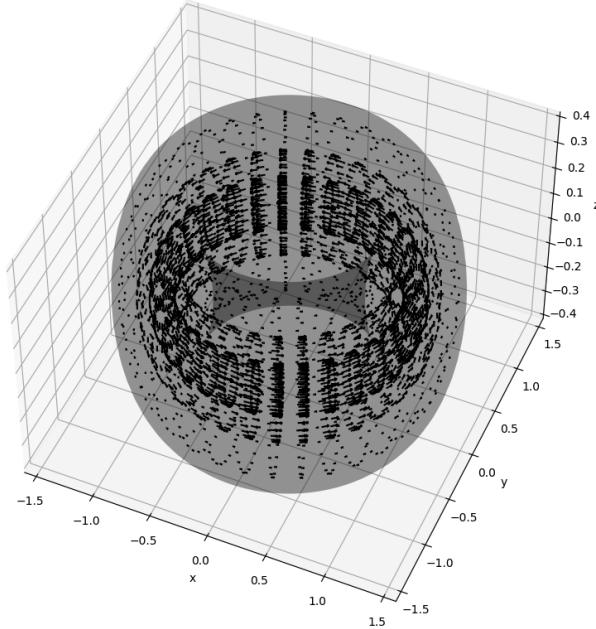
Also, we see that the distance from any point inside the tube and the axis of the tube is given by

$$a^2 = (x - \cos \theta)^2 + (y - \sin \theta)^2 + z^2$$

Then the vector field describing the flow is given by

$$\vec{F} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} -(r^2 - a^2) \sin \theta \\ (r^2 - a^2) \cos \theta \\ 0 \end{bmatrix}$$

The equation can be written only in terms of x, y and z using that $\theta = \arctan(y/x)$. The vector field looks like the following



□

1.2 - Introducing the Actors: Matrices

Proof. 1.2.16

Let A be a $n \times m$ matrix, we write A as

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,m} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,m} \end{bmatrix}$$

Then A^T is

$$A^T = \begin{bmatrix} a_{1,1} & a_{2,1} & a_{3,1} & \dots & a_{n,1} \\ a_{1,2} & a_{2,2} & a_{3,2} & \dots & a_{n,2} \\ a_{1,3} & a_{2,3} & a_{3,3} & \dots & a_{n,3} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{1,m} & a_{2,m} & a_{3,m} & \dots & a_{n,m} \end{bmatrix}$$

So $A^T A$ is given by

$$A^T A = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,m} \\ A_{2,1} & A_{2,2} & \dots & A_{2,m} \\ \vdots & \vdots & \dots & \vdots \\ A_{m,1} & A_{m,2} & \dots & A_{m,m} \end{bmatrix}$$

Where

$$A_{1,1} = a_{1,1}^2 + a_{2,1}^2 + a_{3,1}^2 + \dots + a_{n,1}^2$$

$$A_{1,2} = a_{1,1}a_{1,2} + a_{2,1}a_{2,2} + a_{3,1}a_{3,2} + \dots + a_{n,1}a_{n,2}$$

$$A_{1,m} = a_{1,1}a_{1,m} + a_{2,1}a_{2,m} + a_{3,1}a_{3,m} + \dots + a_{n,1}a_{n,m}$$

$$A_{2,1} = a_{1,2}a_{1,1} + a_{2,2}a_{2,1} + a_{3,2}a_{3,1} + \dots + a_{n,2}a_{n,1}$$

$$A_{2,2} = a_{1,2}^2 + a_{2,2}^2 + a_{3,2}^2 + \dots + a_{n,2}^2$$

$$A_{2,m} = a_{1,2}a_{1,m} + a_{2,2}a_{2,m} + a_{3,2}a_{3,m} + \dots + a_{n,2}a_{n,m}$$

$$A_{m,1} = a_{1,m}a_{1,1} + a_{2,m}a_{2,1} + a_{3,m}a_{3,1} + \dots + a_{n,m}a_{n,1}$$

$$A_{m,2} = a_{1,m}a_{1,2} + a_{2,m}a_{2,2} + a_{3,m}a_{3,2} + \dots + a_{n,m}a_{n,2}$$

$$A_{m,m} = a_{1,m}^2 + a_{2,m}^2 + a_{3,m}^2 + \dots + a_{n,m}^2$$

As we can see $A_{1,2} = A_{2,1}$, $A_{m,1} = A_{1,m}$ and $A_{m,2} = A_{2,m}$, and hence $A^T A$ is a symmetric matrix.

In the same way, if we let AA^T be

$$AA^T = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ A_{n,1} & A_{n,2} & \dots & A_{n,n} \end{bmatrix}$$

Where

$$A_{1,1} = a_{1,1}^2 + a_{1,2}^2 + a_{1,3}^2 + \dots + a_{1,m}^2$$

$$A_{1,2} = a_{1,1}a_{2,1} + a_{1,2}a_{2,2} + a_{1,3}a_{2,3} + \dots + a_{1,m}a_{2,m}$$

$$A_{1,n} = a_{1,1}a_{n,1} + a_{1,2}a_{n,2} + a_{1,3}a_{n,3} + \dots + a_{1,m}a_{n,m}$$

$$A_{2,1} = a_{2,1}a_{1,1} + a_{2,2}a_{1,2} + a_{2,3}a_{1,3} + \dots + a_{2,m}a_{1,m}$$

$$A_{2,2} = a_{2,1}^2 + a_{2,2}^2 + a_{2,3}^2 + \dots + a_{2,m}^2$$

$$A_{2,n} = a_{2,1}a_{n,1} + a_{2,2}a_{n,2} + a_{2,3}a_{n,3} + \dots + a_{2,m}a_{n,m}$$

$$A_{n,1} = a_{n,1}a_{1,1} + a_{n,2}a_{1,2} + a_{n,3}a_{1,3} + \dots + a_{n,m}a_{1,m}$$

$$A_{n,2} = a_{n,1}a_{2,1} + a_{n,2}a_{2,2} + a_{n,3}a_{2,3} + \dots + a_{n,m}a_{2,m}$$

$$A_{n,n} = a_{n,1}^2 + a_{n,2}^2 + a_{n,3}^2 + \dots + a_{n,m}^2$$

We see again that $A_{1,2} = A_{2,1}$, $A_{n,1} = A_{1,n}$ and $A_{n,2} = A_{2,n}$. Therefore AA^T is also a symmetric matrix. \square

Proof. 1.2.20

Let $z = x + iy \in \mathbb{C}$ and the matrix M_z be

$$M_z = \begin{bmatrix} x & y \\ -y & x \end{bmatrix}$$

Then $M_{z_1+z_2}$ for $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ in \mathbb{C} is

$$M_{z_1+z_2} = \begin{bmatrix} x_1 + x_2 & y_1 + y_2 \\ -(y_1 + y_2) & x_1 + x_2 \end{bmatrix}$$

But we see that

$$M_{z_1} + M_{z_2} = \begin{bmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{bmatrix} + \begin{bmatrix} x_2 & y_2 \\ -y_2 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 & y_1 + y_2 \\ -(y_1 + y_2) & x_1 + x_2 \end{bmatrix}$$

Therefore $M_{z_1+z_2} = M_{z_1} + M_{z_2}$.

On the other hand, let us compute $M_{z_1 z_2}$ where

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$$

So

$$M_{z_1 z_2} = \begin{bmatrix} x_1 x_2 - y_1 y_2 & x_1 y_2 + x_2 y_1 \\ -(x_1 y_2 + x_2 y_1) & x_1 x_2 - y_1 y_2 \end{bmatrix}$$

Now we compute $M_{z_1} M_{z_2}$ as follows

$$M_{z_1} M_{z_2} = \begin{bmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{bmatrix} \cdot \begin{bmatrix} x_2 & y_2 \\ -y_2 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 - y_1 y_2 & x_1 y_2 + y_1 x_2 \\ -(y_1 x_2 + x_1 y_2) & -y_1 y_2 + x_1 x_2 \end{bmatrix}$$

Hence $M_{z_1 z_2} = M_{z_1} M_{z_2}$. □

1.3 - What the Actors do: Matrix multiplication as a linear transformation

Proof. 1.3.4

a. Let T be a linear transformation such that

$$T \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

We know that applying T to \vec{e}_1 we get the first column of the matrix T then

$$T\vec{e}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

So in the same way we get that the matrix of this linear transformation is

$$T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

b. In the case of

$$T \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_2 \\ v_1 + 2v_2 \\ v_3 + v_1 \end{bmatrix}$$

We get by the same procedure that

$$T = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

□

Proof. 1.3.8

- a. Given that the linear transformation is $T : \mathbb{R}^5 \rightarrow \mathbb{R}^6$ we know that the matrix representing T is of size 6×5 so to reconstitute the matrix T we need to ask what is the output of $T\vec{e}_i$ for $i \in \{1, 2, 3, 4, 5\}$ so we get as the output the values of i-th column of T . Therefore we need to ask 5 questions to reconstitute it.
- b. If $T : \mathbb{R}^6 \rightarrow \mathbb{R}^5$ then the matrix representing it is of size 5×6 and hence following the same procedure as in part "a" we need to ask what is the output of $T\vec{e}_i$ for $i \in \{1, 2, 3, 4, 5, 6\}$ i.e. we need to ask 6 questions.
- c. Given that \vec{e}_i is a standard basis vectors in \mathbb{R}^5 or \mathbb{R}^6 depending on the case, if we multiply T times a different basis vector we will get a different output but the answer still represents the ith-column of T in that basis, so there is not only one right answer to the question, it depends on the basis. Although, in any case, we still need to ask five or six questions.

□

Proof. 1.3.9

We know that

$$T \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 4 \end{bmatrix}$$

We can write that

$$\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

If T were linear then must happen that

$$T \left(2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = 2T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The LHS is

$$T \left(2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = T \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 4 \end{bmatrix}$$

And the RHS is

$$2T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 5 \end{bmatrix}$$

Then the LHS is different from the RHS, therefore T cannot be linear. □

Proof. 1.3.10

If $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a linear transformation then must be that

$$T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = T \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Then

$$\begin{aligned} \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} \\ T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \\ T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \end{aligned}$$

But also must be that

$$T \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Then

$$\begin{aligned} \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} + T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} &= \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} \\ T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} &= \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} \\ T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} &= \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

Therefore, there is a linear transformation T with the following matrix describing it

$$[T] = \begin{bmatrix} 3 & 1 & -2 \\ 0 & 2 & 1 \\ 0 & 3 & -1 \end{bmatrix}$$

Where we used that $T(\vec{e}_1)$, $T(\vec{e}_2)$, and $T(\vec{e}_3)$ are the first, second and third column respectively of the matrix $[T]$. \square

Proof. 1.3.11

We know from Example 1.3.9 that the transformation R_θ rotates the input by an angle θ counterclockwise around the origin then the transformation R' that rotates the input by an angle θ clockwise is $R_{-\theta}$, therefore

$$R' = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

□

Proof. 1.3.12

- a. For a linear transformation $S : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that reflects a point across the plane $x = y$ must be that

$$S \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ x \\ z \end{bmatrix}$$

Then the matrix of S must be

$$[S] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now, for a linear transformation $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that reflects points across the plane $y = z$ must be that

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ z \\ y \end{bmatrix}$$

Then the matrix of T must be

$$[T] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Finally, the matrix of $S \circ T$ by Theorem 1.3.10 is

$$[S \circ T] = [S][T] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

- b. On the other hand, matrix $[T \circ S]$ is

$$[T \circ S] = [T][S] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Which is different from $[S \circ T]$ so $[S]$ and $[T]$ do not commute.

□

1.4 - The Geometry of \mathbb{R}^n

Proof. 1.4.13

Let \vec{a} and \vec{b} be two vectors pointing in the same direction then must be that $\vec{b} = k\vec{a}$ for some scalar k then

$$\vec{a} \times \vec{b} = \vec{a} \times k\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times k \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_2(ka_3) - a_3(ka_2) \\ -a_1(ka_3) + a_3(ka_1) \\ a_1(ka_2) - a_2(ka_1) \end{bmatrix} = \vec{0}$$

□

Proof. 1.4.14 Let $\vec{u}, \vec{v}, \vec{w}$ be vectors

a. Let us compute $\vec{u} \times (\vec{v} \times \vec{w})$ as follows

$$\begin{aligned}\vec{u} \times (\vec{v} \times \vec{w}) &= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times \left(\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right) \\ &= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 2-1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\end{aligned}$$

On the other hand, $(\vec{u} \times \vec{v}) \times \vec{w}$ is

$$\begin{aligned}(\vec{u} \times \vec{v}) \times \vec{w} &= \left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right) \times \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}\end{aligned}$$

b. Let us compute $\vec{v} \cdot (\vec{v} \times \vec{w})$ as follows

$$\begin{aligned}\vec{v} \cdot (\vec{v} \times \vec{w}) &= \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \cdot \left(\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right) \\ &= \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} \\ &= 2 \cdot 0 + 0 \cdot 3 + 1 \cdot 0 \\ &= 0\end{aligned}$$

Geometrically, this implies that \vec{v} is orthogonal to $\vec{v} \times \vec{w}$.

c. Let $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ then we see that

$$\begin{aligned}
\vec{a} \times (\vec{b} + \vec{c}) &= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 + c_1 \\ b_2 + c_2 \\ b_3 + c_3 \end{bmatrix} \\
&= \begin{bmatrix} a_2(b_3 + c_3) - a_3(b_2 + c_2) \\ -a_1(b_3 + c_3) + a_3(b_1 + c_1) \\ a_1(b_2 + c_2) - a_2(b_1 + c_1) \end{bmatrix} \\
&= \begin{bmatrix} a_2b_3 + a_2c_3 - a_3b_2 - a_3c_2 \\ -a_1b_3 - a_1c_3 + a_3b_1 + a_3c_1 \\ a_1b_2 + a_1c_2 - a_2b_1 - a_2c_1 \end{bmatrix}
\end{aligned}$$

But also we see that

$$\begin{aligned}
(\vec{a} \times \vec{b}) + (\vec{a} \times \vec{c}) &= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \\
&= \begin{bmatrix} a_2b_3 - a_3b_2 \\ -a_1b_3 + a_3b_1 \\ a_1b_2 - a_2b_1 \end{bmatrix} + \begin{bmatrix} a_2c_3 - a_3c_2 \\ -a_1c_3 + a_3c_1 \\ a_1c_2 - a_2c_1 \end{bmatrix} \\
&= \begin{bmatrix} a_2b_3 - a_3b_2 + a_2c_3 - a_3c_2 \\ -a_1b_3 + a_3b_1 - a_1c_3 + a_3c_1 \\ a_1b_2 - a_2b_1 + a_1c_2 - a_2c_1 \end{bmatrix}
\end{aligned}$$

Therefore $\vec{a} \times (\vec{b} + \vec{c}) = (\vec{a} \times \vec{b}) + (\vec{a} \times \vec{c})$.

□

Proof. 1.4.15 Let \vec{v} and \vec{w} be vectors in \mathbb{R}^3 then we know that

$$\vec{v} \times \vec{w} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} v_2 w_3 - v_3 w_2 \\ -v_1 w_3 + v_3 w_1 \\ v_1 w_2 - v_2 w_1 \end{bmatrix}$$

Also, we see that

$$-(\vec{w} \times \vec{v}) = - \left[\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right] = - \begin{bmatrix} w_2 v_3 - w_3 v_2 \\ -w_1 v_3 + w_3 v_1 \\ w_1 v_2 - w_2 v_1 \end{bmatrix} = \begin{bmatrix} w_3 v_2 - w_2 v_3 \\ -w_3 v_1 + w_1 v_3 \\ w_2 v_1 - w_1 v_2 \end{bmatrix}$$

Therefore $\vec{v} \times \vec{w} = -(\vec{w} \times \vec{v})$. □

Proof. 1.4.18

- a. Let the coordinates of \mathbf{a} and \mathbf{b} be ≥ 0 then the area spanned by them is $(a_1 + b_1)(a_2 + b_2)$ minus the area of the pieces 1-6 shown in the figure i.e.

$$(a_1 + b_1)(a_2 + b_2) - \underbrace{a_1 b_2}_{\text{area of 1}} - \underbrace{b_1 b_2 / 2}_{\text{area of 2}} - \underbrace{a_2 a_1 / 2}_{\text{area of 3}} - \underbrace{a_1 a_2 / 2}_{\text{area of 4}} - \underbrace{b_1 b_2 / 2}_{\text{area of 5}} - \underbrace{a_1 b_2}_{\text{area of 6}}$$

$$= a_1 a_2 + a_1 b_2 + b_1 a_2 + b_1 b_2 - a_1 b_2 - b_1 b_2 - a_2 a_1 - a_1 a_2$$

$$= b_1 a_2 - a_1 b_2$$

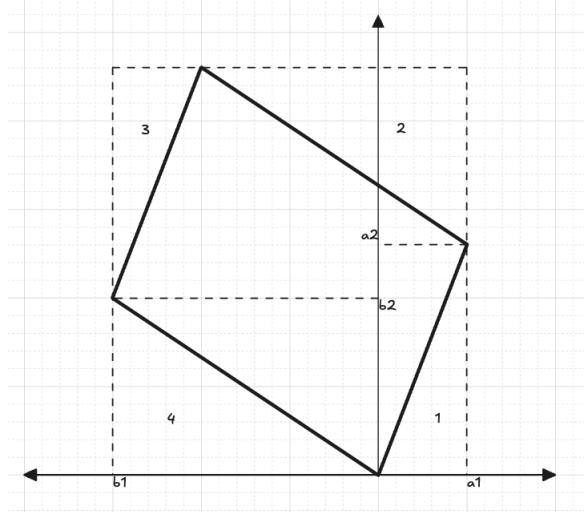
We see that $b_1 a_2 - a_1 b_2 \geq 0$ since $b_1 a_2 > a_1 b_2$.

On the other hand, we see that

$$\left| \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \right| = |a_1 b_2 - a_2 b_1| = a_2 b_1 - a_1 b_2$$

Therefore we see that part 1 of Proposition 1.4.14 holds.

- b. Suppose now that b_1 is negative then we are in the situation described by the following figure



So, to compute the spanned area between \mathbf{a} and \mathbf{b} we need to subtract the areas of the triangles 1-4. let us note that the triangles 1 and 3 have the same area and the same is true for the triangles 2 and 4. Then we compute what follows

$$(a_1 - b_1)(a_2 + b_2) - 2 \cdot \frac{a_1 a_2}{2} - 2 \cdot \frac{(-b_1)b_2}{2} =$$

$$= a_1 a_2 + a_1 b_2 - b_1 a_2 - b_1 b_2 - a_1 a_2 + b_1 b_2$$

$$= a_1 b_2 - b_1 a_2$$

We see that $a_1b_2 - b_1a_2 \geq 0$ since $-b_1a_2 > a_1b_2 > 0$. On the other hand, we see that

$$\left| \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \right| = |a_1b_2 - a_2b_1| = a_1b_2 - a_2b_1$$

Therefore we see that part 1 of Proposition 1.4.14 holds also in this case.

□

Proof. 1.4.24 Let $\vec{v} \in \mathbb{R}^n$ be a nonzero vector, and denote $\vec{v}^\perp \subset \mathbb{R}^n$ the set of vectors $\vec{w} \in \mathbb{R}^n$ such that $\vec{v} \cdot \vec{w} = 0$

a. Let $\vec{w}, \vec{u} \in \vec{v}^\perp$, we see that

$$\vec{v} \cdot (\vec{w} + \vec{u}) = \vec{v} \cdot \vec{w} + \vec{v} \cdot \vec{u} = 0$$

Where we used that the dot product is distributive. Then $\vec{w} + \vec{u} \in \vec{v}^\perp$.

On the other hand, let $a \in \mathbb{R}$ and let us consider the vector $a\vec{w}$ then

$$\vec{v} \cdot a\vec{w} = a(\vec{v} \cdot \vec{w}) = 0$$

Hence $a\vec{w} \in \vec{v}^\perp$. Therefore \vec{v}^\perp is a subspace of \mathbb{R}^n .

b. Let $\vec{a} \in \mathbb{R}^n$ then let us compute what follows

$$\begin{aligned} \vec{v} \cdot \left(\vec{a} - \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2} \vec{v} \right) &= \vec{v} \cdot \vec{a} - \vec{v} \cdot \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2} \vec{v} \\ &= \vec{v} \cdot \vec{a} - \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2} (\vec{v} \cdot \vec{v}) \\ &= \vec{v} \cdot \vec{a} - \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2} |\vec{v}|^2 \\ &= \vec{v} \cdot \vec{a} - \vec{a} \cdot \vec{v} \\ &= \vec{v} \cdot \vec{a} - \vec{v} \cdot \vec{a} \\ &= 0 \end{aligned}$$

Where we used that the dot product is distributive and associative. Therefore by the definition of \vec{v}^\perp we have that

$$\vec{a} - \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2} \vec{v} \in \vec{v}^\perp$$

c. Let

$$P_{\vec{v}^\perp}(\vec{a}) = \vec{a} - \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2} \vec{v}$$

be the projection of \vec{a} onto \vec{v}^\perp .

Suppose there are two numbers $t_1(\vec{a}), t_2(\vec{a})$ such that $(\vec{a} + t_1(\vec{a})\vec{v}) \in \vec{v}^\perp$ and $(\vec{a} + t_2(\vec{a})\vec{v}) \in \vec{v}^\perp$ then by definition must be that

$$\vec{v} \cdot (\vec{a} + t_1(\vec{a})\vec{v}) = 0 \quad \text{and} \quad \vec{v} \cdot (\vec{a} + t_2(\vec{a})\vec{v}) = 0$$

So

$$\begin{aligned} \vec{v} \cdot (\vec{a} + t_1(\vec{a})\vec{v}) &= \vec{v} \cdot (\vec{a} + t_2(\vec{a})\vec{v}) \\ \vec{v} \cdot \vec{a} + t_1(\vec{a})|\vec{v}|^2 &= \vec{v} \cdot \vec{a} + t_2(\vec{a})|\vec{v}|^2 \\ t_1(\vec{a})|\vec{v}|^2 &= t_2(\vec{a})|\vec{v}|^2 \\ t_1(\vec{a}) &= t_2(\vec{a}) \end{aligned}$$

Where we subtracted from both sides $\vec{v} \cdot \vec{a}$ in the third step and we divided by $|\vec{v}|^2$ on both sides in the last step.

Therefore there is a unique number $t(\vec{a})$ such that $(\vec{a} + t(\vec{a})\vec{v}) \in \vec{v}^\perp$.

Since $(\vec{a} + t(\vec{a})\vec{v}) \in \vec{v}^\perp$ then must be that

$$\vec{v} \cdot (\vec{a} + t(\vec{a})\vec{v}) = 0$$

Then

$$\begin{aligned}\vec{v} \cdot \vec{a} + t(\vec{a})|\vec{v}|^2 &= 0 \\ t(\vec{a})|\vec{v}|^2 &= -\vec{v} \cdot \vec{a} \\ t(\vec{a}) &= -\frac{\vec{v} \cdot \vec{a}}{|\vec{v}|^2}\end{aligned}$$

Therefore

$$\vec{a} + t(\vec{a})\vec{v} = \vec{a} - \frac{\vec{a} \cdot \vec{v}}{|\vec{v}|^2}\vec{v} = P_{\vec{v}^\perp}(\vec{a})$$

□

Proof. 1.4.27 Let $\vec{a} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ be a unit vector in \mathbb{R}^3

- a. Let $T_{\vec{a}}$ be $T_{\vec{a}}(\vec{v}) = \vec{v} - 2(\vec{a} \cdot \vec{v})\vec{a}$ also let $\vec{v}, \vec{w} \in \mathbb{R}^3$ and α, β be scalars then

$$\begin{aligned} T_{\vec{a}}(\alpha\vec{v} + \beta\vec{w}) &= \alpha\vec{v} + \beta\vec{w} - 2(\vec{a} \cdot (\alpha\vec{v} + \beta\vec{w}))\vec{a} \\ &= \alpha\vec{v} + \beta\vec{w} - 2(\alpha(\vec{a} \cdot \vec{v}) + \beta(\vec{a} \cdot \vec{w}))\vec{a} \\ &= \alpha\vec{v} + \beta\vec{w} - 2\alpha(\vec{a} \cdot \vec{v})\vec{a} - 2\beta(\vec{a} \cdot \vec{w})\vec{a} \\ &= \alpha(\vec{v} - 2(\vec{a} \cdot \vec{v})\vec{a}) + \beta(\vec{w} - 2(\vec{a} \cdot \vec{w})\vec{a}) \\ &= \alpha T_{\vec{a}}(\vec{v}) + \beta T_{\vec{a}}(\vec{w}) \end{aligned}$$

Where we used that the dot product is distributive. Therefore, $T_{\vec{a}}$ is a linear transformation.

- b. $T_{\vec{a}}(\vec{a})$ by definition is

$$T_{\vec{a}}(\vec{a}) = \vec{a} - 2(\vec{a} \cdot \vec{a})\vec{a} = \vec{a} - 2|\vec{a}|^2\vec{a} = \vec{a} - 2\vec{a} = -\vec{a}$$

Where we used that $|\vec{a}|^2 = 1$ because \vec{a} is a unit vector. So $T_{\vec{a}}(\vec{a})$ returns the opposite vector to \vec{a} .

Let $\vec{v} \in \mathbb{R}^3$ be orthogonal to \vec{a} then $T_{\vec{a}}(\vec{v})$ is

$$T_{\vec{a}}(\vec{v}) = \vec{v} - 2(\vec{a} \cdot \vec{v})\vec{a} = \vec{v} - 0 = \vec{v}$$

Where we used that $\vec{a} \cdot \vec{v} = 0$ because \vec{v} is orthogonal to \vec{a} .

We can name $T_{\vec{a}}$ as the linear transformation that reflects \vec{v} across the plane orthogonal to \vec{a} .

- c. To determine the matrix M of $T_{\vec{a}}$ we need to evaluate the output of $T_{\vec{a}}$ when the input are the standard basis vectors \vec{e}_1, \vec{e}_2 and \vec{e}_3 as follows

$$T_{\vec{a}}(\vec{e}_1) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 2\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 2a \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 - 2a^2 \\ -2ab \\ -2ac \end{bmatrix}$$

$$T_{\vec{a}}(\vec{e}_2) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - 2\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - 2b \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -2ab \\ 1 - 2b^2 \\ -2bc \end{bmatrix}$$

$$T_{\vec{a}}(\vec{e}_3) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 2\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 2c \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -2ac \\ -2bc \\ 1 - 2c^2 \end{bmatrix}$$

Hence M is

$$M = \begin{bmatrix} 1 - 2a^2 & -2ab & -2ac \\ -2ab & 1 - 2b^2 & -2bc \\ -2ac & -2bc & 1 - 2c^2 \end{bmatrix}$$

Finally since M is a symmetric matrix M^2 is going to be symmetric as well.

□

Proof. 1.4.28 Let M be an $n \times m$ matrix

$$M = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1m} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2m} \\ \vdots & \vdots & & \vdots \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nm} \end{bmatrix}$$

Then, by definition

$$|M|^2 = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij}^2$$

Where μ_{ij} are the entries of the matrix M .

Now, let us focus on the diagonal of $M^T M$, we see that the entries of the diagonal are

$$\begin{aligned} \nu_{11} &= \mu_{11}^2 + \mu_{21}^2 + \dots + \mu_{n1}^2 \\ \nu_{22} &= \mu_{12}^2 + \mu_{22}^2 + \dots + \mu_{n2}^2 \\ &\vdots \\ \nu_{mm} &= \mu_{1m}^2 + \mu_{2m}^2 + \dots + \mu_{nm}^2 \end{aligned}$$

So we see that $\text{tr}(M^T M)$ is

$$\text{tr}(M^T M) = \nu_{11} + \nu_{22} + \dots + \nu_{mm} = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij}^2$$

Therefore $|M|^2 = \text{tr}(M^T M)$. □

1.5 - Limits and Continuity

Proof. 1.5.1

- a. $U = \{x \in \mathbb{R} | 0 < x \leq 1\}$ as a subset of \mathbb{R}

This set is not open because if we take $x = 1 \in U$ then for any $r > 0$ the ball $B_r(1)$ is not completely in U .

For this set to be closed must happen that $\mathbb{R} - U$ is open but

$$\mathbb{R} - U = (-\infty, 0] \cup (1, \infty)$$

which is not open because $(-\infty, 0]$ is not open. Therefore U is not closed either.

- b. $U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | \sqrt{x^2 + y^2} < 1 \right\}$ as a subset of \mathbb{R}^2

Let $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in U$, we can take r such that $r < 1 - \sqrt{x^2 + y^2}$ then the ball $B_r(\mathbf{x}) \subset U$ so U is open.

We see that $\mathbb{R}^2 - U$ is

$$\mathbb{R}^2 - U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | \sqrt{x^2 + y^2} \geq 1 \right\}$$

so if we take $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^2 - U$ we see that there is no $r > 0$ such that $B_r(\mathbf{x})$ is completely in $\mathbb{R}^2 - U$ then $\mathbb{R}^2 - U$ is not open and therefore, U is not closed.

- c. $(0, 1]$ as a subset of \mathbb{R}

This set can be written as $\{x \in \mathbb{R} | 0 < x \leq 1\}$ which is the same set we considered in part "a" so $(0, 1]$ is neither open nor closed.

- d. $U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | \sqrt{x^2 + y^2} \leq 1 \right\}$ as a subset of \mathbb{R}^2

Let $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in U$, we see that there is no $r > 0$ such that $B_r(\mathbf{x})$ is completely in U so U is not open.

We see that $\mathbb{R}^2 - U$ is defined as

$$\mathbb{R}^2 - U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 | \sqrt{x^2 + y^2} > 1 \right\}$$

Let $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 - U$, we can take r such that $r < \sqrt{x^2 + y^2} - 1$ then the ball $B_r(\mathbf{x}) \subset \mathbb{R}^2 - U$, hence $\mathbb{R}^2 - U$ is open.

Therefore U is closed.

e. $U = \{x \in \mathbb{R} | 0 \leq x \leq 1\}$ as a subset of \mathbb{R}

This set can be written as $[0, 1] \subset \mathbb{R}$ and we saw in Example 1.5.3 that any set of the form $[a, b]$ for $a, b \in \mathbb{R}$ is not open, so $[0, 1]$ is not open.

We saw also in Example 1.5.3 that any interval of the form $(-\infty, b)$ and (a, ∞) is open so $\mathbb{R} - [0, 1] = (-\infty, 0) \cup (1, \infty)$ is open. Therefore $[0, 1]$ is closed.

f. $U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \sqrt{x^2 + y^2 + z^2} \leq 1, \text{ and } x, y, z \neq 0 \right\}$ as a subset of \mathbb{R}^3

Let $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in U$, we see that there is no $r > 0$ such that $B_r(\mathbf{x})$ is completely in U so U is not open.

Let $\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^3 - U$, we see that there is no $r > 0$ such that $B_r(\mathbf{x})$

is completely in $\mathbb{R}^3 - U$ so $\mathbb{R}^3 - U$ is not open and therefore U is not closed.

g. the empty set \emptyset as a subset of \mathbb{R}

An open set must contain a ball around each of its points. Since there are no points in \emptyset , this condition is satisfied vacuously, and thus \emptyset is open.

We see that $\mathbb{R} - \emptyset = \mathbb{R}$ and \mathbb{R} is open, then \emptyset is closed.

□

Proof. 1.5.2

a. (x, y) -plane in \mathbb{R}^3

Let $\mathbf{x} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in (x, y)$ -plane, then for $r > 0$ we know that the ball around \mathbf{x} is defined as

$$B_r(\mathbf{x}) = \left\{ \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \in \mathbb{R}^3 \text{ such that } \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} < r \right\}$$

so the ball $B_r(\mathbf{x})$ have values outside the (x, y) -plane and hence the (x, y) -plane is not open in \mathbb{R}^3 .

Let $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 - (x, y)$ -plane i.e. $z \neq 0$. We see that taking $r < |z|$

then the ball $B_r(\mathbf{x})$ is inside $\mathbb{R}^3 - (x, y)$ -plane, hence $\mathbb{R}^3 - (x, y)$ -plane is open and therefore the (x, y) -plane is closed.

b. $\mathbb{R} \subset \mathbb{C}$

Let $x + iy \in \mathbb{R}$ then for $r > 0$ we know that the ball around $x + iy$ is defined as

$$B_r(x + iy) = \{x' + iy' \in \mathbb{C} \text{ such that } \sqrt{(x - x')^2 + (0 - y')^2} < r\}$$

so the ball $B_r(x + iy)$ contains points in \mathbb{C} with both real and imaginary values. Hence \mathbb{R} is not open in \mathbb{C} .

Let $x + iy \in \mathbb{C} - \mathbb{R}$ i.e. $y \neq 0$. Taking $r < |y|$ the ball $B_r(x + iy)$ is inside $\mathbb{C} - \mathbb{R}$, hence $\mathbb{C} - \mathbb{R}$ is open and therefore \mathbb{R} is closed in \mathbb{C} .

c. the line $x = 5$ in the (x, y) -plane

Let $(5, y_0)$ be any point in the line $x = 5$ then for $r > 0$ we know that the ball around $(5, y_0)$ is defined as

$$B_r(5, y_0) = \{(x, y) \in \mathbb{R}^2 \text{ such that } \sqrt{(5 - x)^2 + (y_0 - y)^2} < r\}$$

so the ball $B_r(5, y_0)$ contains points in \mathbb{R}^2 which are not in the line $x = 5$ i.e. they have $x \neq 5$ then the line $x = 5$ is not open in \mathbb{R}^2 .

Let $(x_0, y_0) \in \mathbb{R}^2 - \{(x, y) \in \mathbb{R}^2 | x = 5\}$ i.e. $x_0 \neq 5$.

Taking $r < |x_0 - 5|$ the ball $B_r(x_0, y_0)$ is inside $\mathbb{R}^2 - \{(x, y) \in \mathbb{R}^2 | x = 5\}$, hence $\mathbb{R}^2 - \{(x, y) \in \mathbb{R}^2 | x = 5\}$ is open and therefore the line $x = 5$ is closed in \mathbb{R}^2 .

d. $(0, 1) \subset \mathbb{C}$

Let $x_0 + i0 \in (0, 1)$ i.e. $x_0 \in (0, 1)$ then for $r > 0$ we know that the ball around $x_0 + i0$ is defined as

$$B_r(x_0 + i0) = \{x' + iy' \in \mathbb{C} \text{ such that } \sqrt{(x_0 - x')^2 + (0 - y')^2} < r\}$$

so the ball $B_r(x_0 + i0)$ contains points in \mathbb{C} with both real and imaginary values. Hence $(0, 1)$ is not open in \mathbb{C} .

From part (b) we know that $\mathbb{C} - \mathbb{R}$ is open so we need to check only what happens for points like $x_0 + i0$ where $x_0 \notin (0, 1)$. Suppose $x_0 = 1$ then there is no $r > 0$ such that $B_r(x_0 + i0) \subset \mathbb{C} - (0, 1)$ then $\mathbb{C} - (0, 1)$ is not open and therefore $(0, 1)$ is not closed.

e. $\mathbb{R}^n \subset \mathbb{R}^n$

Let $\mathbf{x} \in \mathbb{R}^n$ then a ball $B_r(\mathbf{x}) \subset \mathbb{R}^n$ therefore \mathbb{R}^n is open in \mathbb{R}^n .

We see that $\mathbb{R}^n - \mathbb{R}^n = \emptyset$ and we saw that \emptyset is open so \mathbb{R}^n is closed.

f. the unit sphere in \mathbb{R}^3 or $U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \sqrt{x^2 + y^2 + z^2} = 1 \right\}$

Let $\mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \in U$, then for $r > 0$ we know that the ball around \mathbf{x}_0 is defined as

$$B_r(\mathbf{x}_0) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \text{ such that } \sqrt{(x_0 - x)^2 + (y_0 - y)^2 + (z_0 - z)^2} < r \right\}$$

so the ball $B_r(\mathbf{x}_0)$ have values outside U and hence U is not open in \mathbb{R}^3 .

Let $\mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \in \mathbb{R}^3 - U$, then taking $r < \sqrt{(x_0 - x)^2 + (y_0 - y)^2 + (z_0 - z)^2}$

the ball $B_r(\mathbf{x}_0)$ is in $\mathbb{R}^3 - U$, hence $\mathbb{R}^3 - U$ is open and therefore U is closed in \mathbb{R}^3 .

□

Proof. 1.5.3

- a. Let us define two open sets U_1 and U_2 in \mathbb{R}^n , let $\mathbf{x} \in U_1 \cup U_2$ then if $\mathbf{x} \in U_1$ then there is a ball $B_r^{U_1}(\mathbf{x})$ which is contained in U_1 since U_1 is an open set. In the same way, if $\mathbf{x} \in U_2$ then there is a ball $B_r^{U_2}(\mathbf{x})$ contained in U_2 . Therefore $U_1 \cup U_2$ is an open set.

If we have the union of infinitely many open sets $U_1 \cup U_2 \cup \dots$ the same process can be applied. Therefore the union of open sets is open.

- b. Let us define two open sets U_1 and U_2 in \mathbb{R}^n , let $\mathbf{x} \in U_1 \cap U_2$ then there is a ball $B_{r_1}^{U_1}(\mathbf{x})$ which is contained in U_1 and there is a ball $B_{r_2}^{U_2}(\mathbf{x})$ which is contained in U_2 , since both U_1 and U_2 are open sets. Now, we select a radius $r = \min(r_1, r_2)$ then the ball $B_r(\mathbf{x})$ is contained in $U_1 \cap U_2$. Therefore $U_1 \cap U_2$ is open.

By the same procedure, the same is true for a finite number of open sets.

- c. Let us consider an infinite number of open sets of the form $U_n = (-1/n, 1/n) \subset \mathbb{R}$ where $n \in \mathbb{N}$.

We see that $U_1 \cap U_2 \cap U_3 \cap \dots = \{0\}$ but $\{0\}$ is not open.

Therefore an infinite intersection of open sets is not necessarily open.

□

Proof. 1.5.4

- a. Let B be the biggest open set contained in A , and suppose $B \not\subseteq \overset{\circ}{A}$. We want to arrive at a contradiction.

Let $\mathbf{x} \in B$, then by the definition of open set, there must be $r > 0$ such that $B_r(\mathbf{x}) \subset B$ also, $B_r(\mathbf{x}) \subset A$ but by defintion of interior must be that $B_r(\mathbf{x}) \subset \overset{\circ}{A}$ hence $B \subseteq \overset{\circ}{A}$, a contradiction. Therefore $\overset{\circ}{A}$ is the biggest open set contained in A .

- b. Let C be a closed set such that $A \subset C$. Let $\mathbf{x} \in \overline{A}$ but $\mathbf{x} \notin C$. We want to arrive at a contradiction.

Since C is closed then $\mathbb{R}^n \setminus C$ is open. Then by defintion of open set there is $r' > 0$ such that $B_{r'}(\mathbf{x}) \in \mathbb{R}^n \setminus C$ but also since $\mathbf{x} \in \overline{A}$ for all $r > 0$ must be that $B_r(\mathbf{x}) \cap A \neq \emptyset$. Therefore $B_{r'}(\mathbf{x})$ must intersect A so we have a contradiction and must be that $\mathbf{x} \in C$ and hence $C \subseteq \overline{A}$ which implies that \overline{A} is the smallest closed set containing A .

- c. By defintion $A \subset \overline{A}$ also if $\mathbf{x} \in \partial A$ then for all $r > 0$ we have that $B_r(\mathbf{x}) \cap A \neq \emptyset$ then $A \cup \partial A \subseteq \overline{A}$.

On the other hand, if $\mathbf{x} \in \overline{A}$ then might be that $\mathbf{x} \in A$ as well.

But suppose $\mathbf{x} \notin A$ then by definition of closure for all $r > 0$ we have that $B_r(\mathbf{x}) \cap A \neq \emptyset$.

Also $B_r(\mathbf{x}) \cap \mathbb{R}^n \setminus A \neq \emptyset$ since at least \mathbf{x} is not in A hence $\mathbf{x} \in \partial A$ this implies that $\overline{A} \subseteq A \cup \partial A$. Therefore joining both results $\overline{A} = A \cup \partial A$.

- d. Let $\mathbf{x} \in \overline{A} - \overset{\circ}{A}$, since $\mathbf{x} \in \overline{A}$, by the definition, for all $r > 0$ we have that $B_r(\mathbf{x}) \cap A \neq \emptyset$. Also, since $\mathbf{x} \notin \overset{\circ}{A}$ for all $r > 0$ must happen that $B_r(\mathbf{x}) \cap \mathbb{R}^n \setminus A \neq \emptyset$. Hence $\mathbf{x} \in \partial A$ and $\overline{A} - \overset{\circ}{A} \subseteq \partial A$.

Let now, $\mathbf{x} \in \partial A$. Then by definition for all $r > 0$ we have that $B_r(\mathbf{x}) \cap A \neq \emptyset$ so $\mathbf{x} \in \overline{A}$. Also by definition, for all $r > 0$ we have that $B_r(\mathbf{x}) \cap \mathbb{R}^n \setminus A \neq \emptyset$, then $\mathbf{x} \notin \overset{\circ}{A}$. Therefore $\mathbf{x} \in \overline{A} - \overset{\circ}{A}$ so $\partial A \subseteq \overline{A} - \overset{\circ}{A}$ and joining both results we get that $\partial A = \overline{A} - \overset{\circ}{A}$.

□

Proof. 1.5.5

a. $U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2 \right\}$

Let $\mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in U$, we can take r such that

$$r < \min \left(\sqrt{x_0^2 + y_0^2} - 1, 2 - \sqrt{x_0^2 + y_0^2} \right)$$

then the ball $B_r(\mathbf{x}_0) \subset U$ so U is open.

We see that $\mathbb{R}^2 - U$ is

$$\mathbb{R}^2 - U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 \geq 2 \right\} \cup \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1 \right\}$$

so if we take $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^2 - U$ we see that there is no $r > 0$ such that $B_r(\mathbf{x}_0)$ is completely in $\mathbb{R}^2 - U$ then $\mathbb{R}^2 - U$ is not open and therefore, U is not closed.

b. $U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid xy \neq 0 \right\}$

We see that U can be described as \mathbb{R}^2 minus the x and y axes.

Let $\mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in U$, we can take r such that $r < \min(|x_0|, |y_0|)$ then the ball $B_r(\mathbf{x}_0) \subset U$ so U is open.

We see that $\mathbb{R}^2 - U$ is the union of the x and y axes, but neither of which is open in \mathbb{R}^2 so $\mathbb{R}^2 - U$ is not open and hence U is not closed.

c. $U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 0 \right\}$

In this case, U can be described as the x -axis over \mathbb{R}^2 .

We saw in problem 1.5.2-c that the line $x = 5$ over \mathbb{R}^2 is not open but closed over \mathbb{R}^2 .

This is essentially the same problem but for the line $y = 0$, and hence the same result applies.

d. $\{\mathbb{Q} \subset \mathbb{R}\}$

Let $q \in \mathbb{Q}$ then for any $r > 0$ the ball $B_r(q) \subset \mathbb{R}$ contains an irrational number. Therefore \mathbb{Q} is not open.

Let $x \in \mathbb{R} - \mathbb{Q}$ then for any $r > 0$ the ball $B_r(x) \subset \mathbb{R}$ contains a rational number. Therefore $\mathbb{R} - \mathbb{Q}$ is not open and hence \mathbb{Q} is not closed.

□

Proof. 1.5.6 Let $A \subset \mathbb{R}^n$ and let $\mathbf{x} \in \partial A$ then by definition for all $r > 0$ we have that $B_r(\mathbf{x}) \cap A \neq \emptyset$ and that $B_r(\mathbf{x}) \cap \mathbb{R}^n \setminus A \neq \emptyset$. This implies by the definition of closure that $\mathbf{x} \in \overline{A}$ but also that $\mathbf{x} \in \overline{\mathbb{R}^n \setminus A}$, hence $\partial A \subseteq \overline{A} \cap \overline{\mathbb{R}^n \setminus A}$.

On the other hand, let $\mathbf{x} \in \overline{A} \cap \overline{\mathbb{R}^n \setminus A}$, then by definition for all $r > 0$ we have that $B_r(\mathbf{x}) \cap A \neq \emptyset$ and that $B_r(\mathbf{x}) \cap \mathbb{R}^n \setminus A \neq \emptyset$ but this implies that $\mathbf{x} \in \partial A$, so $\overline{A} \cap \overline{\mathbb{R}^n \setminus A} \subseteq \partial A$.

Therefore joining both results we get that $\partial A = \overline{A} \cap \overline{\mathbb{R}^n \setminus A}$. \square

Proof. 1.5.7

- a. The natural domain of $\sin 1/xy$ is \mathbb{R}^2 minus the x -axis and the y -axis since neither x nor y can be 0.

The natural domain in this case is open but not closed.

- b. The natural domain of $\ln \sqrt{x^2 - y}$ are all x and y such that $x > 0$ and $0 \leq y < x^2$.

Given that the x -axis is included and any ball there will include points that are not in the natural domain we can conclude it's not open. Also, it's not closed because the side $y < x^2$ does not contain all its boundary points.

- c. The natural domain of $\ln \ln x$ is $(1, \infty)$, then the natural domain is open but not closed.
- d. The natural domain of $\arcsin x$ is $[-1, 1]$ so this implies that for our function $\arcsin 3/(x^2 + y^2)$ must happen that $x^2 + y^2 \geq -3$ or that $x^2 + y^2 \geq 3$. Since $x^2 + y^2$ is always positive the first constraint can be dropped and therefore the natural domain of our function is

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \geq 3\}$$

The natural domain is therefore closed and not open.

- e. For the function $\sqrt{e^{\cos xy}}$ we see that $e^{\cos xy}$ is always positive for every $(x, y) \in \mathbb{R}^2$, so the natural domain of $\sqrt{e^{\cos xy}}$ is \mathbb{R}^2 .

Therefore the natural domain of this function is open and closed.

- f. The natural domain of $1/xyz$ is $(\mathbb{R} \setminus \{0\}) \times (\mathbb{R} \setminus \{0\}) \times (\mathbb{R} \setminus \{0\})$ i.e \mathbb{R}^3 minus the planes $x = 0$, $y = 0$ and $z = 0$ since x, y and z cannot be 0.

Therefore the natural domain of this function is open but not closed.

\square

Proof. 1.5.8

a. Let

$$A = \begin{bmatrix} 0 & -\epsilon & -\epsilon \\ 0 & 0 & -\epsilon \\ 0 & 0 & 0 \end{bmatrix}$$

Then we see that

$$I - A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -\epsilon & -\epsilon \\ 0 & 0 & -\epsilon \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & \epsilon & \epsilon \\ 0 & 1 & \epsilon \\ 0 & 0 & 1 \end{bmatrix} = B$$

Now, let us compute a few powers of A

$$\begin{aligned} A^2 &= \begin{bmatrix} 0 & -\epsilon & -\epsilon \\ 0 & 0 & -\epsilon \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -\epsilon & -\epsilon \\ 0 & 0 & -\epsilon \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \epsilon^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ A^3 &= \begin{bmatrix} 0 & 0 & \epsilon^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -\epsilon & -\epsilon \\ 0 & 0 & -\epsilon \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

On the other hand, we know that $S = I + A + A^2 + A^3 + \dots$ converges to $(I - A)^{-1} = B^{-1}$ but since $A^3 = 0$ then S becomes $S = I + A + A^2$ and it's given by

$$\begin{aligned} S &= I + A + A^2 \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -\epsilon & -\epsilon \\ 0 & 0 & -\epsilon \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \epsilon^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -\epsilon & \epsilon^2 - \epsilon \\ 0 & 1 & -\epsilon \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

b. Let A be

$$A = \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix}$$

Then

$$I - A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix} = \begin{bmatrix} 1 & -\epsilon \\ \epsilon & 1 \end{bmatrix} = C$$

Now, let us compute a few powers of A

$$\begin{aligned} A^2 &= \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix} = \begin{bmatrix} -\epsilon^2 & 0 \\ 0 & -\epsilon^2 \end{bmatrix} = -\epsilon^2 I \\ A^3 &= \begin{bmatrix} -\epsilon^2 & 0 \\ 0 & -\epsilon^2 \end{bmatrix} \cdot \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\epsilon^3 \\ \epsilon^3 & 0 \end{bmatrix} = -\epsilon^2 A \\ A^4 &= \begin{bmatrix} 0 & -\epsilon^3 \\ \epsilon^3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix} = \begin{bmatrix} \epsilon^4 & 0 \\ 0 & \epsilon^4 \end{bmatrix} = \epsilon^4 I \\ A^5 &= \begin{bmatrix} \epsilon^4 & 0 \\ 0 & \epsilon^4 \end{bmatrix} \cdot \begin{bmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{bmatrix} = \begin{bmatrix} 0 & \epsilon^5 \\ -\epsilon^5 & 0 \end{bmatrix} = \epsilon^4 A \end{aligned}$$

Then we can simplify $S = I + A + A^2 + A^3 + \dots$ as follows

$$\begin{aligned} S &= I + A + A^2 + A^3 + A^4 + A^5 + \dots \\ &= I + A - \epsilon^2 I - \epsilon^2 A + \epsilon^4 I + \epsilon^4 A + \dots \\ &= (1 - \epsilon^2 + \epsilon^4 - \dots)I + (1 - \epsilon^2 + \epsilon^4 - \dots)A \\ &= \frac{1}{1 + \epsilon^2}(I + A) \end{aligned}$$

Where we used that by the geometric series we see that $1 - \epsilon^2 + \epsilon^4 - \dots = 1/(1 + \epsilon^2)$.

Therefore the inverse of C is

$$C^{-1} = \frac{1}{1 + \epsilon^2}(I + A) = \begin{bmatrix} \frac{1}{1+\epsilon^2} & \frac{\epsilon}{1+\epsilon^2} \\ -\frac{\epsilon}{1+\epsilon^2} & \frac{1}{1+\epsilon^2} \end{bmatrix}$$

□

Proof. 1.5.9 Let $\sum_{i=1}^{\infty} \mathbf{x}_i$ be a convergent series in \mathbb{R}^n .

If $\sum_{i=1}^{\infty} |\mathbf{x}_i|$ does not converge then the triangle inequality is true since

$$\left| \sum_{i=1}^{\infty} \mathbf{x}_i \right| < \infty = \sum_{i=1}^{\infty} |\mathbf{x}_i|$$

Then let us suppose $\sum_{i=1}^{\infty} |\mathbf{x}_i|$ converges. We know that in \mathbb{R}^n the triangle inequality holds for finite sums so we can write that

$$\left| \sum_{i=1}^m \mathbf{x}_i \right| \leq \sum_{i=1}^m |\mathbf{x}_i| \leq \sum_{i=1}^{\infty} |\mathbf{x}_i|$$

Now, let us take the limit as m goes to infinity

$$\lim_{m \rightarrow \infty} \left| \sum_{i=1}^m \mathbf{x}_i \right| \leq \sum_{i=1}^{\infty} |\mathbf{x}_i|$$

Because the continuity of the norm $|\cdot|$ and the convergence of $(\sum_{i=1}^n \mathbf{x}_i)$ we get that

$$\lim_{m \rightarrow \infty} \left| \sum_{i=1}^m \mathbf{x}_i \right| = \left| \sum_{i=1}^{\infty} \mathbf{x}_i \right|$$

Therefore

$$\left| \sum_{i=1}^{\infty} \mathbf{x}_i \right| \leq \sum_{i=1}^{\infty} |\mathbf{x}_i|$$

□

Proof. 1.5.11

Let $\varphi : (0, \infty) \rightarrow (0, \infty)$ be a function such that $\lim_{\epsilon \rightarrow 0} \varphi(\epsilon) = 0$.

- a. Let $i \rightarrow \mathbf{a}_i$ be a convergent sequence that converges to \mathbf{a} in \mathbb{R}^n .

Let $\varphi(\epsilon) = \epsilon' > 0$ then we know that there is $N \in \mathbb{N}$ such that when $n \geq N$ we have that $|\mathbf{a}_i - \mathbf{a}| < \epsilon' = \varphi(\epsilon)$.

On the other hand, if given $\epsilon > 0$ there is $N \in \mathbb{N}$ such that for $n > N$, we have that $|\mathbf{a}_i - \mathbf{a}| < \varphi(\epsilon)$. Then by Proposition 1.5.14 (1) we have that the sequence $i \rightarrow \mathbf{a}_i$ converges to \mathbf{a} .

- b. The limit $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ exists if and only if for any $\epsilon > 0$ there is some $\delta > 0$ such that when $|\mathbf{x} - \mathbf{x}_0| < \delta$ we have that $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \varphi(\epsilon)$.

□

Proof. 1.5.12

Let u be a strictly positive function of $\epsilon > 0$, such that $u(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$, and let U be a subset of \mathbb{R}^n .

1. A function $f : U \rightarrow \mathbb{R}$ has the limit a at \mathbf{x}_0 if $\mathbf{x}_0 \in \overline{U}$ and if for all $\epsilon > 0$ there exists $\delta > 0$ such that when $|\mathbf{x} - \mathbf{x}_0| < \delta$ and $\mathbf{x} \in U$ then $|f(\mathbf{x}) - a| < \epsilon$.

Let $\epsilon > 0$, then if we take $\epsilon' = u(\epsilon)$ since the function has the limit a at \mathbf{x}_0 then we have that there exists some $\delta' > 0$ such that when $|\mathbf{x} - \mathbf{x}_0| < \delta'$ and $\mathbf{x} \in U$ then $|f(\mathbf{x}) - a| < \epsilon' = u(\epsilon)$. Therefore point 1 is equivalent to point 2.

2. On the other hand, the second point states that: A function $f : U \rightarrow \mathbb{R}$ has the limit a at \mathbf{x}_0 if $\mathbf{x}_0 \in \overline{U}$ and if for all $\epsilon > 0$ there exists $\delta > 0$ such that when $|\mathbf{x} - \mathbf{x}_0| < \delta$ and $\mathbf{x} \in U$ then $|f(\mathbf{x}) - a| < u(\epsilon)$.

Let $\epsilon > 0$, then if we name $u(\epsilon) = \epsilon'$ since the function has the limit a at \mathbf{x}_0 then we have that there exists some $\delta' > 0$ such that when $|\mathbf{x} - \mathbf{x}_0| < \delta'$ and $\mathbf{x} \in U$ then $|f(\mathbf{x}) - a| < u(\epsilon) = \epsilon'$. Therefore point 2 is equivalent to point 1.

Finally, joining both points, we see that both definitions are equivalent. \square

Proof. 1.5.13

Let every convergent sequence in $C \subset \mathbb{R}^n$ converge to a point in C .

By definition we know that $C \subseteq \overline{C}$.

On the other hand, suppose there is a point $\mathbf{c} \in \overline{C}$ such that $\mathbf{c} \notin C$ we want to arrive at a contradiction. By definition we know that $B_\epsilon(\mathbf{c}) \cap C \neq \emptyset$ for every $\epsilon > 0$. So we can take a sequence $\{\mathbf{c}_n\}$ in C such that $\mathbf{c}_n \in B_{1/n}(\mathbf{c})$ for all $n \in \mathbb{N}$ then $|\mathbf{c}_n - \mathbf{c}| < 1/n$ which implies that the sequence $i \rightarrow \mathbf{c}_i$ converges to \mathbf{c} as $n \rightarrow \infty$ so \mathbf{c} must be in C , a contradiction.

Therefore $\overline{C} \subseteq C$ but joining the previous result we see that $\overline{C} = C$ and hence C is closed. \square

Proof. 1.5.14

$$\text{a. } \lim_{(x,y) \rightarrow (1,2)} \frac{x^2}{x+y}$$

Let $U = (0, 2) \subset \mathbb{R}$ and $V = (1, 3) \subset \mathbb{R}$ then the function x^2 is continuous over U and $1/(x+y)$ is continuous over $U \times V$ then because of the product of continuous functions is continuous we have that $x^2/(x+y)$ is continuous over $U \times V$ and hence since $(1, 2) \in U \times V$ we get that

$$\lim_{(x,y) \rightarrow (1,2)} \frac{x^2}{x+y} = \frac{1^2}{1+2} = \frac{1}{3}$$

$$\text{b. } \lim_{(x,y) \rightarrow (0,0)} \frac{\sqrt{|xy|}}{x^2+y^2}$$

Let us consider the sequences $(x_n, y_n) = (1/n, 0)$ which converge to $(0, 0)$ then we see that

$$\frac{\sqrt{|x_n|}y_n}{x_n^2+y_n^2} = \frac{\sqrt{|1/n|} \cdot 0}{(1/n)^2 + 0^2} = 0$$

But if we consider the sequences $(x_n, y_n) = (1/n, 1/n)$ which also converge to $(0, 0)$ we get that

$$\frac{\sqrt{|x_n|}y_n}{x_n^2+y_n^2} = \frac{\sqrt{|1/n|} \cdot 1/n}{(1/n)^2 + (1/n)^2} = \frac{1/\sqrt{n} \cdot 1/n}{2/n^2} = \frac{n}{2\sqrt{n}} = \frac{\sqrt{n}}{2}$$

And $\sqrt{n} \neq 0$ for $n \neq 0$. Therefore the limit as $(x, y) \rightarrow (0, 0)$ of $\sqrt{|xy|}/(x^2+y^2)$ doesn't exist.

$$\text{c. } \lim_{(x,y) \rightarrow (0,0)} \frac{\sqrt{|xy|}}{\sqrt{x^2+y^2}}$$

Let us consider the sequences $(x_n, y_n) = (1/n, 0)$ which converge to $(0, 0)$ then we see that

$$\frac{\sqrt{|x_n y_n|}}{\sqrt{x_n^2+y_n^2}} = \frac{\sqrt{|1/n \cdot 0|}}{\sqrt{(1/n)^2 + 0^2}} = 0$$

But if we consider the sequences $(x_n, y_n) = (1/n, 1/n)$ which also converge to $(0, 0)$ we get that

$$\frac{\sqrt{|x_n y_n|}}{\sqrt{x_n^2+y_n^2}} = \frac{\sqrt{|1/n \cdot 1/n|}}{\sqrt{(1/n)^2 + (1/n)^2}} = \frac{\sqrt{1/n^2}}{\sqrt{2/n^2}} = \frac{1/n}{\sqrt{2}/n} = \frac{1}{\sqrt{2}}$$

Therefore the limit as $(x, y) \rightarrow (0, 0)$ of $\sqrt{|xy|}/\sqrt{x^2+y^2}$ doesn't exist.

d. $\lim_{(x,y) \rightarrow (1,2)} x^2 + y^3 - 3$

Let $U = (0, 2) \subset \mathbb{R}$ and $V = (1, 3) \subset \mathbb{R}$ then the function x^2 is continuous over U and $y^3 - 3$ is continuous over V then because the sum of continuous functions is continuous we have that $x^2 + y^3 - 3$ is continuous over $U \times V$ and hence since $(1, 2) \in U \times V$ we get that

$$\lim_{(x,y) \rightarrow (1,2)} x^2 + y^3 - 3 = 1^2 + 2^3 - 3 = 6$$

□

Proof. 1.5.16

- a. Let $D^* \subset \mathbb{R}^2$ be the region $0 < x^2 + y^2 < 1$, and let $f : D^* \rightarrow \mathbb{R}$ be a function. Then the assertion

$$\lim_{\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}} f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = a$$

implies that the limit of f as we approach $(0, 0)$ exists and is equal to a . Also, since the function is not defined at $(0, 0)$ we could extend the domain defining that $f(0, 0) = a$.

- b. Let us consider the limit as $(x, y) \rightarrow (0, 0)$ of the function

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \frac{\sin(x+y)}{\sqrt{x^2+y^2}}$$

Let us consider the sequences $(x_n, y_n) = (1/n, 0)$ which converge to $(0, 0)$ then we see that

$$\frac{\sin(1/n+0)}{\sqrt{1/n^2+0^2}} = \frac{\sin(1/n)}{1/n} = n \sin(1/n)$$

So as $n \rightarrow \infty$ we see that $n \sin(1/n) \rightarrow 1$.

But if we consider the sequences $(x_n, y_n) = (1/n, 1/n)$ which also converge to $(0, 0)$ we get that

$$\frac{\sin(1/n+1/n)}{\sqrt{1/n^2+1/n^2}} = \frac{\sin(2/n)}{\sqrt{2} \cdot 1/n} = \frac{n \sin(2/n)}{\sqrt{2}}$$

In this case, we see that as $n \rightarrow \infty$ we get that $n \sin(2/n)/\sqrt{2} \rightarrow \sqrt{2}$. Therefore the limit as $(x, y) \rightarrow (0, 0)$ of $\sin(x+y)/\sqrt{x^2+y^2}$ doesn't exist.

Now, let us consider the limit as $(x, y) \rightarrow (0, 0)$ of the function

$$g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = (|x| + |y|) \ln(x^2 + y^4)$$

Let us consider the region such that $0 < x^2 + y^4 < 1$, in this region $\ln(x^2 + y^4)$ is negative and hence

$$(|x| + |y|) \ln(x^2 + y^4) < 0$$

On the other hand, we know that $-1/a^n < \ln(a)$ when $a > 0$ and $n > 0$ so we can write that

$$-\frac{(|x| + |y|)}{(x^2 + y^4)^{1/5}} < (|x| + |y|) \ln(x^2 + y^4) < 0$$

Also, we see that

$$0 < \frac{(|x| + |y|)}{(x^2 + y^4)^{1/5}} = \frac{|x|}{(x^2 + y^4)^{1/5}} + \frac{|y|}{(x^2 + y^4)^{1/5}}$$

Let us analyse each term on the RHS, we see that $x^2 + y^4 \geq x^2$ so $(x^2 + y^4)^{1/5} \geq (x^2)^{1/5}$ hence

$$0 < \frac{|x|}{(x^2 + y^4)^{1/5}} \leq \frac{|x|}{(x^2)^{1/5}} = |x|^{3/5} \rightarrow 0 \quad \text{as } x \rightarrow 0$$

But also we see that $(x^2 + y^4)^{1/5} \geq (y^4)^{1/5}$ so in the same way

$$0 < \frac{|y|}{(x^2 + y^4)^{1/5}} \leq \frac{|y|}{(y^4)^{1/5}} = |y|^{1/5} \rightarrow 0 \quad \text{as } y \rightarrow 0$$

Therefore using the squeeze theorem we see that

$$\frac{(|x| + |y|)}{(x^2 + y^4)^{1/5}} \rightarrow 0 \quad \text{as } (x, y) \rightarrow (0, 0)$$

And hence using the squeeze theorem again we get that

$$(|x| + |y|) \ln(x^2 + y^4) \rightarrow 0 \quad \text{as } (x, y) \rightarrow (0, 0)$$

□

Proof. 1.5.17

1. Let $i \rightarrow \mathbf{a}_i$ and $i \rightarrow \mathbf{b}_i$ be two convergent sequences of points in \mathbb{R}^n . Let us take the j th component $(a_i)_j$ of \mathbf{a}_i and $(b_i)_j$ in \mathbf{b}_i that converge to a_j in \mathbf{a} and b_j in \mathbf{b} respectively. Then we know that given $\epsilon/2 > 0$ there is N_a and N_b such that when $n \geq N_a$ we have that

$$|(a_n)_j - a_j| < \epsilon/2$$

and when $n \geq N_b$ we have that

$$|(b_n)_j - b_j| < \epsilon/2$$

Let us take $N = \max\{N_a, N_b\}$ so when $n \geq N$ we have that

$$|((a_n)_j - a_j) + ((b_n)_j - b_j)| \leq |(a_n)_j - a_j| + |(b_n)_j - b_j| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

But we can write that

$$|((a_n)_j - a_j) + ((b_n)_j - b_j)| = |((a_n)_j + (b_n)_j) - (a_j + b_j)|$$

so

$$|((a_n)_j + (b_n)_j) - (a_j + b_j)| < \epsilon$$

This implies that $\lim_{i \rightarrow \infty} ((a_i)_j + (b_i)_j) = a_j + b_j$ but by Proposition (1.5.13) we know that the convergence of the coordinates implies the convergence of the sequence, therefore

$$\lim_{i \rightarrow \infty} (\mathbf{a}_i + \mathbf{b}_i) = \mathbf{a} + \mathbf{b} = \lim_{i \rightarrow \infty} \mathbf{a}_i + \lim_{i \rightarrow \infty} \mathbf{b}_i$$

2. Let $i \rightarrow \mathbf{a}_i$ and $i \rightarrow c_i$ be two convergent sequences in \mathbb{R}^n and \mathbb{R} respectively. Let us take the j th component $(a_i)_j$ of \mathbf{a}_i that converge to a_j in \mathbf{a} .

Since $c_i \rightarrow c$ let us take $\epsilon = 1$ then there is N_1 such that when $n \geq N_1$ we have that $|c_n - c| < 1$, also

$$|c_n| = |c_n - c + c| \leq |c_n - c| + |c| < 1 + |c|$$

Hence $|c_n| \leq |c| + 1$ when $n \geq N_1$.

Let us fix $\epsilon > 0$ and let us define

$$\epsilon' = \frac{\epsilon}{1 + |c| + |a_j|} > 0$$

Then there is N_a such that when $n \geq N_a$ we have that

$$|(a_n)_j - a_j| < \epsilon'$$

Also, there is some N_c such that $|c_n - c| < \epsilon'$ when $n \geq N_c$.

So, if we take $N = \max\{N_a, N_c, N_1\}$ when $n \geq N$ we have that

$$\begin{aligned} |(a_n)_j c_n - a_j c| &= |(a_n)_j c_n - a_j c + a_j c_n - a_j c_n| \\ &= |c_n((a_n)_j - a_j) + a_j(c_n - c)| \\ &\leq |c_n| |(a_n)_j - a_j| + |a_j| |c_n - c| \\ &< |c_n| \epsilon' + |a_j| \epsilon' \\ &< (1 + |c|) \epsilon' + |a_j| \epsilon' \\ &< \epsilon' ((1 + |c|) + |a_j|) \\ &< \epsilon \end{aligned}$$

Therefore this implies that $\lim_{i \rightarrow \infty} (a_i)_j c_i = a_j c$. Since this holds for every component j of \mathbf{a}_i then the sequence $\mathbf{a}_i c_i$ converges to \mathbf{ac} i.e.

$$\lim_{i \rightarrow \infty} c_i \mathbf{a}_i = \mathbf{ac} = \lim_{i \rightarrow \infty} \mathbf{a}_i \cdot \lim_{i \rightarrow \infty} c_i$$

3. Let $i \rightarrow \mathbf{a}_i$ and $i \rightarrow \mathbf{b}_i$ be two convergent sequences of points in \mathbb{R}^m . We can also see these sequences as sequences of vectors in \mathbb{R}^m i.e. $i \rightarrow \vec{\mathbf{a}}_i$ and $i \rightarrow \vec{\mathbf{b}}_i$.

Let us take the j th component $(a_i)_j$ of $\vec{\mathbf{a}}_i$ and $(b_i)_j$ in $\vec{\mathbf{b}}_i$ that converge to a_j in $\vec{\mathbf{a}}$ and b_j in $\vec{\mathbf{b}}$ respectively.

Then from part (2) we know that given $\epsilon/m > 0$ (m is the dimension) there is N_j such that when $n \geq N_j$ we have that

$$|(a_n)_j (b_n)_j - a_j b_j| < \frac{\epsilon}{m}$$

i.e. $\lim_{i \rightarrow \infty} (a_i)_j (b_i)_j = a_j b_j$.

Since we can do this for every component of \mathbf{a}_i and \mathbf{b}_i taking $N = \max\{N_1, N_2, \dots, N_m\}$ when $n \geq N$ the following inequalities hold

$$\begin{aligned} &|((a_n)_1 (b_n)_1 + (a_n)_2 (b_n)_2 + \dots + (a_n)_m (b_n)_m) - (a_1 b_1 + a_2 b_2 + \dots + a_m b_m)| \\ &= |((a_n)_1 (b_n)_1 - a_1 b_1) + ((a_n)_2 (b_n)_2 - a_2 b_2) + \dots + ((a_n)_m (b_n)_m - a_m b_m)| \\ &\leq |(a_n)_1 (b_n)_1 - a_1 b_1| + |(a_n)_2 (b_n)_2 - a_2 b_2| + \dots + |(a_n)_m (b_n)_m - a_m b_m| \\ &< \epsilon/m + \epsilon/m + \dots + \epsilon/m = \epsilon \end{aligned}$$

Therefore

$$\lim_{i \rightarrow \infty} \vec{\mathbf{a}}_i \cdot \vec{\mathbf{b}}_i = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = \lim_{i \rightarrow \infty} \vec{\mathbf{a}}_i \cdot \lim_{i \rightarrow \infty} \vec{\mathbf{b}}_i$$

4. Let $i \rightarrow \mathbf{a}_i$ be a bounded sequence in \mathbb{R}^m and $i \rightarrow c_i$ be a sequence that converge to 0 in \mathbb{R} .

Let us take the j th component $(a_i)_j$ of \mathbf{a}_i that is bounded by some number M_j so $|(a_i)_j| < M_j$ for every i .

Also, let $\epsilon > 0$ and $\epsilon' = \epsilon/M_j$ then there is N_c such that when $n \geq N_c$ we have that $|c_n - 0| = |c_n| < \epsilon'$.

So we see that

$$|(a_n)_j c_n - 0| = |(a_n)_j c_n| = |(a_n)_j| |c_n| < M_j \epsilon' = \epsilon$$

Which implies that $\lim_{i \rightarrow \infty} (a_i)_j c_i = 0$. But since this holds for every component j of \mathbf{a}_i then the sequence $c_i \mathbf{a}_i$ converges to 0 i.e.

$$\lim_{i \rightarrow \infty} c_i \mathbf{a}_i = 0$$

□

Proof. 1.5.18

Let $i \rightarrow \mathbf{a}_i$ be a sequence convergent to \mathbf{a} and let $j \rightarrow \mathbf{a}_{i(j)}$ be any subsequence of \mathbf{a}_i .

Let us take the k th component $(a_i)_k$ of \mathbf{a}_i that converges to a_k in \mathbf{a} . Also, the k th component of the subsequence $\mathbf{a}_{i(j)}$ is $(a_{i(j)})_k$.

Given some $\epsilon > 0$ then there is N such that when $n \geq N$ we have that $|(a_n)_k - a_k| < \epsilon$.

Also, since for the subsequence we know that $i(j) \rightarrow \infty$ then we can find J such that $i(j) \geq N$ for all $j \geq J$.

Therefore since the subsequence $(a_{i(j)})_k$ are selected elements of $(a_i)_k$ then must be that $|(a_{i(j)})_k - a_k| < \epsilon$ when $i(j) \geq N$, which implies that

$$\lim_{j \rightarrow \infty} (a_{i(j)})_k = a_k$$

Finally, since this holds for every component k of the sequence and the subsequence then the subsequence $j \rightarrow \mathbf{a}_{i(j)}$ converges to \mathbf{a} . \square

1.6 - Five Big Theorems

Proof. 1.6.1

Let a set $X \subset \mathbb{R}^n$ be contained in a ball centered at $\mathbf{x} \in \mathbb{R}^n$ hence

$$X \subset B_{R'}(\mathbf{x}) \text{ for some } R' < \infty$$

We can define $B_R(\mathbf{0})$ as a ball centered at the origin such that $R = R' + |\mathbf{x}|$.

So we see that

$$X \subset B'_R(\mathbf{x}) \subset B_R(\mathbf{0})$$

Therefore X is bounded. \square

Proof. 1.6.2

Let $A \subset \mathbb{R}^n$ be a subset that it's not compact. So A is not closed and/or not bounded.

Let us suppose A is closed but not bounded and let us define

$$f(\mathbf{x}) = |\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

Let us fix $\mathbf{x}_0 \in A$ then given $\epsilon > 0$ if we take $\delta = \epsilon$ we see that whenever $|\mathbf{x} - \mathbf{x}_0| < \delta$ we have that

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = ||\mathbf{x}| - |\mathbf{x}_0|| \leq |\mathbf{x} - \mathbf{x}_0| < \delta = \epsilon$$

Therefore f is continuous in A .

Now, let $M > 0$ in \mathbb{R} then since A is not bounded there is $\mathbf{x} \in A$ such that $|\mathbf{x}| > M$ so we see that

$$|f(\mathbf{x})| = \left| \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \right| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = |\mathbf{x}| > M$$

Therefore f is not bounded.

Let us suppose now that A is bounded but not closed, and let us select some $\mathbf{p} \in \overline{A}/A$ then if we define f as

$$f(\mathbf{x}) = \frac{1}{|\mathbf{x} - \mathbf{p}|}$$

We see that f is continuous in A because $|\mathbf{x}|$ is continuous in A . Also, as \mathbf{x} approaches \mathbf{p} the function blows up so it's not bounded. \square

Proof. 1.6.3

Let $z = x + iy$ where $x, y \in \mathbb{R}$ and let us define the polynomial $p(z) = 1 + x^2y^2$.

We want to prove that $p(z)$ has no roots, so if $p(z) = 0$ we see that must be that $xy = i$ which cannot happen because $x, y \in \mathbb{R}$ and no multiplication of two real numbers gives us the imaginary number i .

Therefore $p(z)$ has no roots, this doesn't contradict the Fundamental theorem of algebra because the theorem requires that the polynomials are of the form

$$p(z) = z^k + a_{k-1}z^{k-1} + \dots + a_0$$

Where $z \in \mathbb{C}$ and $1 + x^2y^2$ cannot be put in this form. □

1.7 - Derivatives in several variables as linear transformations

Proof. 1.7.6

a. Let

$$\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos x \\ x^2y + y^2 \\ \sin(x^2 - y) \end{pmatrix}$$

Then the partial derivative of \mathbf{f} are

$$\overrightarrow{D}_1 \mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\sin x \\ 2xy \\ 2x \cos(x^2 - y) \end{pmatrix} \quad \overrightarrow{D}_2 \mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ x^2 + 2y \\ -\cos(x^2 - y) \end{pmatrix}$$

b. Let

$$\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{x^2 + y^2} \\ xy \\ \sin^2 xy \end{pmatrix}$$

Then the partial derivative of \mathbf{f} are

$$\overrightarrow{D}_1 \mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ y \\ 2y \cos xy \sin xy \end{pmatrix} \quad \overrightarrow{D}_2 \mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{y}{\sqrt{x^2+y^2}} \\ x \\ 2x \cos xy \sin xy \end{pmatrix}$$

□

Proof. 1.7.7

We can write the results of problem 1.7.6 as the Jacobian matrix of both \mathbf{f} as follows

a.

$$\left[\mathbf{J}\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} \right] = \left[\overrightarrow{D_1}\mathbf{f}(x) \quad \overrightarrow{D_2}\mathbf{f}(x) \right] = \begin{bmatrix} -\sin x & 0 \\ 2xy & x^2 + 2y \\ 2x \cos(x^2 - y) & -\cos(x^2 - y) \end{bmatrix}$$

b.

$$\left[\mathbf{J}\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} \right] = \left[\overrightarrow{D_1}\mathbf{f}(x) \quad \overrightarrow{D_2}\mathbf{f}(x) \right] = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ y & x \\ 2y \cos xy \sin xy & 2x \cos xy \sin xy \end{bmatrix}$$

□

Proof. 1.7.9

- a. Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ then the derivative of \mathbf{f} is an $m \times n$ matrix called $[\mathbf{D}\mathbf{f}(\mathbf{x})]$.
- b. Let $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ then the derivative of f is a 1×3 matrix or a row vector with 3 components.
- c. Let $\mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}^4$ then the derivative of \mathbf{f} is a 4×1 matrix or a column vector with 4 components.

□

Proof. 1.7.10

Let $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a function.

a. Let $\mathbf{a}, \mathbf{v} \in \mathbb{R}^2$ and suppose \mathbf{f} is affine, then it can be written as

$$\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$$

where \mathbf{A} is a 2×2 matrix and \mathbf{b} is a vector in \mathbb{R}^2 .

So we have that

$$\mathbf{f}(\mathbf{a} + \mathbf{v}) = \mathbf{A}(\mathbf{a} + \mathbf{v}) + \mathbf{b} = \mathbf{Aa} + \mathbf{Av} + \mathbf{b} = \mathbf{f}(\mathbf{a}) + \mathbf{Av}$$

Let us show now that the derivative of \mathbf{f} at \mathbf{a} is \mathbf{A} by using the definition as follows

$$\begin{aligned} \lim_{|\mathbf{h}| \rightarrow 0} \frac{1}{|\mathbf{h}|} & \left(\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \mathbf{Ah} \right) = \\ &= \lim_{|\mathbf{h}| \rightarrow 0} \frac{1}{|\mathbf{h}|} \left(\mathbf{A}(\mathbf{a} + \mathbf{h}) + \mathbf{b} - (\mathbf{Aa} + \mathbf{b}) - \mathbf{Ah} \right) \\ &= \lim_{|\mathbf{h}| \rightarrow 0} \frac{1}{|\mathbf{h}|} \left(\mathbf{Aa} + \mathbf{Ah} + \mathbf{b} - \mathbf{Aa} - \mathbf{b} - \mathbf{Ah} \right) \\ &= \mathbf{0} \end{aligned}$$

Therefore

$$\mathbf{f}(\mathbf{a} + \mathbf{v}) = \mathbf{f}(\mathbf{a}) + [\mathbf{D}\mathbf{f}(\mathbf{a})]\mathbf{v}$$

b. Let

$$\mathbf{f} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ 0 \end{pmatrix}$$

We see that \mathbf{f} is not affine because it can't be written as $\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$. Since \mathbf{f} is differentiable at \mathbf{a} then we know that $[\mathbf{D}\mathbf{f}(\mathbf{a})] = [\mathbf{J}\mathbf{f}(\mathbf{a})]$ and

$$[\mathbf{J}\mathbf{f}(\mathbf{a})] = \begin{bmatrix} 2a_1 & 0 \\ 0 & 0 \end{bmatrix}$$

If part (a) holds for \mathbf{f} , we must have that

$$\begin{aligned} \mathbf{f} \begin{pmatrix} a_1 + v_1 \\ a_2 + v_2 \end{pmatrix} &= \mathbf{f} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{bmatrix} 2a_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1^2 \\ 0 \end{pmatrix} + \begin{pmatrix} 2a_1 v_1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} a_1^2 + 2a_1 v_1 \\ 0 \end{pmatrix} \end{aligned}$$

But we see that

$$\mathbf{f} \begin{pmatrix} a_1 + v_1 \\ a_2 + v_2 \end{pmatrix} = \begin{pmatrix} (a_1 + v_1)^2 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1^2 + 2a_1 v_1 + v_1^2 \\ 0 \end{pmatrix}$$

Therefore, if \mathbf{f} is not affine, then part (a) doesn't hold. \square

Proof. 1.7.11

a. The Jacobian of $f \begin{pmatrix} x \\ y \end{pmatrix} = \sin(xy)$ is

$$\left[\mathbf{J}f \begin{pmatrix} x \\ y \end{pmatrix} \right] = [y \cos(xy) \quad x \cos(xy)]$$

b. The Jacobian of $f \begin{pmatrix} x \\ y \end{pmatrix} = e^{x^2+y^3}$ is

$$\left[\mathbf{J}f \begin{pmatrix} x \\ y \end{pmatrix} \right] = [2xe^{x^2+y^3} \quad 3y^2e^{x^2+y^3}]$$

c. The Jacobian of $\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ x+y \end{pmatrix}$ is

$$\left[\mathbf{J}\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} \right] = \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix}$$

d. The Jacobian of $\mathbf{f} \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$ is

$$\left[\mathbf{J}\mathbf{f} \begin{pmatrix} r \\ \theta \end{pmatrix} \right] = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}$$

□

Proof. 1.7.15

- a. A mapping $F : \text{Mat}(n, m) \rightarrow \text{Mat}(k, l)$ is differentiable at $A \in \text{Mat}(n, m)$ if exists a linear transformation $[\mathbf{D}F(A)] : \text{Mat}(n, m) \rightarrow \text{Mat}(k, l)$ such that

$$\lim_{H \rightarrow [0]} \frac{1}{|H|} \left((F(A + H) - F(A)) - [\mathbf{D}F(A)]H \right) = [0]$$

- b. Let $F : \text{Mat}(n, m) \rightarrow \text{Mat}(n, n)$ given by $F(A) = AA^T$ then we see that

$$\begin{aligned} & \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((A + H)(A + H)^T - AA^T - [\mathbf{D}F(A)]H \right) = \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((A + H)(A^T + H^T) - AA^T - [\mathbf{D}F(A)]H \right) \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((AA^T + HA^T + AH^T + HH^T) - AA^T - [\mathbf{D}F(A)]H \right) \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left(HA^T + AH^T + HH^T - [\mathbf{D}F(A)]H \right) \end{aligned}$$

So let us define $[\mathbf{D}F(A)]$ as the linear transformation such that

$$[\mathbf{D}F(A)] : H \rightarrow HA^T + AH^T$$

Then we see that

$$|HA^T + AH^T + HH^T - HA^T - AH^T| = |HH^T|$$

This gives

$$\lim_{H \rightarrow [0]} \frac{|HH^T|}{|H|} \leq \lim_{H \rightarrow [0]} \frac{|H||H^T|}{|H|} = \lim_{H \rightarrow [0]} \frac{|H||H|}{|H|} = 0$$

Where we used that $|H| = |H^T|$.

Therefore F is differentiable and the linear transformation

$$[\mathbf{D}F(A)] : H \rightarrow HA^T + AH^T$$

is the derivative of F .

□

Proof. 1.7.18

Let $U \subset \mathbb{R}^n$ be open and let $f : U \rightarrow \mathbb{R}$ be differentiable at $\mathbf{a} \in U$.

We want to show that if \mathbf{v} is a unit vector making an angle θ with the gradient $\vec{\nabla}f(\mathbf{a})$ then

$$[\mathbf{D}f(\mathbf{a})]\mathbf{v} = |\vec{\nabla}f(\mathbf{a})| \cos \theta$$

We know that to evaluate the gradient $\vec{\nabla}f(\mathbf{a})$ at \mathbf{v} we need to compute the dot product between the two vectors i.e. $\vec{\nabla}f(\mathbf{a}) \cdot \mathbf{v}$ and we know that

$$\vec{\nabla}f(\mathbf{a}) \cdot \mathbf{v} = |\vec{\nabla}f(\mathbf{a})||\mathbf{v}| \cos \theta = |\vec{\nabla}f(\mathbf{a})| \cos \theta$$

Where we used that $|\mathbf{v}| = 1$.

On the other hand, we know that $[\mathbf{D}f(\mathbf{a})]$ is defined as the row vector $[D_1f(\mathbf{a}) \dots D_n f(\mathbf{a})]$ so taking the directional derivative in the direction \mathbf{v} implies computing the matrix multiplication between $[\mathbf{D}f(\mathbf{a})]$ and \mathbf{v}

$$[D_1f(\mathbf{a}) \dots D_n f(\mathbf{a})] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = D_1f(\mathbf{a})v_1 + \dots + D_n f(\mathbf{a})v_n$$

But this is no different from taking the dot product between $[\mathbf{D}f(\mathbf{a})]^T = \vec{\nabla}f(\mathbf{a})$ and \mathbf{v} , so we get that

$$[\mathbf{D}f(\mathbf{a})]\mathbf{v} = |\vec{\nabla}f(\mathbf{a})| \cos \theta$$

Evaluating $[\mathbf{D}f(\mathbf{a})]$ in the direction \mathbf{v} gives us the how fast the function changes in the direction of \mathbf{v} and this is going to be maximum when \mathbf{v} points in the direction of $\vec{\nabla}f(\mathbf{a})$ i.e. $\theta = 0$ and hence $\cos \theta = 1$.

So $\vec{\nabla}f(\mathbf{a})$ must point in the direction that f changes the fastest.

Finally, if \mathbf{v} points in the direction of $\vec{\nabla}f(\mathbf{a})$ we know that the directional derivative is

$$[\mathbf{D}f(\mathbf{a})]\mathbf{v} = |\vec{\nabla}f(\mathbf{a})|$$

Then we know that the rate of change of f in the direction of $\vec{\nabla}f(\mathbf{a})$ (or \mathbf{v}) is $|\vec{\nabla}f(\mathbf{a})|$. \square

Proof. 1.7.21

Let us consider the determinant of 2×2 matrices as a function
 $\det : \text{Mat}(2, 2) \rightarrow \mathbb{R}$. We want to show that

$$[\mathbf{D} \det(I)]H = h_{1,1} + h_{2,2}$$

Where I is the identity and H is the increment matrix.

If this is the derivative of \det then the following limit must tend to 0

$$\begin{aligned} & \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((\det(I+H) - \det(I)) - [\mathbf{D} \det(I)]H \right) = \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left(1 + h_{2,2} + h_{1,1} + h_{1,1}h_{2,2} - h_{1,2}h_{2,1} - 1 - h_{1,1} - h_{2,2} \right) \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} (h_{1,1}h_{2,2} - h_{1,2}h_{2,1}) \\ &= \lim_{H \rightarrow [0]} \frac{\det(H)}{|H|} \end{aligned}$$

Where we used that

$$\begin{aligned} \det \begin{pmatrix} 1 + h_{1,1} & h_{1,2} \\ h_{2,1} & 1 + h_{2,2} \end{pmatrix} &= (1 + h_{1,1})(1 + h_{2,2}) - h_{1,2}h_{2,1} \\ &= 1 + h_{2,2} + h_{1,1} + h_{1,1}h_{2,2} - h_{1,2}h_{2,1} \end{aligned}$$

We see that $|h_{1,1}| \leq \sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}$ then

$$0 \leq \frac{|h_{1,1}h_{2,2}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} \leq |h_{2,2}|$$

In the same way

$$0 \leq \frac{|h_{1,2}h_{2,1}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} \leq |h_{2,1}|$$

So, by the squeeze theorem, as $H \rightarrow [0]$ we get that

$$\begin{aligned} & \lim_{H \rightarrow [0]} \frac{|h_{1,1}h_{2,2}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} = 0 \\ & \lim_{H \rightarrow [0]} \frac{|h_{1,2}h_{2,1}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} = 0 \end{aligned}$$

Also, by the triangle inequality we see that

$$0 \leq \frac{|h_{1,1}h_{2,2} - h_{1,2}h_{2,1}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} \leq \frac{|h_{1,1}h_{2,2}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} + \frac{|h_{1,2}h_{2,1}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}}$$

Therefore, again by the squeeze theorem we have that

$$\lim_{H \rightarrow [0]} \frac{|h_{1,1}h_{2,2} - h_{1,2}h_{2,1}|}{\sqrt{h_{1,1}^2 + h_{1,2}^2 + h_{2,1}^2 + h_{2,2}^2}} = 0$$

This implies that the limit of $\det(H)/|H|$ as $H \rightarrow [0]$ is 0 and hence the derivative of $\det : \text{Mat}(2, 2) \rightarrow \mathbb{R}$ is $h_{1,1} + h_{2,2}$. \square

1.8 - Rules for computing derivatives

Proof. 1.8.5

The proof is using the Jacobian matrix to determine the derivative of a function $h(\mathbf{a}) = (fg)(\mathbf{a})$ but we are only allowed to do that when we are sure that h is a differentiable function at \mathbf{a} .

Therefore, if we wanted to complete this proof we must show that h is differentiable at \mathbf{a} . \square

Proof. 1.8.12

- a. Let $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be differentiable and $[\mathbf{D}\mathbf{f}(\mathbf{0})]$ be not invertible. Also, suppose $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a function such that

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{x}$$

We want to prove that \mathbf{g} is not differentiable.

To compute the derivative $[\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{a})]$ we note that

$$\begin{aligned} & \frac{1}{|\mathbf{h}|} \left| (\mathbf{g} \circ \mathbf{f})(\mathbf{a} + \mathbf{h}) - (\mathbf{g} \circ \mathbf{f})(\mathbf{a}) - [\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{a})]\mathbf{h} \right| \\ &= \frac{1}{|\mathbf{h}|} \left| \mathbf{a} + \mathbf{h} - \mathbf{a} - [\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{a})]\mathbf{h} \right| \\ &= \frac{1}{|\mathbf{h}|} \left| \mathbf{h}(\mathbf{1} - [\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{a})]) \right| \\ &= \left| \mathbf{1} - [\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{a})] \right| \end{aligned}$$

So for this to tend to $\mathbf{0}$ when $\mathbf{h} \rightarrow \mathbf{0}$ must be that

$$[\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{a})] = \mathbf{1}$$

or using the chain rule assuming \mathbf{g} is differentiable at $\mathbf{f}(\mathbf{a})$ we get that

$$[\mathbf{D}\mathbf{g}(\mathbf{f}(\mathbf{a}))][\mathbf{D}\mathbf{f}(\mathbf{a})] = \mathbf{1}$$

So we have that

$$[\mathbf{D}\mathbf{g}(\mathbf{f}(\mathbf{a}))] = [\mathbf{D}\mathbf{f}(\mathbf{a})]^{-1}$$

But if $\mathbf{a} = \mathbf{0}$ we know that $[\mathbf{D}\mathbf{f}(\mathbf{0})]$ is not invertible then \mathbf{g} is not differentiable at $\mathbf{f}(\mathbf{0})$.

Therefore, \mathbf{g} is not differentiable, and hence the statement is true i.e. there is no differentiable function $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{x}$.

- b. Let us define $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$ as follows

$$\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} x^2 \sin(\frac{1}{x}) & \text{when } x \neq 0 \\ 0 & \text{when } x = 0 \end{cases}$$

It can be shown that \mathbf{f} is differentiable everywhere.

We see that $D_1\mathbf{f}(\mathbf{x})$ is

$$D_1\mathbf{f}(\mathbf{x}) = \begin{cases} 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}) & \text{when } x \neq 0 \\ 0 & \text{when } x = 0 \end{cases}$$

But $\lim_{\mathbf{x} \rightarrow \mathbf{0}} D_1\mathbf{f}(\mathbf{x})$ is undetermined since $\lim_{x \rightarrow 0} \cos(\frac{1}{x})$ is undetermined.

Therefore $D_1\mathbf{f}(\mathbf{x})$ is not continuous, and hence the statement is false.

□

Proof. 1.8.13

Let $F : U \rightarrow \text{Mat}(n, n)$ be a map given by $F(A) = (A + A^2)^{-1}$ where $U \subset \text{Mat}(n, n)$ is the set of matrices A such that $A + A^2$ is invertible.

We want to compute the derivative of F .

Let us define $G(A) = A + A^2$ then

$$\begin{aligned} [\mathbf{D}G(A)]H &= [\mathbf{D}(A + A^2)]H \\ &= [\mathbf{D}A]H + [\mathbf{D}A^2]H \\ &= H + AH + HA \end{aligned}$$

Also, if we define $H(A) = A^{-1}$ then we note that $F = H \circ G$, hence

$$\begin{aligned} [\mathbf{D}F(A)]H &= [\mathbf{D}(H \circ G)(A)]H \\ &= [\mathbf{D}H(G(A))] [\mathbf{D}G(A)]H \\ &= [\mathbf{D}H(G(A))](H + AH + HA) \\ &= [\mathbf{D}H(A + A^2)](H + AH + HA) \\ &= -(A + A^2)^{-1}(H + AH + HA)(A + A^2)^{-1} \end{aligned}$$

Where we used the chain rule and that $[\mathbf{D}A^{-1}]H = -A^{-1}HA^{-1}$. \square

1.9 - The Mean Value Theorem and Criteria for Differentiability

Proof. 1.9.2

a. Let

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \frac{3x^2y - y^3}{x^2 + y^2} & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$$

Given that f is a rational function whose denominator does not vanish, we can compute both partial derivatives at any point $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Let us compute the partial derivatives of f at $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

We see that

$$\begin{aligned} D_1 f \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \lim_{h \rightarrow 0} \frac{1}{h} \left(f \begin{pmatrix} h \\ 0 \end{pmatrix} - f \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \lim_{h \rightarrow 0} \frac{1}{h} (0 + 0) = 0 \\ D_2 f \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \lim_{h \rightarrow 0} \frac{1}{h} \left(f \begin{pmatrix} 0 \\ h \end{pmatrix} - f \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \lim_{h \rightarrow 0} \frac{1}{h} (-h + 0) = -1 \end{aligned}$$

So all partial derivatives exist everywhere.

Let us suppose f is differentiable everywhere, we want to arrive at a contradiction. Then we can compute the directional derivatives using the jacobian.

Let us consider the derivative of f at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in the direction of the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we see that

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} \left(f \begin{pmatrix} 0 \\ 0 \end{pmatrix} + h \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) - f \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \frac{3h^3 - h^3}{h^2 + h^2} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \frac{2h^3}{2h^2} \\ &= 1 \end{aligned}$$

But if we compute this directional derivative by multiplying the Jacobian matrix by the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ we get that

$$\left[D_1 f \begin{pmatrix} 0 \\ 0 \end{pmatrix}, D_2 f \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = [0 \quad -1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1$$

This is not the same value we got first, therefore we have a contradiction and hence f cannot be differentiable at the origin.

On the other hand, since we cannot compute the directional derivative at $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ using the Jacobian because f is not differentiable there, then, we have to compute it using the definition of directional derivative as follows

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{1}{h} \left(f \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + h \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right) - f \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) \right) = \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{3h^2 v_1^2 h v_2 - h^3 v_2^3}{h^2 v_1^2 + h^2 v_2^2} \right) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{h(3v_1^2 v_2 - v_2^3)}{v_1^2 + v_2^2} \right) \\ &= \frac{3v_1^2 v_2 - v_2^3}{v_1^2 + v_2^2} \end{aligned}$$

Therefore the limit exist for any vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and hence the directional derivative exist at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

This implies that all directional derivatives exists, since at any other point different from the origin we can use the jacobian to determine it.

b. Let

$$g \begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0} \\ 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0} \end{cases}$$

Given that g is a rational function whose denominator does not vanish, we can compute both partial derivatives at any point $\begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}$.

So we can compute the directional derivatives at any point different from $\mathbf{0}$ using the Jacobian, let us check that the directional derivative also exists at $\mathbf{0}$.

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} \left(g \begin{pmatrix} 0 \\ 0 \end{pmatrix} + h \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right) - g \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{h^2 v_1^2 h v_2}{h^4 v_1^4 + h^2 v_2^2} \right) \\ &= \lim_{h \rightarrow 0} \left(\frac{v_1^2 v_2}{h^2 v_1^4 + v_2^2} \right) \\ &= \frac{v_1^2}{v_2} \end{aligned}$$

Hence g has directional derivatives at every point.

On the other hand, we know that $g(\mathbf{0}) = 0$ but if we approach the origin along the curve $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ t^2 \end{pmatrix}$ we see that

$$\lim_{t \rightarrow 0} g \begin{pmatrix} t \\ t^2 \end{pmatrix} = \lim_{t \rightarrow 0} \frac{t^4}{t^4 + t^4} = \frac{1}{2}$$

Therefore g is not continuous at the origin.

c. Let

$$h \begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \frac{x^2y}{x^6+y^2} & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0} \\ 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0} \end{cases}$$

Given that h is a rational function whose denominator does not vanish, we can compute both partial derivatives at any point $\begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}$.

So we can compute the directional derivatives at any point different from $\mathbf{0}$ using the Jacobian, let us check that the directional derivative also exists at $\mathbf{0}$.

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} \left(h \begin{pmatrix} 0 \\ 0 \end{pmatrix} + t \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right) - h \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\frac{t^2 v_1^2 t v_2}{t^6 v_1^6 + t^2 v_2^2} \right) \\ &= \lim_{t \rightarrow 0} \left(\frac{v_1^2 v_2}{t^4 v_1^6 + v_2^2} \right) \\ &= \frac{v_1^2}{v_2} \end{aligned}$$

Hence h has directional derivatives at every point.

On the other hand, suppose we approach the origin along the curve $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ t^3 \end{pmatrix}$ then we see that

$$\lim_{t \rightarrow 0} \left| h \begin{pmatrix} t \\ t^3 \end{pmatrix} \right| = \lim_{t \rightarrow 0} \left| \frac{t^2 t^3}{t^6 + t^6} \right| = \lim_{t \rightarrow 0} \left| \frac{t^5}{2t^6} \right| = \lim_{t \rightarrow 0} \frac{1}{2|t|} = \infty$$

Therefore h cannot be bounded in a neighborhood of the origin. □

1.10 - Review exercises for Chapter 1

Proof. 1.10.28

- a. Let us consider the mapping $F : \text{Mat}(n, n) \rightarrow \text{Mat}(n, n)$ given by $F(A) = A^3$ we want to prove it's differentiable, then we must show that there is a linear mapping $[DF(A)]H$ such that

$$\lim_{H \rightarrow [0]} \frac{1}{|H|} \left(F(A + H) - F(A) - [DF(A)]H \right) = [0]$$

Then

$$\begin{aligned} & \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((A + H)^3 - A^3 - [DF(A)]H \right) = \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((A + H)(A + H)(A + H) - A^3 - [DF(A)]H \right) \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left(A^3 + A^2H + AHA + AH^2 + HA^2 + HAH \right. \\ &\quad \left. + H^2A + H^3 - A^3 - [DF(A)]H \right) \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left(A^2H + AHA + AH^2 + HA^2 + HAH + H^2A + H^3 - [DF(A)]H \right) \end{aligned}$$

So if we let $[DF(A)]$ be a mapping such that

$$[DF(A)] : H \rightarrow A^2H + AHA + HA^2$$

We get that

$$\begin{aligned} & \lim_{H \rightarrow [0]} \frac{1}{|H|} \left((A + H)^3 - A^3 - [DF(A)]H \right) = \\ &= \lim_{H \rightarrow [0]} \frac{1}{|H|} \left(AH^2 + HAH + H^2A + H^3 \right) \end{aligned}$$

But also, we have that

$$\begin{aligned} \frac{1}{|H|} (|AH^2 + HAH + H^2A + H^3|) &\leq \frac{1}{|H|} (|AH^2| + |HAH| + |H^2A| + |H^3|) \\ &\leq \frac{1}{|H|} (|A||H|^2 + |H||A||H| + |H|^2|A| + |H|^3) \\ &= |A||H| + |A||H| + |H||A| + |H^2| \\ &= 3|A||H| + |H^2| \end{aligned}$$

And

$$\lim_{H \rightarrow [0]} 3|A||H| + |H^2| = [0]$$

Therefore $F(A) = A^3$ is differentiable and its derivative is $A^2H + AHA + HA^2$.

- b. Let $F : \text{Mat}(n, n) \rightarrow \text{Mat}(n, n)$ given by $F(A) = A^k$ for $k \geq 1$. Then the derivative $[DF(A)]$ must be a mapping such that

$$\lim_{H \rightarrow [0]} \frac{1}{|H|} \left((A + H)^k - A^k - [DF(A)]H \right) = [0]$$

Let $[DF(A)]$ be the mapping such that

$$H \rightarrow A^{k-1}H + A^{k-2}HA + A^{k-3}HA^2 + \dots + HA^{k-1}$$

Since $(A + H)^k$ contains all combinations of powers between A and H such that the sum of the powers of each term is k then the expression

$$(A + H)^k - A^k - [DF(A)]H$$

will have all terms with powers of H bigger or equal than 2, then using the properties the matrix norm such as $|AH| \leq |A||H|$ and $|H^n| \leq |H|^n$ we get that

$$\frac{1}{|H|} \left| (A + H)^k - A^k - [DF(A)]H \right| \leq 0 \quad \text{as } H \rightarrow [0]$$

Therefore the derivative of the mapping $F(A) = A^k$ is

$$A^{k-1}H + A^{k-2}HA + A^{k-3}HA^2 + \dots + HA^{k-1}$$

□

Proof. 1.10.32

- a. Let the map $A \rightarrow A^{-3}$, so, if we let $f(A) = A^3$ and $g(A) = A^{-1}$ then we can compute the derivative using the chain rule as follows

$$\begin{aligned} [\mathbf{D}(f \circ g)(A)]H &= [\mathbf{D}f(g(A))][\mathbf{D}g(A)]H \\ &= [\mathbf{D}f(A^{-1})](-A^{-1}HA^{-1}) \\ &= A^{-2}(-A^{-1}HA^{-1}) + A^{-1}(-A^{-1}HA^{-1})A^{-1} + (-A^{-1}HA^{-1})A^{-2} \\ &= -(A^{-3}HA^{-1} + A^{-2}HA^{-2} + A^{-1}HA^{-3}) \end{aligned}$$

- b. Let the map $A \rightarrow A^{-n}$, so, if we let $f(A) = A^n$ and $g(A) = A^{-1}$ then we can compute the derivative using the chain rule as follows

$$\begin{aligned} [\mathbf{D}(f \circ g)(A)]H &= \\ &= [\mathbf{D}f(g(A))][\mathbf{D}g(A)]H \\ &= [\mathbf{D}f(A^{-1})](-A^{-1}HA^{-1}) \\ &= -[A^{-n+1}(A^{-1}HA^{-1}) + A^{-n+2}(A^{-1}HA^{-1})A^{-1} + A^{-n+3}(A^{-1}HA^{-1})A^{-2} + \\ &\quad + \dots + (A^{-1}HA^{-1})A^{-n+1}] \\ &= -[A^{-n}HA^{-1} + A^{-n+1}HA^{-2} + A^{-n+2}HA^{-3} + \dots + A^{-1}HA^{-n}] \end{aligned}$$

□