Supplementary Notes on Linear Algebra

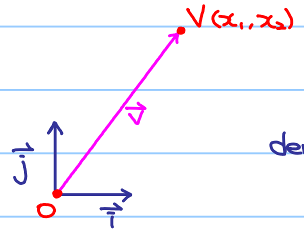
§ 1 Vectors in $\mathbb{R}^n$

Definition 1.1

A vector in $\mathbb{R}^n$ is an element of $\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n) : x_1, x_2, \cdots, x_n \in \mathbb{R}\}$.

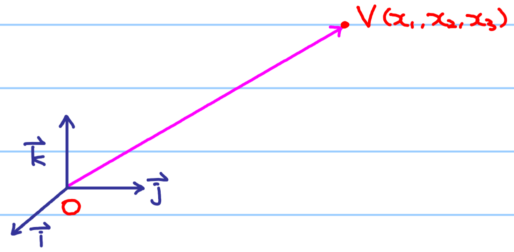
Example 1.1

A vector in $\mathbb{R}^2$ can be written as (x_1, x_2) or $x_1\vec{i} + x_2\vec{j}$.



denoted by $\vec{v}$ or $\overrightarrow{OV}$

A vector in $\mathbb{R}^3$ can be written as (x_1, x_2, x_3) or $x_1\vec{i} + x_2\vec{j} + x_3\vec{k}$.



A vector in $\mathbb{R}^n$ can be written as $(x_1, x_2, \cdots, x_n)$ or $x_1\vec{e_1} + x_2\vec{e_2} + \cdots + x_n\vec{e_n}$.

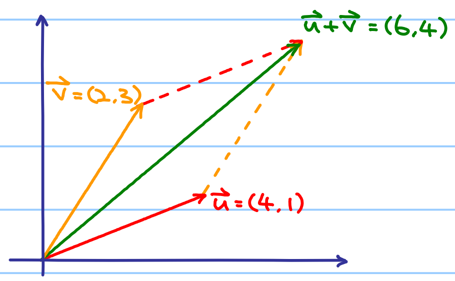
$\vec{0} = (0, 0, \cdots, 0) = 0\vec{e_1} + 0\vec{e_2} + \cdots + 0\vec{e_n}$ is said to the zero vector.

Definition 1.2 (Vector Addition)

If $\vec{u} = (u_1, u_2, \cdots, u_n)$, $\vec{v} = (v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$, $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \cdots, u_n + v_n)$.

Example 1.2

If $\vec{u} = (4, 1)$, $\vec{v} = (2, 3) \in \mathbb{R}^2$

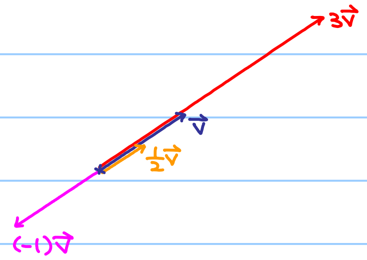


Definition 1.3 (Scalar Multiplication)

If $\vec{v} = (v_1, v_2, \cdots, v_n)$, $t \in \mathbb{R}$ (called scalar), $t\vec{v} = (tv_1, tv_2, \cdots, tv_n)$.

Example 1.3

If $\vec{v} = (4,2) \in \mathbb{R}^2$, $3\vec{v} = (12,6)$, $\frac{1}{2}\vec{v} = (2,1)$, $(-1)\vec{v} = (-4,-2)$.



Definition 1.4

$\vec{v}, \vec{w} \in \mathbb{R}^n$ are said to be parallel if $\vec{v} = t\vec{w}$ for some $t \in \mathbb{R}$.

Definition 1.5

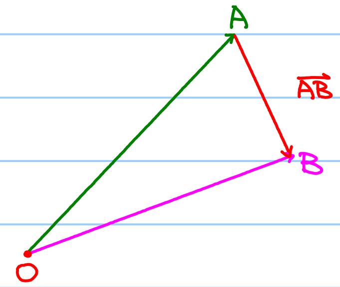
Let $\vec{v}, \vec{w} \in \mathbb{R}^n$.

$-\vec{v}$ is defined as $(-1)\vec{v}$ and $\vec{u} - \vec{v}$ is defined as $\vec{u} + (-\vec{v})$.

Example 1.4

If $\vec{OA} = \vec{a} = 3\vec{i} + 5\vec{j}$ and $\vec{OB} = \vec{b} = 4\vec{i} + 2\vec{j}$,

then $\vec{AB} = \vec{OB} - \vec{OA} = \vec{b} - \vec{a} = (4\vec{i} + 2\vec{j}) - (3\vec{i} + 5\vec{j}) = \vec{i} - 3\vec{j}$.



Proposition 1.1

Let $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, $s, t \in \mathbb{R}$.

1) (Commutative Law of Vector Addition) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

2) (Associative Law of Vector Addition) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$

3) (Existence of Additive Identity) $\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}$

4) (Existence of Additive Inverse) $\vec{v} + (-\vec{v}) = (-\vec{v}) + \vec{v} = \vec{0}$

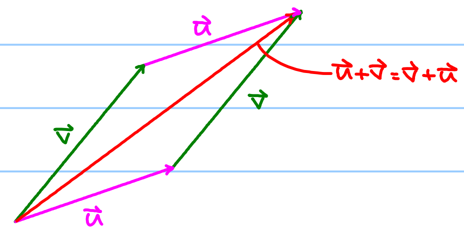
5) (Existence of Multiplicative Identity) $1\vec{v} = \vec{v}$ where $1 \in \mathbb{R}$

6) (Associative Law of Scalar Multiplication) $s(t\vec{v}) = (st)\vec{v}$

7) (Distributive Law of Scalar Multiplication) $s(\vec{u} + \vec{v}) = s\vec{u} + s\vec{v}$ and $(s+t)\vec{v} = s\vec{v} + t\vec{v}$.

Remark: $\mathbb{R}^n$ is a vector space proof of (1):

Let $\vec{u}=(u_1,u_2,\cdots,u_n)$, $\vec{v}=(v_1,v_2,\cdots,v_n)\in\mathbb{R}^n$.



$\vec{u}+\vec{v}=(u_1+v_1,u_2+v_2,\cdots,u_n+v_n)$

$\quad=(v_1+u_1,v_2+u_2,\cdots,v_n+u_n)\quad(\because u_i,v_i\in\mathbb{R},\ u_i+v_i=v_i+u_i)$

$\quad=\vec{v}+\vec{u}$

Definition 1.6

If $\vec{v}=(v_1,v_2,\cdots,v_n)$, length of $\vec{v}$, $|\vec{v}|=\left(\sum_{i=1}^{n}v_i^2\right)^{\frac{1}{2}}=\sqrt{v_1^2+v_2^2+\cdots+v_n^2}$.

Exercise 1.1

Let $\vec{v}\in\mathbb{R}^n$, $k\in\mathbb{R}$. Show that $|k\vec{v}|=|k||\vec{v}|$.

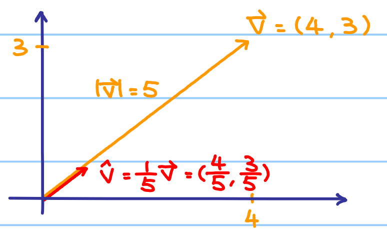
If we let $\hat{v}=\frac{1}{|\vec{v}|}\vec{v}$, then $\hat{v}\parallel\vec{v}$ and $|\hat{v}|=1$. $\hat{v}$ is said to be the unit vector of $\vec{v}$.

Idea : A vector $\vec{v}$ in $\mathbb{R}^n$ is a quantity with direction and magnitude.

$\vec{v}=|\vec{v}|\hat{v}$ where $\hat{v}$ and $|\vec{v}|$ give the direction and magnitude of $\vec{v}$ respectively.

Example 1.5

If $\vec{v}=(4,3)\in\mathbb{R}^2$, $|\vec{v}|=\sqrt{4^2+3^2}=5$ (Pyth thm.) and $\hat{v}=\frac{4}{5}\vec{i}+\frac{3}{5}\vec{j}$.



Definition 1.7 (Dot Product)

If $\vec{u}=(u_1,u_2,\cdots,u_n)$, $\vec{v}=(v_1,v_2,\cdots,v_n)\in\mathbb{R}^n$, $\vec{u}\cdot\vec{v}=\sum_{i=1}^{n}u_iv_i=u_1v_1+u_2v_2+\cdots+u_nv_n$.

In particular, $\vec{v}\cdot\vec{v}=\sum_{i=1}^{n}v_i^2=|\vec{v}|^2$.

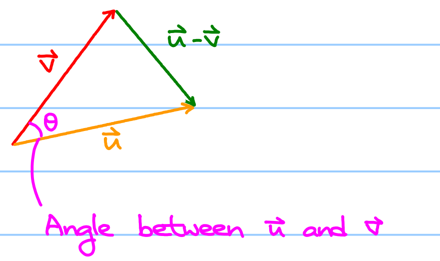
Example 1.5

If $\vec{u}=(4,2,3)$, $\vec{v}=(-1,6,-2)\in\mathbb{R}^3$, $\vec{u}\cdot\vec{v}=4\cdot(-1)+2\cdot6+3\cdot(-2)=2$.

Geometrical meaning ?

Cosine Law: $|\vec{u}-\vec{v}|^2 = |\vec{u}|^2+|\vec{v}|^2-2|\vec{u}||\vec{v}|\cos\theta$

Triangle spanned by $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$.



$$\sum_{i=1}^{n}(u_i-v_i)^2 = \sum_{i=1}^{n}u_i^2+\sum_{i=1}^{n}v_i^2-2|\vec{u}||\vec{v}|\cos\theta$$

$$\sum_{i=1}^{n}u_i^2-2\sum_{i=1}^{n}u_iv_i+\sum_{i=1}^{n}v_i^2 = \sum_{i=1}^{n}u_i^2+\sum_{i=1}^{n}v_i^2-2|\vec{u}||\vec{v}|\cos\theta$$

$$\sum_{i=1}^{n}u_iv_i = |\vec{u}||\vec{v}|\cos\theta$$

$$\vec{u}\cdot\vec{v} = |\vec{u}||\vec{v}|\cos\theta$$

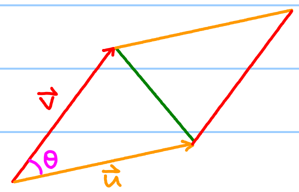
Direct consequence:

1) $\vec{v}$ is perpendicular (or orthogonal) to $\vec{v}$, i.e. $\vec{v}\perp\vec{w} \Leftrightarrow \theta=\frac{\pi}{2} \Leftrightarrow \vec{v}\cdot\vec{w}=0$

2) $\vec{e_i}\cdot\vec{e_j} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}$

Furthermore, let $\vec{u}=u_1\vec{i}+u_2\vec{j}$, $\vec{v}=v_1\vec{i}+v_2\vec{j}\in\mathbb{R}^2$.

The area of parallelogram spanned by $\vec{u}$ and $\vec{v}$



$= |\vec{u}||\vec{v}||\sin\theta|$

$= \sqrt{|\vec{u}|^2|\vec{v}|^2(1-\cos^2\theta)}$

$= \sqrt{|\vec{u}|^2|\vec{v}|^2-|\vec{u}\cdot\vec{v}|^2}$

$= \sqrt{(u_1^2+u_2^2)(v_1^2+v_2^2)-(u_1v_1+u_2v_2)^2}$

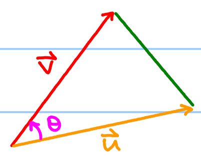
$= \sqrt{(u_1v_2-u_2v_1)^2}$

$= |u_1v_2-u_2v_1|$

Remark: Assume that θ is the angle measured from $\vec{u}$ to $\vec{v}$.

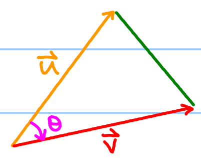
The signed area of parallelogram spanned by $\vec{u}$ and $\vec{v}$ $= |\vec{u}||\vec{v}|\sin\theta = u_1v_2-u_2v_1 = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$

$u_1v_2-u_2v_1>0$



$\theta>0$

$u_1v_2-u_2v_1<0$



$\theta<0$

Proposition 1.2

Let $\vec{u},\vec{v},\vec{w}\in\mathbb{R}^n$, $t\in\mathbb{R}$.

1) (Commutative Law of Dot Product) $\vec{u}\cdot\vec{v}=\vec{v}\cdot\vec{u}$

2) (Distributive Law of Dot Product) $\vec{u}\cdot(\vec{v}+\vec{w})=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w}$

3) $(t\vec{u})\cdot\vec{v}=\vec{u}\cdot(t\vec{v})=t(\vec{u}\cdot\vec{v})$

proof of (2):

Let $\vec{u} = (u_1, u_2, \cdots, u_n)$, $\vec{v} = (v_1, v_2, \cdots, v_n)$ and $\vec{w} = (w_1, w_2, \cdots, w_n)$

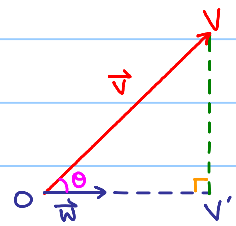
$\vec{u}\cdot(\vec{v}+\vec{w}) = \sum_{i=1}^{n} u_i(v_i+w_i) = \sum_{i=1}^{n}(u_iv_i+u_iw_i) = \sum_{i=1}^{n} u_iv_i + \sum_{i=1}^{n} u_iw_i = \vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w}$

Furthermore, $(\vec{v}+\vec{w})\cdot\vec{u} = \vec{u}\cdot(\vec{v}+\vec{w}) = \vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w} = \vec{v}\cdot\vec{u}+\vec{w}\cdot\vec{u}$.

Projection of $\vec{v}$ along $\vec{w}$:

length of $\vec{OV'} = |\vec{v}|\cos\theta$

$\text{proj}_{\vec{w}}(\vec{v}) = \vec{OV'} = \underbrace{(|\vec{v}|\cos\theta)}_{\text{magnitude}}\underbrace{\hat{w}}_{\text{direction}} = \frac{|\vec{v}||\vec{w}|\cos\theta}{|\vec{w}|^2}\vec{w} = \frac{\vec{v}\cdot\vec{w}}{|\vec{w}|^2}\vec{w}$



which is the projection of $\vec{v}$ along $\vec{w}$

$\vec{OV} = \vec{v}$ can be expressed as $\vec{OV'}+\vec{V'V}$

where $\vec{OV'} = \text{proj}_{\vec{w}}(\vec{v}) = \frac{\vec{v}\cdot\vec{w}}{|\vec{w}|^2}\vec{w}$ and $\vec{V'V} = \vec{OV}-\vec{OV'} = \vec{v}-\frac{\vec{v}\cdot\vec{w}}{|\vec{w}|^2}\vec{w}$

Furthermore, $\vec{OV'} \parallel \vec{w}$ and $\vec{V'V}\cdot\vec{w} = (\vec{v}-\frac{\vec{v}\cdot\vec{w}}{|\vec{w}|^2}\vec{w})\cdot\vec{w} = 0$, so $\vec{V'V}\perp\vec{w}$.

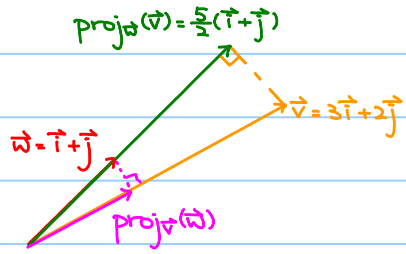
Example 1.6

Let $\vec{v} = 3\vec{i}+2\vec{j}$, $\vec{w} = \vec{i}+\vec{j} \in \mathbb{R}^2$.

$\text{proj}_{\vec{w}}(\vec{v}) = \frac{\vec{v}\cdot\vec{w}}{|\vec{w}|^2}\vec{w} = \frac{5}{2}(\vec{i}+\vec{j})$

Exercise: $\text{proj}_{\vec{v}}(\vec{w}) = ?$

Answer: $\text{proj}_{\vec{v}}(\vec{w}) = \frac{5}{13}(3\vec{i}+2\vec{j})$



Example 1.7

Let $\vec{v} = 2\vec{e}_1-3\vec{e}_2+\vec{e}_3+4\vec{e}_4$, $\vec{w} = \vec{e}_1+2\vec{e}_2-\vec{e}_3+\vec{e}_4 \in \mathbb{R}^4$

$|\vec{v}| = \sqrt{2^2+(-3)^2+1^2+4^2} = \sqrt{30}$, $|\vec{w}| = \sqrt{1^2+2^2+(-1)^2+1^2} = \sqrt{7}$

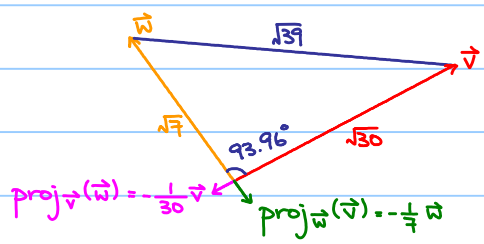
Distance between $\vec{v}$ and $\vec{w}$ $= |\vec{v}-\vec{w}| = |\vec{e}_1-5\vec{e}_2+2\vec{e}_3+3\vec{e}_4| = \sqrt{39}$

$\vec{v}\cdot\vec{w} = 2\cdot1+(-3)\cdot2+1\cdot(-1)+4\cdot1 = -1$

$|\vec{v}||\vec{w}|\cos\theta = \vec{v}\cdot\vec{w}$

$\sqrt{30}\,\sqrt{7}\cdot\cos\theta = -1$

$\theta = \cos^{-1}\left(\frac{-1}{\sqrt{210}}\right) \approx 93.96°$



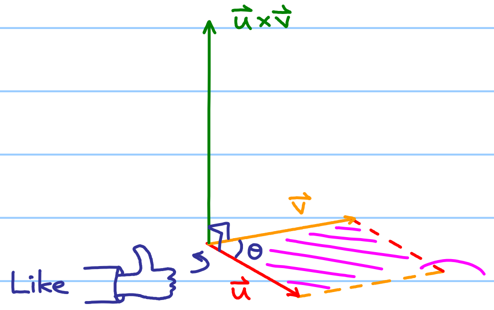
$\therefore$ Angle between $\vec{v}$ and $\vec{w}$ $\approx 93.96°$

$\text{proj}_{\vec{w}}(\vec{v}) = \frac{\vec{v}\cdot\vec{w}}{|\vec{w}|^2}\vec{w} = -\frac{1}{7}\vec{w}$

$\text{proj}_{\vec{v}}(\vec{w}) = \frac{\vec{w}\cdot\vec{v}}{|\vec{v}|^2}\vec{v} = -\frac{1}{30}\vec{v}$

Definition 1.8 (Cross Product in $\mathbb{R}^3$)

Let $\vec{u}, \vec{v} \in \mathbb{R}^3$, $\vec{u}\times\vec{v}$ is defined as the following:



parallelogram spanned by $\vec{u}$ and $\vec{v}$

area $= |\vec{u}||\vec{v}|\sin\theta$

(Right hand rule)

Caution: Cross product is only defined in $\mathbb{R}^3$ but NOT any other dimension.

Magnitude: $|\vec{u}\times\vec{v}| = |\vec{u}||\vec{v}|\sin\theta$

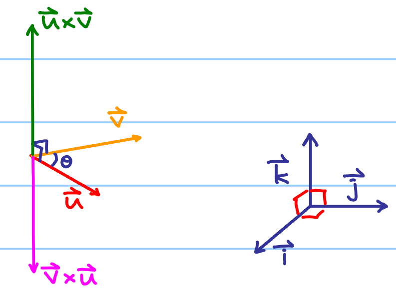
Direction: $\vec{u}\times\vec{v} \perp \vec{u}$ and $\vec{u}\times\vec{v} \perp \vec{v}$ with right hand rule.

By definition, we have:

1) $\vec{u}\times\vec{v} = -(\vec{v}\times\vec{u})$

2) $\vec{i}\times\vec{j} = \vec{k}$, $\vec{j}\times\vec{k} = \vec{i}$, $\vec{k}\times\vec{i} = \vec{j}$

$\vec{i}\times\vec{i} = \vec{j}\times\vec{j} = \vec{k}\times\vec{k} = \vec{0}$ (NOT just the number 0)



How to compute $\vec{u}\times\vec{v}$ if $\vec{u} = u_1\vec{i}+u_2\vec{j}+u_3\vec{k}$ and $\vec{v} = v_1\vec{i}+v_2\vec{j}+v_3\vec{k}$?

$\vec{u}\times\vec{v} = (u_1\vec{i}+u_2\vec{j}+u_3\vec{k})\times(v_1\vec{i}+v_2\vec{j}+v_3\vec{k})$ (Assume distributive law)

$= u_1v_1\,\vec{i}\times\vec{i} + u_1v_2\,\vec{i}\times\vec{j} + u_1v_3\,\vec{i}\times\vec{k} +$

$u_2v_1\,\vec{j}\times\vec{i} + u_2v_2\,\vec{j}\times\vec{j} + u_2v_3\,\vec{j}\times\vec{k} +$

$u_3v_1\,\vec{k}\times\vec{i} + u_3v_2\,\vec{k}\times\vec{j} + u_3v_3\,\vec{k}\times\vec{k}$

$= (u_2v_3-u_3v_2)\vec{i} - (u_1v_3-u_3v_1)\vec{j} + (u_1v_2-u_2v_1)\vec{k}$

$$= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Example 1.14

If $\vec{u} = \vec{i}+2\vec{k}$, $\vec{v} = 2\vec{i}-3\vec{j}+\vec{k}$,

then $\vec{u}\times\vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & 2 \\ 2 & -3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ -3 & 1 \end{vmatrix}\vec{i} - \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix}\vec{j} + \begin{vmatrix} 1 & 0 \\ 2 & -3 \end{vmatrix}\vec{j} = 6\vec{i}+3\vec{j}-3\vec{k}$

Proposition 13

Let $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$, $t \in \mathbb{R}$.

1) $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$

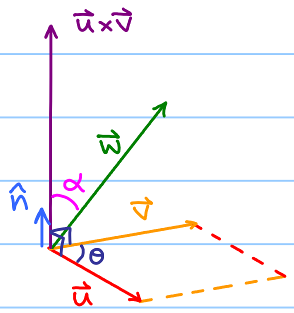
2) (Distributive Law of Cross Product) $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$

3) $(t\vec{u}) \times \vec{v} = \vec{u} \times (t\vec{v}) = t(\vec{u} \times \vec{v})$

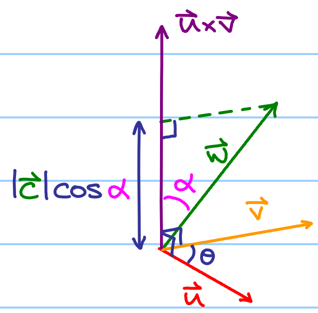
Note that if $\vec{u}, \vec{v} \in \mathbb{R}^3$, then $\vec{u} \times \vec{v} \in \mathbb{R}^3$.

Suppose $\vec{w} \in \mathbb{R}^3$, then we know that $(\vec{u} \times \vec{v}) \cdot \vec{w}$ is well-defined and it is just a scalar.

$(\vec{u} \times \vec{v}) \cdot \vec{w}$ is called scalar triple product, but does it have any geometrical meaning?



$\hat{n}$: unit vector of $\vec{u} \times \vec{v}$

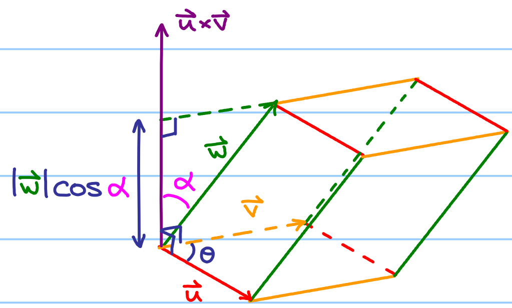


$\vec{u} \times \vec{v} = |\vec{u} \times \vec{v}| \, \hat{n}$

$(\vec{u} \times \vec{v}) \cdot \vec{w} = |\vec{u} \times \vec{v}| \, \hat{n} \cdot \vec{w}$

$= \underbrace{|\vec{u} \times \vec{v}|}_{\text{base area}} \underbrace{(|\vec{w}| \cos \alpha)}_{\text{height}}$

= (signed) volume of the parallelepiped spanned by $\vec{u}$, $\vec{v}$ and $\vec{w}$.



Remark: If $\frac{\pi}{2} < \alpha < \pi$, $\cos \alpha < 0$

If $\vec{u} = u_1\vec{i} + u_2\vec{j} + u_3\vec{k}$, $\vec{v} = v_1\vec{i} + v_2\vec{j} + v_3\vec{k}$ and $\vec{w} = w_1\vec{i} + w_2\vec{j} + w_3\vec{k}$

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = [(u_2v_3 - u_3v_2)\vec{i} - (u_1v_3 - u_3v_1)\vec{j} + (u_1v_2 - u_2v_1)\vec{k}] \cdot (w_1\vec{i} + w_2\vec{j} + w_3\vec{k})$$

$$= (u_2v_3 - u_3v_2)w_1 - (u_1v_3 - u_3v_1)w_2 + (u_1v_2 - u_2v_1)w_3 = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

From the properties of determinants:

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

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

differ by a minus sign.

§ 2 Straight Lines and Planes

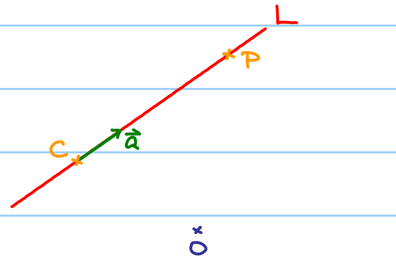
Straight line L in $\mathbb{R}^3$:

Let $C = (c_1, c_2, c_3)$ be a fixed point

$P = (x, y, z)$ be a movable point

$\vec{a} = (a_1, a_2, a_3)$ be a fixed vector (direction vector)

L be a straight line passes through C and goes along direction $\vec{a}$.



Then, we have $\overrightarrow{CP} \parallel \vec{a}$, i.e $\overrightarrow{CP} = t\vec{a}$, $t \in \mathbb{R}$

$(x-c_1, y-c_2, z-c_3) = t(a_1, a_2, a_3)$

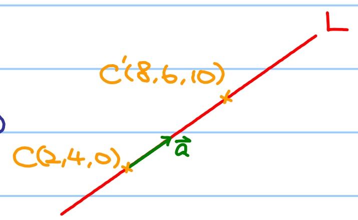
$$\therefore \begin{cases} x = c_1 + ta_1 \\ y = c_2 + ta_2 \\ z = c_3 + ta_3 \end{cases}$$ (parametric equation of L)

Eliminate t: $\frac{x-c_1}{a_1} = \frac{y-c_2}{a_2} = \frac{z-c_3}{a_3}$ if $a_1, a_2, a_3 \neq 0$.

(Think. If $a_1, a_2 \neq 0$, but $a_3 = 0$, then the equation becomes: $\frac{x-c_1}{a_1} = \frac{y-c_2}{a_2}$ and $z = c_3$.)

Example 2.1

If the equation of a straight line L in $\mathbb{R}^3$ is $\frac{x-2}{3} = y-4 = \frac{z}{5}$, then L passes through $(2,4,0)$ and goes along the direction $\vec{a} = (3,1,5)$

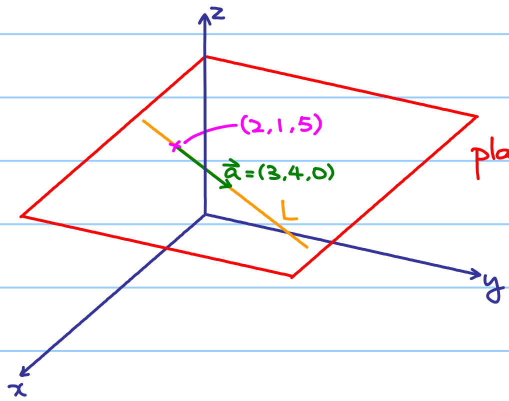


However, L also passes through the point $C' = (2,4,0) + 2(3,1,5) = (8,6,10)$.

Therefore, $\frac{x-8}{3} = y-6 = \frac{z-10}{5}$ is also an equation of L

Example 2.2

If the equation of a straight line L in $\mathbb{R}^3$ is $\frac{x-2}{3} = \frac{y-1}{4}$ and $z=5$, then L passes through $(2,1,5)$ and goes along the direction $\vec{a} = (3,4,0)$



plane Π : $z = 5$

Note: L lies on Π

Example 2.3

If L is a straight line in $\mathbb{R}^3$ given by the equation $\frac{x-2}{3}=\frac{y+1}{-2}=\frac{z-1}{2}$,

$Q=(10,-3,4)$ is a fixed point.

What is the shortest distance between L and Q?

L passes through $P(2,-1,1)$

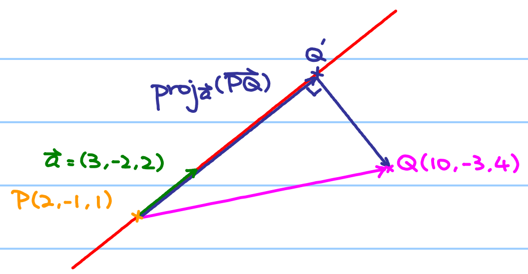
Direction vector of $L=\vec{a}=(3,-2,2)$

$\overrightarrow{PQ}=(10,-3,4)-(2,-1,1)=(8,-2,3)$

$\overrightarrow{PQ'}=\text{proj}_{\vec{a}}(\overrightarrow{PQ})=\frac{\overrightarrow{PQ}\cdot\vec{a}}{|\vec{a}|^2}\vec{a}=\frac{34}{17}\vec{a}=2\vec{a}=(6,-4,4)$

$\overrightarrow{Q'Q}=\overrightarrow{PQ}-\overrightarrow{PQ'}=(8,-2,3)-(6,-4,4)=(2,2,-1)$

Shortest distance between L and $Q=|\overrightarrow{Q'Q}|=\sqrt{2^2+2^2+(-1)^2}=3$



Follow the idea of the discussion of straight lines in $\mathbb{R}^3$, figure out the equation of straight lines in $\mathbb{R}^n$

In general, if L is a straight line in $\mathbb{R}^n$ which passes through a fixed point $\vec{c}=(c_1,c_2,\cdots,c_n)$ and goes along the direction $\vec{a}=(a_1,a_2,\cdots,a_n)$.

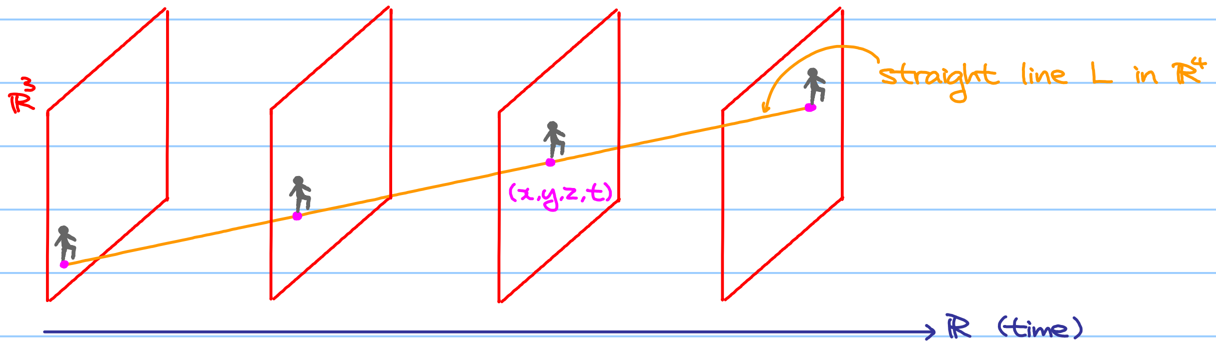
$\vec{x}=\vec{c}+t\vec{a}$, $t\in\mathbb{R}$ is a parametric equation of L, where $\vec{x}=(x_1,x_2,\cdots,x_n)$.

If $a_i\neq 0$ for all i, by eliminating t, we obtain $\frac{x_1-c_1}{a_1}=\frac{x_2-c_2}{a_2}=\cdots=\frac{x_n-c_n}{a_n}$.

(Think: What does the equation look like if some $a_i=0$?)

Need some imagination:

Somebody is walking in $\mathbb{R}^3$.

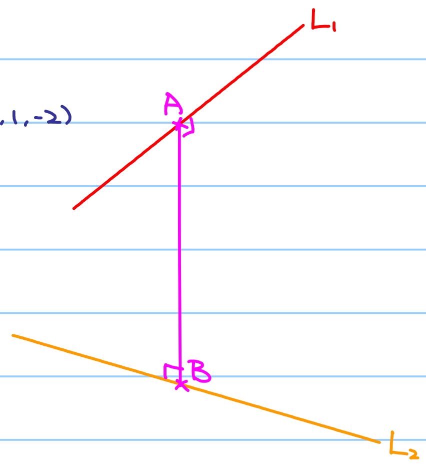


Example 2.4

Let $\vec{c_1} = (1,9,9,6)$, $\vec{a_1} = (2,-1,-3,2)$, $\vec{c_2} = (2,3,-2,7)$, $\vec{a_2} = (1,2,1,-2)$

Let $L_1 : \vec{x} = \vec{c_1} + t\vec{a_1}$ and $L_2 . \vec{x} = \vec{c_2} + s\vec{a_2}$, $t, s \in \mathbb{R}$, be two straight lines in $\mathbb{R}^4$.

Find the shortest distance between L_1 and L_2



Let $\vec{OA} = \vec{c_1} + t_o\vec{a_1}$, $\vec{OB} = \vec{c_2} + s_o\vec{a_2}$ for some $t_o, s_o \in \mathbb{R}$.

Then $\vec{BA} = \vec{OA} - \vec{OB} = (\vec{c_1} - \vec{c_2}) + t_o\vec{a_1} - s_o\vec{a_2} = (-1,6,11,-1) + t_o(2,-1,-3,2) - s_o(1,2,1,-2)$

Note : $\vec{BA} \perp \vec{a_1}$ and $\vec{BA} \perp \vec{a_2}$ give two equations:

$$\left.\begin{array}{lll} \vec{BA}\cdot\vec{a_1} = 0 & \Rightarrow & -43 + 7s_o + 18t_o = 0 \\ \vec{BA}\cdot\vec{a_2} = 0 & \Rightarrow & 24 - 10s_o - 7t_o = 0 \end{array}\right. \Rightarrow \begin{cases} s_o = 1 \\ t_o = 2 \end{cases}$$

Idea : 2 equations , 2 unknowns , it suffices to know s_o and t_o

$\therefore$ $\vec{AB} = (-1,6,11,-1) + 2(2,-1,-3,2) - (1,2,1,-2) = (2,2,4,5)$

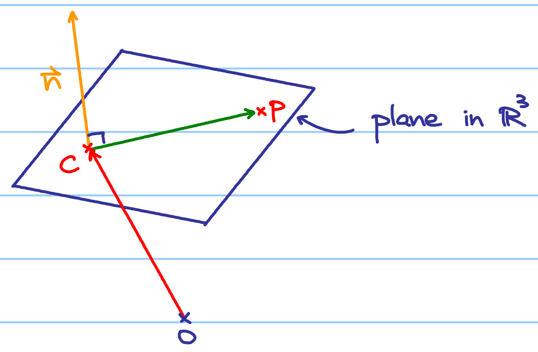
and the shortest distance between L_1 and $L_2 = |\vec{AB}| = \sqrt{2^2+2^2+4^2+5^2} = \sqrt{49} = 7$

Planes in $\mathbb{R}^3$:

Let $C = (c_1, c_2, c_3)$ be a fixed point on the plane.

$P = (x, y, z)$ be a movable point on the plane.

$\vec{n} = (A, B, C)$ be a normal of the plane.



Then, we have $\vec{n} \perp \vec{CP}$

i.e. $\vec{n} \cdot \vec{CP} = 0$

$$A(x-c_1) + B(y-c_2) + C(z-c_3) = 0$$

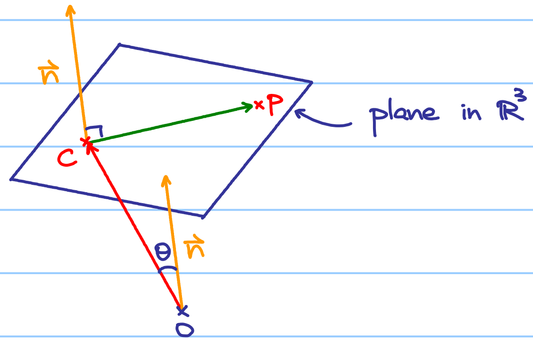
$$Ax + By + Cz + \underbrace{(-Ac_1 - Bc_2 - Cc_3)}_{\text{denote it by } D} = 0$$

$\therefore$ The equation of a plane in $\mathbb{R}^3$ is of the form $Ax + By + Cz + D = 0$

where $\vec{n} = A\vec{i} + B\vec{j} + C\vec{k}$ is a normal.

Furthermore, if d is the distance between O and the plane $\pi: Ax+By+Cz+D=0$ then $d=|\vec{c}|\cos\theta|$ where θ is the angle between $\vec{n}$ and $\vec{c}$.

$$d=\left||\vec{c}|\cos\theta\right|$$
$$=\left|\frac{|\vec{n}||\vec{c}|\cos\theta}{|\vec{n}|}\right|$$
$$=\frac{|\vec{n}\cdot\vec{c}|}{|\vec{n}|}$$
$$=\frac{|D|}{\sqrt{A^2+B^2+C^2}}$$



Example 2.5

$\pi: 2x-2y-z-3=0$ is a plane in $\mathbb{R}^3$ with a normal $2\vec{i}-2\vec{j}-\vec{k}$ in $\mathbb{R}^3$.

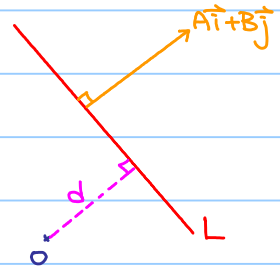
The distance between O and $\pi = \frac{|-3|}{\sqrt{2^2+(-2)^2+(-1)^2}} = 1$.

Exercise 2.1 (Revisit of straight lines in $\mathbb{R}^2$)

Follow the idea of the discussion of planes in $\mathbb{R}^3$, show that if $L: Ax+By+C=0$ is a straight line in $\mathbb{R}^2$, then

a) $\vec{n}=A\vec{i}+B\vec{j}$ gives a normal of L;

b) the distance between O and $L = d = \frac{|C|}{\sqrt{A^2+B^2}}$.



Example 2.6

Let $L: \frac{x-1}{2}=\frac{y-2}{-1}=\frac{z}{2}$ be a straight line and $\pi: x+y+z=0$ be a plane in $\mathbb{R}^3$.

a) Find the intersection of L and π

b) Find the angle between L and π

c) Find the projection of L on π

a) If P is a point lying on L, $P=(1,2,0)+t(2,-1,2)=(1+2t,2-t,2t)$, $t\in\mathbb{R}$.

Suppose that P further lies on π, $(1+2t)+(2-t)+2t=0$

$$3t+3=0$$
$$t=-1$$

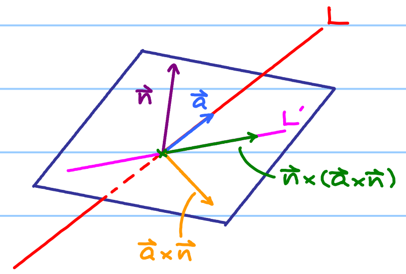
$\therefore$ L and π intersect at $(-1,3,-2)$.

b) Note: $\vec{a} = (2,-1,2)$ is a direction vector of L

$\vec{n} = (1,1,1)$ is a normal of π

The angle between L and $\vec{n}$ $= \cos^{-1}\left(\frac{\vec{a}\cdot\vec{n}}{|\vec{a}||\vec{n}|}\right) = \cos^{-1}\left(\frac{1}{\sqrt{3}}\right)$

$\therefore$ The angle between L and $\pi = \frac{\pi}{2} - \cos^{-1}\left(\frac{1}{\sqrt{3}}\right)$



c) Question: How to find a direction vector of L'?

$$\vec{a}\times\vec{n} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & -1 & 2 \\ 1 & 1 & 1 \end{vmatrix} = -3\vec{i} + 3\vec{k}$$

$$\vec{n}\times(\vec{a}\times\vec{n}) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 1 & 1 \\ -3 & 0 & 3 \end{vmatrix} = 3\vec{i} - 6\vec{j} + 3\vec{k} = 3(\vec{i} - 2\vec{j} + \vec{k})$$

$\therefore$ $\vec{i} - 2\vec{j} + \vec{k}$ is a direction vector of L'.

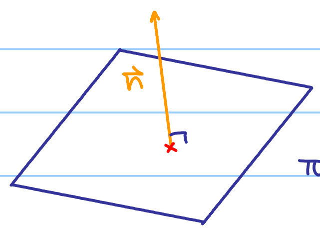
Equation of L': $x+1 = \frac{y-3}{-2} = z-2$

Follow the idea of the discussion of planes in $\mathbb{R}^3$,

the equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n + b = 0$ in $\mathbb{R}^n$ gives a "plane" in $\mathbb{R}^n$,

which is said to be an affine hyperplane π

The vector $\vec{n} = (a_1, a_2, \cdots, a_n)$ is a normal of the affine hyperplane π.



π: $(n-1)$-dim affine hyperplane in $\mathbb{R}^n$, n-dim space.

1-dim affine hyperplane in $\mathbb{R}^2$ is just an usual straight line in $\mathbb{R}^2$.

2-dim affine hyperplane in $\mathbb{R}^3$ is just an usual straight line in $\mathbb{R}^3$.

Example 2.7

Let π be a plane in $\mathbb{R}^4$ given by $2x_1 + x_2 - x_3 + 3x_4 = 4$ and let $P = (1,2,3,1)$ be a point on π.

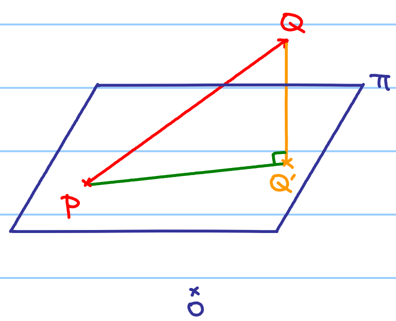
Also, let $Q = (2,5,7,4)$ be a point which does not lie on π.

What is the projection Q' of Q on π?

Note: $\vec{n} = (2,1,-1,3)$ is normal to π, so

$$\overrightarrow{QQ'} = \mathrm{proj}_{\vec{n}}(\overrightarrow{PQ}) = \frac{\overrightarrow{PQ}\cdot\vec{n}}{|\vec{n}|^2}\vec{n} = \frac{10}{15}\vec{n} = \frac{2}{3}(2,1,-1,3)$$

$$\therefore \overrightarrow{OQ'} = \overrightarrow{OQ} + \overrightarrow{QQ'} = \left(\frac{10}{3}, \frac{17}{3}, \frac{19}{3}, 6\right)$$



§3 Matrices and Determinants

Definition 3.1

A $m \times n$ real matrix is a rectangular array of real numbers (called entities)

with m rows and n columns.

The set of all $m \times n$ real matrices is denoted by $M_{m\times n}(\mathbb{R})$.

Let $A \in M_{m\times n}(\mathbb{R})$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$, the entry at i-th row, j-th column is denoted by a_{ij} or $[A]_{ij}$.

In particular, $A \in M_{n\times n}(\mathbb{R})$ is called a square matrix and we simply write $A \in M_n(\mathbb{R})$.

Example 3.1

$A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix}$ is a 2×3 real matrix (or simply 2×3 matrix)

We write $A \in M_{2\times 3}(\mathbb{R})$.

While $a_{12} = 3$, $a_{21} = 0$.

$O_{mn} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in M_{m\times n}(\mathbb{R})$ is said to be zero matrix.

Sometimes, it is simply denoted by O if no confusion occurs.

Definition 3.2 (Matrix Addition)

Let $A, B \in M_{m\times n}(\mathbb{R})$.

$A + B \in M_{m\times n}(\mathbb{R})$ which is defined by $[A+B]_{ij} = [A]_{ij} + [B]_{ij}$.

Example 3.2

Let $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 3 & 2 \\ 4 & 5 & 1 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$

Then $A + B = \begin{bmatrix} 2+(-1) & 3+3 & 1+2 \\ 0+4 & 4+5 & 5+1 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 3 \\ 4 & 9 & 6 \end{bmatrix}$

Definition 3.3 (Scalar Multiplication)

Let $A \in M_{m\times n}(\mathbb{R})$ and $r \in \mathbb{R}$.

$rA \in M_{m\times n}(\mathbb{R})$ which is defined by $[rA]_{ij} = r[A]_{ij}$.

Example 3.3

Let $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix} \in M_{2\times3}(\mathbb{R})$. Then $3A = \begin{bmatrix} 6 & 9 & 3 \\ 0 & 12 & 15 \end{bmatrix}$.

Definition 3.4

Let $A, B \in M_{m\times n}(\mathbb{R})$.

$-B$ is defined as $(-1)B$ and $A-B$ is defined as $A+(-B)$.

Proposition 3.1

Let $A, B, C \in M_{m\times n}(\mathbb{R})$, $s, t \in \mathbb{R}$.

1) (Commutative Law of Matrix Addition) $A+B = B+A$
2) (Associative Law of Matrix Addition) $(A+B)+C = A+(B+C)$
3) (Existence of Additive Identity) $O_{mn}+A = A+O_{mn} = A$
4) (Existence of Additive Inverse) $A+(-A) = (-A)+A = O_{mn}$
5) (Existence of Multiplicative Identity) $1A = A$
6) (Associative Law of Scalar Multiplication) $(st)A = s(tA)$
7) (Distributive Law of Scalar Multiplication) $s(A+B) = sA+sB$ and $(s+t)A = sA+tA$

Definition 3.5

Let $A \in M_{m\times n}(\mathbb{R})$

Each row can be regarded as a vector in $\mathbb{R}^n$, called a row vector

Each column can be regarded as a vector in $\mathbb{R}^m$, called a column vector.

$$\begin{array}{ccc} & \text{j-th column vector} & \\ A = & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2j} & & a_{2n} \\ \vdots & & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} & \begin{array}{l} \\ \\ \\ \text{i-th row vector} \\ \\ \\ \end{array} \end{array}$$

Therefore, $A \in M_{m\times n}(\mathbb{R})$ has m row vector (in $\mathbb{R}^n$) and n column vectors (in $\mathbb{R}^m$).

In particular, if $b \in M_{m\times1}(\mathbb{R})$, we may regard it as a vector in $\mathbb{R}^m$, and usually we write $\vec{b}$.

Definition 3.6 (Matrix Multiplication)

Let $A \in M_{m\times n}(\mathbb{R})$, $B \in M_{n\times p}(\mathbb{R})$

AB is defined as $C \in M_{m\times p}(\mathbb{R})$ with $c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$

$$C = m\left\{\overbrace{\begin{bmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{bmatrix}}^{p}\right. = m\left\{\overbrace{\begin{bmatrix} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \\ \end{bmatrix}}^{n}\right. n\left\{\overbrace{\begin{bmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{nj} & \end{bmatrix}}^{p}\right.$$

$$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

= dot product of i-th row vector of A and j-th column vector of B.

Example 3.4

Let $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix} \in M_{2\times3}(\mathbb{R})$, $B = \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 2 & 4 \end{bmatrix} \in M_{3\times2}(\mathbb{R})$, $C = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \in M_{2\times2}(\mathbb{R})$

$$AB = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} (2)(1)+(3)(0)+(1)(2) & (2)(3)+(3)(1)+(1)(4) \\ (0)(1)+(4)(0)+(5)(2) & (0)(3)+(4)(1)+(5)(4) \end{bmatrix} = \begin{bmatrix} 4 & 13 \\ 10 & 24 \end{bmatrix} \in M_{2\times2}(\mathbb{R})$$

$$BA \overset{\text{Ex}}{=} \begin{bmatrix} 2 & 15 & 16 \\ 0 & 4 & 5 \\ 4 & 22 & 22 \end{bmatrix} \in M_{3\times3}(\mathbb{R})$$

Remark. In general, $AB \neq BA$ (even they have different dimensions!)

$CA \overset{\text{Ex.}}{=} \begin{bmatrix} 4 & 10 & 7 \\ 6 & 25 & 23 \end{bmatrix}$ but AC is undefined.

Example 3.5

Let $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in M_{2\times2}(\mathbb{R})$.

$A, B \neq 0$, but $AB = 0$

Definition 3.7

Let $I_n = \begin{bmatrix} 1 & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{bmatrix} \in M_n(\mathbb{R})$, i.e. $[I_n]_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}$

then I_n is said to be identity matrix.

(Sometimes, we simply write I if no confusion occurs.)

Proposition 3.2

1) (Associative Law of Matrix Multiplication)

Let $A \in M_{m\times n}(\mathbb{R})$, $B \in M_{n\times p}(\mathbb{R})$, $C \in M_{p\times q}(\mathbb{R})$. Then $(AB)C = A(BC)$.

2) (Existence of Multiplicative Identity)

Let $A \in M_{m\times n}(\mathbb{R})$. Then $I_m A = A I_n = A$.

3) (Distributive Law of Matrix Multiplication)

Let $A \in M_{m\times n}(\mathbb{R})$, $B_1, B_2 \in M_{n\times p}(\mathbb{R})$, $C \in M_{p\times q}(\mathbb{R})$.

Then $A(B_1+B_2) = AB_1 + AB_2$ and $(B_1+B_2)C = B_1C + B_2C$.

4) Let $A \in M_{m\times n}(\mathbb{R})$, $B \in M_{n\times p}(\mathbb{R})$, $s \in \mathbb{R}$. Then $(sA)B = s(AB) = A(sB)$

proof of (1):

$$[(AB)C]_{ij} = \sum_{s=1}^{p} [AB]_{is}[C]_{sj} = \sum_{s=1}^{p}\left(\left(\sum_{r=1}^{n}[A]_{ir}[B]_{rs}\right)[C]_{sj}\right) = \sum_{\substack{1\le r\le n \\ 1\le s\le p}} [A]_{ir}[B]_{rs}[C]_{sj}$$

summing over all possible pairs of r, s.

$$[A(BC)]_{ij} = \sum_{r=1}^{n} [A]_{ir}[BC]_{rj} = \sum_{r=1}^{n}\left([A]_{ir}\left(\sum_{s=1}^{p}[B]_{rs}[C]_{sj}\right)\right) = \sum_{\substack{1\le r\le n \\ 1\le s\le p}} [A]_{ir}[B]_{rs}[C]_{sj}$$

proof of (2):

$$[I_m A]_{ij} = \sum_{r=1}^{m} [I_m]_{ir}[A]_{rj} = \underbrace{[I_m]_{i1}}_{0}[A]_{1j} + \cdots + \underbrace{[I_m]_{ii}}_{1}[A]_{ij} + \cdots + \underbrace{[I_m]_{im}}_{0}[A]_{mj} = [A]_{ij}$$

Matrix multiplication seems strange, but think:

$$(S)\begin{cases} 2x + 3y = 1 \\ 4x + 5y = 6 \end{cases}$$

(S) is a system of linear equations.

Let $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$

Then (S) can be written as $A\vec{x} = \vec{b}$.

Definition 3.8

Let $A \in M_{m\times n}(\mathbb{R})$

The transpose $A^T \in M_{n\times m}(\mathbb{R})$ of A is defined by $[A^T]_{ij} = [A]_{ji}$.

Example 3.6

Let $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$.

$$A^T = \begin{bmatrix} 2 & 0 \\ 3 & 4 \\ 1 & 5 \end{bmatrix}$$

Idea : diagonal

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 5 \end{bmatrix} \qquad A^T = \begin{bmatrix} 2 & 0 \\ 3 & 4 \\ 1 & 5 \end{bmatrix}$$

e.g. $[A^T]_{31} = [A]_{13} = 1$

Flipping along the diagonal!

Example 3.7

If $\vec{u} = (u_1, u_2, \cdots, u_n)$, $\vec{v} = (v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$, we reformulate them as $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in M_{n\times 1}(\mathbb{R})$

then $\vec{u}^T \vec{v} = [u_1 \ u_2 \ \cdots \ u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = [\underbrace{u_1v_1 + u_2v_2 + \cdots + u_nv_n}_{\vec{u}\cdot\vec{v}}] \in M_1(\mathbb{R})$

Therefore, sometimes we write $\vec{u}\cdot\vec{v}$ as $\vec{u}^T\vec{v}$ by regarding a 1×1 real matrix as a real number.

If we accept this, we have $|\vec{u}|^2 = \vec{u}^T\vec{u}$.

Proposition 3.3

Let $A, B \in M_{m\times n}(\mathbb{R})$, $C \in M_{n\times p}(\mathbb{R})$, $s \in \mathbb{R}$. Then

1) $(A^T)^T = A$

2) $(A+B)^T = A^T + B^T$

3) $(sA)^T = sA^T$

4) $(AC)^T = C^T A^T$

proof of (4) :

$$[(AC)^T]_{ij} = [AC]_{ji} = \sum_{r=1}^{n} [A]_{jr}[C]_{ri} = \sum_{r=1}^{n} [A^T]_{rj}[C^T]_{ir} = \sum_{r=1}^{n} [C^T]_{ir}[A^T]_{rj} = [C^T A^T]_{ij}$$

Definition 3.9

Let $A \in M_n(\mathbb{R})$.

A is said to be symmetric if $A^T = A$;

A is said to be antisymmetric (or skew symmetric) if $A^T = -A$.

Definition 3.10 (Determinant of a Square Matrix)

Determinant of a Square Matrix A is denoted by $\det(A)$ or $|A|$, which is defined by:

1) Let $A \in M_1(\mathbb{R})$, i.e. $A = [a_{11}]$

$\det(A) = a_{11}$.

(Expanding along the first row)

2) Let $A \in M_2(\mathbb{R})$, i.e. $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

Draw a table: $\begin{vmatrix} + & - \\ - & + \end{vmatrix}$

delete delete

$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

$\det A = +a_{11}|a_{22}| - a_{12}|a_{21}| = a_{11}a_{22} - a_{12}a_{21}$

NOT abs. value but det!

$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ × × −

3) Let $A \in M_3(\mathbb{R})$, i.e. $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Draw a table: $\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$

delete

$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

$$\det A = +a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{23}a_{31}$$

$$= \sum_{\sigma \in S_3} \operatorname{sign}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)}$$ (Fancy way)

Summing over all permutation of 1, 2, 3

$\det(A)$ is defined inductively.

Example 3.8

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 2 & 5 & 3 \end{bmatrix} \in M_3(\mathbb{R})$

Draw a table $\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$

Actually, the sign at i-th row, j-th column $= (-1)^{i+j}$

Expanding along the first row:

$$\det(A) = +1\begin{vmatrix} 1 & 4 \\ 5 & 3 \end{vmatrix} - 2\begin{vmatrix} 0 & 4 \\ 2 & 3 \end{vmatrix} + 3\begin{vmatrix} 0 & 1 \\ 2 & 5 \end{vmatrix} = 1(-17) - 2(-8) + 3(-2) = -7$$

Expanding along the second row:

$$\det(A) = -0\begin{vmatrix} 2 & 3 \\ 5 & 3 \end{vmatrix} + 1\begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix} - 4\begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = 1(-3) - 4(1) = -7$$

Expanding along the first column:

$$\det(A) = +1\begin{vmatrix} 1 & 4 \\ 5 & 3 \end{vmatrix} - 0\begin{vmatrix} 2 & 3 \\ 5 & 3 \end{vmatrix} + 2\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} = 1(-17) + 2(5) = -7$$

No matter which row or column we expand along, the answers are always the same!

($\therefore$ Just pick a row or column with more zeros!)

Definition 3.11

Let $A \in M_n(\mathbb{R})$.

Minor M_{ij} of A is defined as the determinant of the $(n-1)\times(n-1)$ submatrix obtained from deleting i-th row and j-th column of A.

delete j-th column

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2j} & & a_{2n} \\ \vdots & & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

delete i-th row

$$M_{ij} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1\,j-1} & \cdots & a_{1\,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2\,j-1} & \cdots & a_{2\,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ a_{i-1\,1} & a_{i-1\,2} & \cdots & a_{i-1\,j-1} & \cdots & a_{i-1\,j+1} & \cdots & a_{in} \\ a_{i+1\,1} & a_{i+1\,2} & \cdots & a_{i+1\,j-1} & \cdots & a_{i+1\,j+1} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m\,j-1} & \cdots & a_{m\,j+1} & \cdots & a_{mn} \end{vmatrix}$$

With the above definition, we have

$$\det A = \sum_{r=1}^{n} [A]_{ir} (-1)^{i+r} M_{ir} = \sum_{r=1}^{n} a_{ir} (-1)^{i+r} M_{ir}$$ (Expanding along the i-th row)

$$= \sum_{r=1}^{n} [A]_{rj} (-1)^{r+j} M_{rj} = \sum_{r=1}^{n} a_{rj} (-1)^{r+j} M_{rj}$$ (Expanding along the j-th column)

Question : Any meaning of det(A) ?

1) If $\vec{v} = a_{11}\vec{i} \in \mathbb{R}^1$, regard it as a row vector of $A = [a_{11}] \in M_1(\mathbb{R})$.

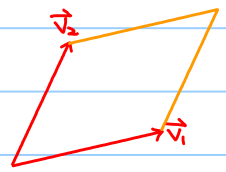
det $A = a_{11}$ = sided length of $\vec{v}$.

2) If $\vec{v_1} = a_{11}\vec{i} + a_{12}\vec{j}$, $\vec{v_2} = a_{21}\vec{i} + a_{22}\vec{j} \in \mathbb{R}^2$, regard them as row vectors of $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

det $A = a_{11}a_{22} - a_{12}a_{21}$

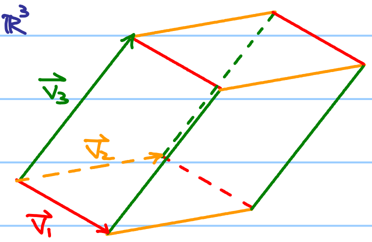
= signed area of parallelogram

spanned by $\vec{v_1}$ and $\vec{v_2}$



3) If $\vec{v_1} = a_{11}\vec{i} + a_{12}\vec{j} + a_{13}\vec{k}$, $\vec{v_2} = a_{21}\vec{i} + a_{22}\vec{j} + a_{23}\vec{k}$, $\vec{v_3} = a_{31}\vec{i} + a_{32}\vec{j} + a_{33}\vec{k} \in \mathbb{R}^3$

regard them as row vectors of $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$



det $A = (\vec{v_1} \times \vec{v_2}) \cdot \vec{v_3}$

= signed volume of the parallelepiped spanned by $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$.

In general, if $A \in M_n(\mathbb{R})$, det(A) = signed volume of parallelotope in $\mathbb{R}^n$ spanned by row vectors

Proposition 3.4

Let $A, B \in M_n(\mathbb{R})$. Then

1) det A^T = det A.

2) det(AB) = (det A)(det B).

proof of (1) :

Prove by induction on n.

(1) When $n=1$, let $A = (a) \in M_1(\mathbb{R})$.

Then $A^T = A = (a)$ and so det A^T = det $A = a$.

(2) Assume that for any $B \in M_{n-1}(\mathbb{R})$, we have det B^T = det B.

$$\det A = \sum_{r=1}^{n} [A]_{ir}(-1)^{i+r} M_{ir} \qquad \text{(Expanding along the i-th row)}$$

$$= \sum_{r=1}^{n} [A^T]_{ri}(-1)^{r+i} M_{ri} \qquad (M_{ir} = M_{ri} \text{ by induction assumption})$$

$$= \det(A^T) \qquad \text{(Expanding along the i-th column)}$$

Remark :

Let $A, B \in M_n(\mathbb{R})$. AB may not equal to BA, but det(AB) = det(A) det(B) = det(B) det(A) = det(BA).

Definition 3.12

Let $A \in M_n(\mathbb{R})$.

A is said to be an upper (a lower) triangular matrix if $a_{ij} = 0$ for $i > j$ $(j < i)$

A is said to be a diagonal matrix if $a_{ij} = 0$ for $i \neq j$.

$$\begin{bmatrix} \ddots & * \\ 0 & \ddots \end{bmatrix} \text{diagonal} \qquad \begin{bmatrix} \ddots & 0 \\ * & \ddots \end{bmatrix} \text{diagonal} \qquad \begin{bmatrix} \ddots & 0 \\ 0 & \ddots \end{bmatrix} \text{diagonal}$$

upper triangular matrix lower triangular matrix diagonal matrix

Exercise 3.1

1) Let $A \in M_n(\mathbb{R})$. If A is an upper triangular, a lower triangular or a diagonal matrix, show that $\det(A) = \prod_{i=1}^{n} a_{ii} = a_{11}a_{22}\cdots a_{nn}$.

   In particular, $\det(I_n) = 1$.

2) Let $A \in M_n(\mathbb{R})$. If n is odd and A is antisymmetric, show that $\det A = 0$

3) If $A \in M_n(\mathbb{R})$ such that $A^TA = A^TA = I$, then A is said to be an orthogonal matrix.

   In this case, show that $\det A = \pm 1$.

   Remark: Note that $A^TA = I \Rightarrow A^T = A^{-1} \Rightarrow AA^T = I$, so if we only know $A^TA = I$ (or $AA^T = I$), it suffices to conclude that A is orthogonal.

Elementary Matrices:

Let $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$ where $k \neq 0$.

What are they? Recall: $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

1) E_1 is obtained by multiplying the third row of I by k.
2) E_2 is obtained by swapping the first and second row of I.
3) E_3 is obtained from I by multiplying the second row by k and then adding to the third row.

Exercise 3.2

1) Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in M_3(\mathbb{R})$.

Find E_1A, E_2A, E_3A and compare them with A.

$E_1A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{bmatrix}$ $E_2A = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ $E_3A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ka_{21}+a_{31} & ka_{22}+a_{32} & ka_{23}+a_{33} \end{bmatrix}$

1) E_1A is obtained by multiplying the third row of A by k.

2) E_2A is obtained by swapping the first and second row of A.

3) E_3A is obtained from I by multiplying the second row by k and then adding to the third row.

2) Compute $\det(E_1)$, $\det(E_2)$ and $\det(E_3)$.

$\det(E_1) = k$, $\det(E_2) = -1$ and $\det(E_3) = 1$.

Direct consequence :

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{vmatrix} = \det(E_1A) = \det(E_1)\det(A) = k\det A = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Multiplying a row of A by k, the determinant is multiplied by k.
(Taking out a common factor k from a row)

$$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \det(E_2A) = \det(E_2)\det(A) = -\det A = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Swapping two rows of A, the determinant is changed by a $\pm$ sign

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ka_{21}+a_{31} & ka_{22}+a_{32} & ka_{23}+a_{33} \end{vmatrix} = \det(E_3A) = \det(E_3)\det(A) = \det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Multiplying a row by k and then adding to another row, the determinant is unchanged.

In general, we have:

Definition 3.13

The following operations are called elementary row operations:

1) Multiplying i-th row by k : $kR_i \to R_i$

2) Swapping i-th and j-th row : $R_i \leftrightarrow R_j$

3) Multiplying j-th row by k and adding to i-th row : $R_i + kR_j \to R_i$

If $E \in M_n(\mathbb{R})$ is a matrix obtained by applying one of the above operations on $I_n \in M_n(\mathbb{R})$, E is called an elementary matrix.

Furthermore, let $A \in M_{n\times m}(\mathbb{R})$,

EA is exactly the matrix obtained by applying the same operation on A.

Proposition 3.5

Let $E_1, E_2, E_3 \in M_n(\mathbb{R})$ be the three types of elementary matrice in definition 3.11.

Then, $\det(E_1) = k$, $\det(E_2) = -1$, $\det(E_3) = 1$.

Proposition 3.6

Let $A \in M_n(\mathbb{R})$. Then,

1) Multiplying a row of A by k, the determinant is multiplied by k $(= \det(E_1A))$

2) Swapping two rows of A, the determinant is changed by a $\pm$ sign $(= \det(E_2A))$

3) Multiplying a row by k and then adding to another row, the determinant is unchanged.
$(= \det(E_3A))$

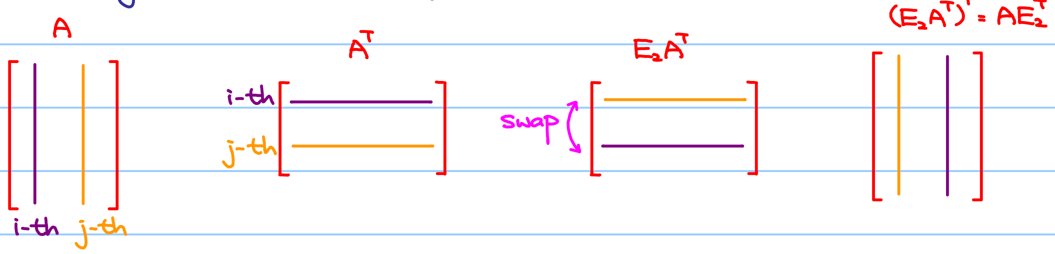
Exercise 3.3

Let $A \in M_{m\times n}(\mathbb{R})$. Then $A^T \in M_{n\times m}(\mathbb{R})$.

Let $E_1, E_2, E_3 \in M_n(\mathbb{R})$ be the three types of elementary matrice in definition 3.11.

Then, show that

1) $(E_1A^T)^T = AE_1^T$ is the matrix of obtained by multiplying i-th column of A by k.
2) $(E_2A^T)^T = AE_2^T$ is the matrix of obtained by swapping i-th and j-th column of A.
3) $(E_3A^T)^T = AE_3^T$ is the matrix of obtained by multiplying j-th column of A by k and adding to i-th column of A.



Proposition 3.7

Let $A \in M_n(\mathbb{R})$. Then,

1) Multiplying a column of A by k, the determinant is multiplied by k $(= \det(AE_1^T))$
2) Swapping two columns of A, the determinant is changed by a $\pm$ sign $(= \det(AE_2^T))$
3) Multiplying a column by k and then adding to another column, the determinant is unchanged $(= \det(AE_3^T))$

Example 3.9

Let $A = \begin{bmatrix} 0 & 4 & 4 \\ 1 & 2 & 3 \\ 2 & 7 & 15 \end{bmatrix} \in M_3(\mathbb{R})$

$$\begin{vmatrix} 0 & 4 & 4 \\ 1 & 2 & 3 \\ 2 & 7 & 15 \end{vmatrix} \overset{R_1 \leftrightarrow R_2}{=} - \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 4 \\ 2 & 7 & 15 \end{vmatrix} \overset{R_3+(-2)R_1 \to R_3}{=} - \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 4 \\ 0 & 3 & 9 \end{vmatrix} = -(4)(3) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 1 & 3 \end{vmatrix}$$

$$\overset{R_3+(-1)R_2 \to R_3}{=} -(4)(3) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{vmatrix} = -24$$

Alternative method:

$$\begin{vmatrix} 0 & 4 & 4 \\ 1 & 2 & 3 \\ 2 & 7 & 15 \end{vmatrix} \overset{C_2+(-1)C_3 \to C_2}{=} \begin{vmatrix} 0 & 0 & 4 \\ 1 & -1 & 3 \\ 2 & -8 & 15 \end{vmatrix} = 4 \begin{vmatrix} 1 & -1 \\ 2 & -8 \end{vmatrix} = 4(-6) = -24$$

Exercise 3.4

Let $A \in M_n(\mathbb{R})$. Show that

1) If there is a row or column of A with all zeros, then $\det A = 0$.

(Hint: Compute $\det(A)$ by expanding along that row or column.)

2) If there are two rows or columns of A with same entities, then $\det A = 0$.

(Hint: If $R_i = R_j$, perform $R_i + (-1)R_j \to R_i$.)

3) If $k \in \mathbb{R}$, then $\det(kA) = k^n \det(A)$.

Definition 3.14

Let $A \in M_n(\mathbb{R})$. If there exists $B \in M_n(\mathbb{R})$ such that $AB = I = BA$,

then B is said to be an inverse of A (symmetrically, A is an inverse of B)

In this case, A (also B) is said to be an invertible matrix.

Question:

1) Existence? How to find?

2) Uniqueness?

Definition 3.15

Let $A \in M_n(\mathbb{R})$, M_{ij} are minors of A (see definition 2.8).

The cofactor matrix $\mathrm{cof}(A)$ of A is defined by $[\mathrm{cof}(A)]_{ij} = (-1)^{i+j} M_{ij}$ and

the adjugate matrix $\mathrm{adj}(A)$ of A is defined by $\mathrm{adj}(A) = \mathrm{cof}(A)^T$, i.e. $[\mathrm{adj}(A)]_{ij} = (-1)^{j+i} M_{ji}$.

$$\mathrm{cof}(A) = \begin{bmatrix} M_{11} & -M_{12} & \cdots & (-1)^{1+n} M_{1n} \\ -M_{21} & M_{22} & \cdots & (-1)^{2+n} M_{2n} \\ \vdots & \vdots & & \vdots \\ (-1)^{n+1} M_{n1} & (-1)^{n+2} M_{n2} & \cdots & M_{nn} \end{bmatrix} \qquad \mathrm{adj}(A) = \begin{bmatrix} M_{11} & -M_{21} & \cdots & (-1)^{n+1} M_{n1} \\ -M_{12} & M_{22} & \cdots & (-1)^{n+2} M_{n2} \\ \vdots & \vdots & & \vdots \\ (-1)^{1+n} M_{1n} & (-1)^{2+n} M_{2n} & \cdots & M_{nn} \end{bmatrix}$$

Proposition 3.8

$A\ \mathrm{adj}(A) = (\det A) I = \mathrm{adj}(A)\ A$

proof of the first equality:

$$[A\ \mathrm{adj}(A)]_{ij} = \sum_{r=1}^{n} a_{ir} [\mathrm{adj}\ A]_{rj} = \sum_{r=1}^{n} a_{ir} (-1)^{j+r} M_{jr} = \begin{cases} \det A & \text{if } i=j \\ 0 & \text{if } i \neq j \quad (*) \end{cases}$$

Recall:

$$\det A = \sum_{r=1}^{n} a_{ir} (-1)^{i+r} M_{ir}$$

(Expanding along the i-th row)

$$= \sum_{r=1}^{n} a_{rj} (-1)^{r+j} M_{rj}$$

(Expanding along the j-th column)

$\therefore A\ \mathrm{adj}(A) = (\det A) I$

Why $(*)$? For $i \neq j$,

$$A = \begin{bmatrix} \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & & & \vdots \end{bmatrix} \xrightarrow[\text{i-th row}]{\text{Replace j-th row by}} \tilde{A} = \begin{bmatrix} \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \end{bmatrix} \begin{array}{l} \\ \leftarrow \text{i-th row} \\ \\ \leftarrow \text{j-th row} \\ \\ \end{array}$$

Note: $\tilde{M}_{jr} = M_{jr}$

$$\sum_{r=1}^{n} a_{ir} (-1)^{j+r} M_{jr} = \sum_{r=1}^{n} a_{ir} (-1)^{j+r} \tilde{M}_{jr} \quad \text{(Expanding along the j-th row of } \tilde{A}\text{)}$$

$$= \det \tilde{A} = 0$$

Direct consequence / Answer of (1):

If $\det A \neq 0$, $A\left(\frac{1}{\det A} \mathrm{adj}(A)\right) = \left(\frac{1}{\det A} \mathrm{adj}(A)\right) = I$, i.e. $\frac{1}{\det A} \mathrm{adj}(A)$ is an inverse of A.

However, if $\det A = 0$, does it imply that A has no inverse? (Answer later!)

Answer of (2):

Proposition 3.9

Let $A \in M_n(\mathbb{R})$. If B and C are both inverse matrices of A, then $B = C$

proof. By assumption $AB = I = BA$, $AC = I = CA$.

Then, $AB = I$

$$\underbrace{(CA)}_{I}B = C(AB) = C$$

$$\therefore B = C$$

Therefore, once inverse of A exists, it must be unique, and we denote it by A^{-1}.

Proposition 3.10

Let $A \in M_n(\mathbb{R})$. A is invertible if and only if $\det A \neq 0$.

proof:

$(\Rightarrow)$: If A is invertible, i.e there exists $A^{-1} \in M_n(\mathbb{R})$ such that $AA^{-1} = I = A^{-1}A$

then $\det A \det A^{-1} = \det(AA^{-1}) = \det I = 1$.

$\therefore \det A \neq 0$

$(\Leftarrow)$: If $\det A \neq 0$, from proposition 3.8, we have $\left(\frac{1}{\det A} \operatorname{adj}(A)\right) \cdot A = I = A \cdot \left(\frac{1}{\det A} \operatorname{adj}(A)\right)$

A^{-1} exists and $A^{-1} = \frac{1}{\det A} \operatorname{adj}(A)$

Remark:

Let $A, B \in M_n(\mathbb{R})$ such that $AB = I$. Is it true that $B = A^{-1}$?

$AB = I \Rightarrow \det A \det B = \det(AB) = \det I = 1 \Rightarrow \det A \neq 0$

Therefore, inverse of A exists and $A^{-1} = \frac{1}{\det A} \operatorname{adj}(A)$.

From $AB = I$

$A^{-1}(AB) = A^{-1} I$

$\therefore B = A^{-1}$

From now on, if we want to check if B is the inverse of A, it suffices to check $AB = I$ (or $BA = I$).

Example 3.10

Let $A = \begin{bmatrix} 8 & 3 \\ 5 & 2 \end{bmatrix}$

$\det A = \begin{vmatrix} 8 & 3 \\ 5 & 2 \end{vmatrix} = 1 \neq 0 \Rightarrow A^{-1}$ exists.

$\operatorname{cof}(A) = \begin{bmatrix} 2 & -5 \\ -3 & 8 \end{bmatrix}$ and $\operatorname{adj}(A) = \begin{bmatrix} 2 & -3 \\ -5 & 8 \end{bmatrix}$

$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A) = \begin{bmatrix} 2 & -3 \\ -5 & 8 \end{bmatrix}$

Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.

$\det A = 1 \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1 \neq 0 \Rightarrow A^{-1}$ exists.

$\operatorname{cof}(A) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & -1 & -1 \\ -2 & 0 & 1 \end{bmatrix}$ and $\operatorname{adj}(A) = \begin{bmatrix} 1 & 2 & -2 \\ 0 & -1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$

$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A) = \begin{bmatrix} -1 & -2 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$

§4 System of Linear Equations

A system of linear equations :

Given a_{ij}'s and b_i's, we want to find x_i's which satisfy the following equations .

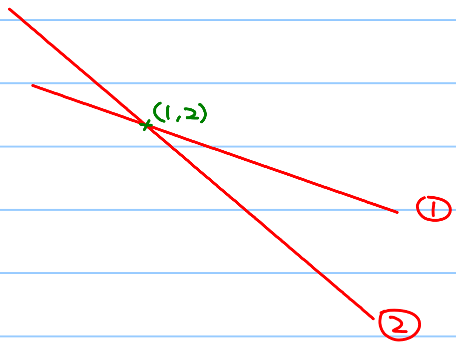
$$(S):\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

or simply write $A\vec{x} = \vec{b}$ where $A \in M_{m\times n}(\mathbb{R})$, $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in M_{n\times 1}(\mathbb{R})$, $\vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in M_{m\times 1}(\mathbb{R})$.

Idea : We have m linear equations which define m affine hyperplanes in $\mathbb{R}^n$,
a solution of (S) is an intersection point of those m affine hyperplanes

Example 4.1

Let $(S): \begin{cases} x + 3y = 7 & — ① \\ 2x + 5y = 12 & — ② \end{cases}$



$② + (-2)\times ① \to ②$: $\begin{cases} x + 3y = 7 & — ① \\ -y = -2 & — ③ \end{cases}$

$(-1)\times ③ \to ③$: $\begin{cases} x + 3y = 7 & — ① \\ y = 2 & — ④ \end{cases}$

$\therefore y = 2$, put $y = 2$ into ① , $x + 3(2) = 7$ and so $x = 1$

Actually we do not have to keep track on unknowns :

$$\left[\begin{array}{cc:c} 1 & 3 & 7 \\ 2 & 5 & 12 \end{array}\right]$$

$$R_2 + (-2)\times R_1 \to R_2 \left[\begin{array}{cc:c} 1 & 3 & 7 \\ 0 & -1 & -2 \end{array}\right]$$

$$(-1)\times R_2 \to R_2 \left[\begin{array}{cc:c} 1 & 3 & 7 \\ 0 & 1 & 2 \end{array}\right]$$

Another interpretation:

Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 7 \\ 12 \end{bmatrix}$

$$\left[\begin{array}{cc|c} 1 & 3 & 7 \\ 2 & 5 & 12 \end{array}\right] \quad \longleftarrow \text{called augmented matrix} \qquad [A \,\vdots\, \vec{b}]$$

$$R_2 + (-2) \times R_1 \to R_2 \left[\begin{array}{cc|c} 1 & 3 & 7 \\ 0 & -1 & -2 \end{array}\right] \qquad \text{Let } E_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \qquad [E_1 A \,\vdots\, E_1 \vec{b}]$$

$$(-1) \times R_2 \to R_2 \left[\begin{array}{cc|c} 1 & 3 & 7 \\ 0 & 1 & 2 \end{array}\right] \qquad \text{Let } E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad [E_2 E_1 A \,\vdots\, E_2 E_1 \vec{b}]$$

$$\begin{cases} x + 3y = 7 \\ \quad\;\;\, y = 2 \end{cases}$$

In fact, we can go further:

$$R_1 + (-3) \times R_2 \to R_1 \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right] \qquad \text{Let } E_3 = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \qquad [\underbrace{E_3 E_2 E_1 A}_{= I} \,\vdots\, \underbrace{E_3 E_2 E_1 \vec{b}}_{= \begin{bmatrix} 1 \\ 2 \end{bmatrix}}]$$

$\therefore \quad x = 1, \; y = 2$ $\qquad \vec{x} = I\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Definition 4.1

Let $A \in M_{m \times n}(\mathbb{R})$. A is in row echelon form if

1) all nonzero rows are above any rows of all zeros;
2) the leading (nonzero) entry of a nonzero row must be 1;
3) the leading 1 of a nonzero row is always strictly to the left of the leading 1 of the next nonzero row.

A is in reduced row echelon form if it further satisfies

4) For each leading 1, it is the only nonzero entry in its column.

Example 4.2

The following matrices are in row echelon form

$$\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

in reduced row echelon form

but the following are not

$$\begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Gaussian Elimination :

Goal : For a system of linear equations $[A \,\vdots\, \vec{b}]$, we perform elementary row operations, transform it as $[A_0 \,\vdots\, \vec{b}_0]$ such that A_0 is in row echelon form (or even reduced row echelon form) and solve.

Example 4.3

$$\left[\begin{array}{ccc|c} 0 & 3 & 1 & -1 \\ 2 & 1 & 0 & 1 \\ 1 & 4 & 2 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & 4 & 2 & 1 \\ 2 & 1 & 0 & 1 \\ 0 & 3 & 1 & -1 \end{array}\right] \xrightarrow{R_2+(-2)R_1 \to R_2} \left[\begin{array}{ccc|c} 1 & 4 & 2 & 1 \\ 0 & -7 & -4 & -1 \\ 0 & 3 & 1 & -1 \end{array}\right] \xrightarrow{(-\frac{1}{7})R_2 \to R_2} \left[\begin{array}{ccc|c} 1 & 4 & 2 & 1 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 3 & 1 & -1 \end{array}\right]$$

$$\xrightarrow{R_3+(-3)R_2 \to R_3} \left[\begin{array}{ccc|c} 1 & 4 & 2 & 1 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 0 & -5/7 & -10/7 \end{array}\right] \xrightarrow{(-\frac{7}{5})R_3 \to R_3} \left[\begin{array}{ccc|c} 1 & 4 & 2 & 1 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

row echelon form

i.e. $\begin{cases} x_1 + 4x_2 + 2x_3 = 1 \\ x_2 + \frac{4}{7}x_3 = \frac{1}{7} \\ x_3 = 2 \end{cases}$ $\Rightarrow x_1 + 4(-1) + 2(2) = 1 \Rightarrow x_1 = 1$; $\Rightarrow x_2 + \frac{4}{7}(2) = \frac{1}{7} \Rightarrow x_2 = -1$ $\quad \therefore x_1 = 1,\ x_2 = -1,\ x_3 = 2$

backward substitution

OR do further

$$\left[\begin{array}{ccc|c} 1 & 4 & 2 & 1 \\ 0 & 1 & 4/7 & 1/7 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow[R_1+(-2)R_3 \to R_1]{R_2+(-\frac{4}{7})R_3 \to R_2} \left[\begin{array}{ccc|c} 1 & 4 & 0 & -3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{R_1+(-4)R_2 \to R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right] \quad \therefore x_1 = 1,\ x_2 = -1,\ x_3 = 2$$

reduced row echelon form

Example 4.4

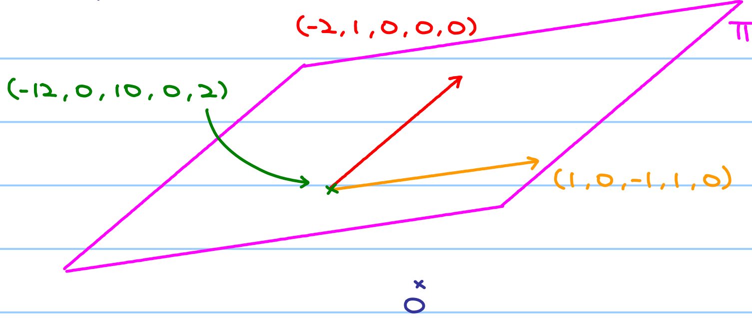
$$\left[\begin{array}{ccccc|c} 1&2&1&0&2&2\\ 2&4&3&1&0&6\\ 3&6&4&1&5&14 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1&2&1&0&2&2\\ 0&0&1&1&-4&2\\ 0&0&1&1&-1&8 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1&2&1&0&2&2\\ 0&0&1&1&-4&2\\ 0&0&0&0&3&6 \end{array}\right] \longrightarrow \underbrace{\left[\begin{array}{ccccc|c} 1&2&1&0&2&2\\ 0&0&1&1&-4&2\\ 0&0&0&0&1&2 \end{array}\right]}_{\text{row echelon form}}$$

$$\therefore \begin{cases} x_1+2x_2+x_3+2x_5=2 & \Rightarrow x_1+2x_2+(10-t)+2(2)=2 \Rightarrow x_1=-12-2x_2+t=-12-2s+t \ (\text{let } x_2=s\in\mathbb{R}) \\ x_3+x_4-4x_5=2 & \Rightarrow x_3+x_4-4(2)=2 \Rightarrow x_3=10-x_4=10-t \quad (\text{let } x_4=t\in\mathbb{R}) \\ x_5=2 \end{cases}$$

(x_2, x_4 can be any real number, so they are called free variables.)

$$\therefore (x_1,x_2,x_3,x_4,x_5)=(-12-2s+t,\ s,\ 10-t,\ t,\ 2)$$
$$=(-12,0,10,0,2)+s(-2,1,0,0,0)+t(1,0,-1,1,0)$$
where $t,s\in\mathbb{R}$

In $\mathbb{R}^5$,



π is a 2-dimensional affine subspace passing through $(-12,0,10,0,2)$ spanned by $(-2,1,0,0,0)$ and $(1,0,-1,1,0)$.

Every point on π is a solution.

OR do further

$$\left[\begin{array}{ccccc|c} 1&2&1&0&2&2\\ 0&0&1&1&-4&2\\ 0&0&0&0&1&2 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1&2&1&0&0&-2\\ 0&0&1&1&0&10\\ 0&0&0&0&1&2 \end{array}\right] \longrightarrow \underbrace{\left[\begin{array}{ccccc|c} 1&2&0&-1&0&-12\\ 0&0&1&1&0&10\\ 0&0&0&0&1&2 \end{array}\right]}_{\text{reduced row echelon form}}$$

$$\therefore \begin{cases} x_1+2x_2-x_4=-12 & \Rightarrow x_1=-12-2x_2+x_4=-12-2s+t \ (\text{let } x_2=s\in\mathbb{R}) \\ x_3+x_4=10 & \Rightarrow x_3=10-x_4=10-t \quad (\text{let } x_4=t\in\mathbb{R}) \\ x_5=2 \end{cases}$$

$$\therefore (x_1,x_2,x_3,x_4,x_5)=(-12-2s+t,\ s,\ 10-t,\ t,\ 2)$$

Example 4.5

$$\left[\begin{array}{ccc|c} 1&2&1&1\\ 1&3&2&0\\ 2&1&1&3\\ 3&8&4&3 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1&2&1&1\\ 0&1&1&-1\\ 0&-3&-1&1\\ 0&2&1&0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1&2&1&1\\ 0&1&1&-1\\ 0&0&2&-2\\ 0&0&-1&2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1&2&1&1\\ 0&1&1&-1\\ 0&0&1&-1\\ 0&0&-1&2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1&2&1&1\\ 0&1&1&-1\\ 0&0&1&-1\\ 0&0&0&1 \end{array}\right]$$

The last equation is $0x_1+0x_2+0x_3=1$, which is impossible!

There is no solution!

The elementary row operation also gives another way to find the inverse of a matrix.

Example 4.6

Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \in M_3(\mathbb{R})$.

$$\underset{A \qquad\qquad I}{\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]} \longrightarrow \underset{E_1A \qquad\qquad E_1}{\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{array}\right]} \longrightarrow \underset{E_2E_1A \qquad\qquad E_2E_1}{\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array}\right]} \longrightarrow \underset{E_3E_2E_1A \qquad\qquad E_3E_2E_1}{\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right]} \longrightarrow \underset{E_4E_3E_2E_1A \qquad\qquad E_4E_3E_2E_1}{\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & -1 & -2 & 2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right]}$$

Note that $E_4E_3E_2E_1A = I$, so $A^{-1} = E_4E_3E_2E_1 = \begin{bmatrix} -1 & -2 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$

(Compare to example 3.10)

Let $\vec{x}_1$ and $\vec{x}_2$ be solutions of $A\vec{x} = \vec{b}$.

If we let $\vec{u} = \vec{x}_1 + t(\vec{x}_2 - \vec{x}_1)$, $t \in \mathbb{R}$, then $A\vec{u} = A\vec{x}_1 + tA(\vec{x}_2 - \vec{x}_1) = A\vec{x}_1 + t(A\vec{x}_2 - A\vec{x}_1) = \vec{b} + t(\vec{b} - \vec{b}) = \vec{b}$.

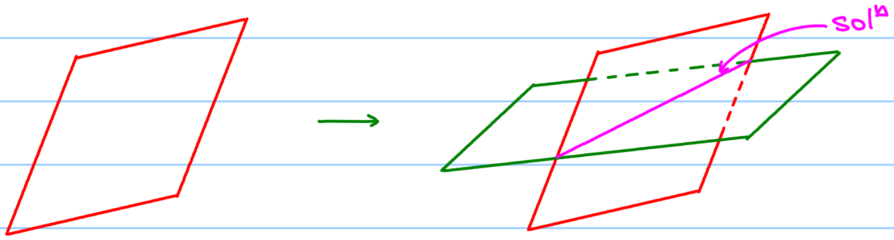
Therefore, every point lies on the line joining $\vec{x}_1$ and $\vec{x}_2$ is also a solution of $A\vec{x} = \vec{b}$.

That means $A\vec{x} = \vec{b}$ cannot have a solution set consisting of more than one discrete points.

Idea :

Impose an equation "=" Impose a constrain

Impose an extra equation "=" Impose an extra constrain



Cut down the dimension of solution set ? Depends on how those planes intersect !

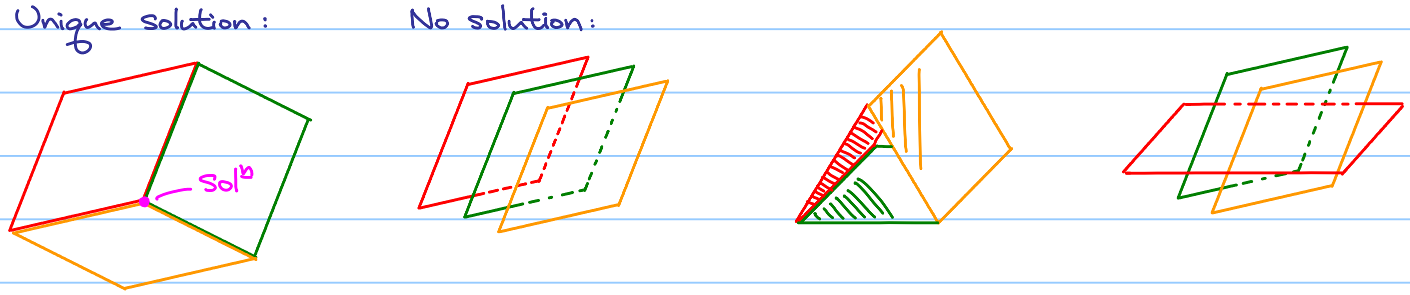
Usually (but NOT necessary) the dimension of solution set is reduced by 1 if we impose an extra equation.

If we insert n affine hyperplanes (n equations) in $\mathbb{R}^n$ (n unknown), the intersection of them (solution set) is usually of 0 dimension (unique solution).

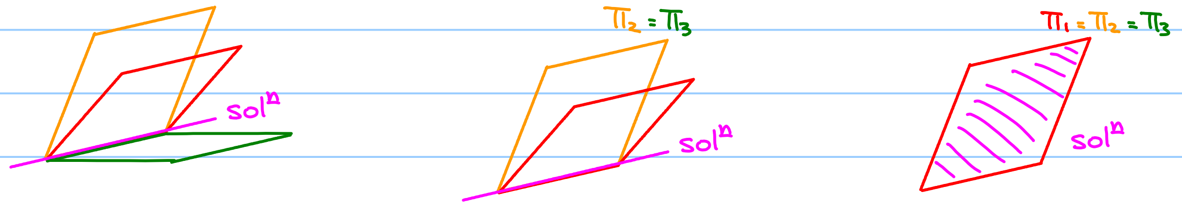
For example, in $\mathbb{R}^3$,

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 & - \Pi_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 & - \Pi_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 & - \Pi_3 \end{cases}$$

Unique solution: No solution:



Infinitely many solutions:



Question. Given $A \in M_n(\mathbb{R})$, $\vec{b} \in M_{n\times 1}(\mathbb{R})$, how to determine if

$A\vec{x} = \vec{b}$ has unique solution $\vec{x} \in M_{n\times 1}(\mathbb{R})$?

Proposition 4.1

Let $A \in M_n(\mathbb{R})$. $A\vec{x} = \vec{b}$ has unique solution if and only if A is invertible.

proof:

$(\Leftarrow)$: If A is invertible,

$$A\vec{x} = \vec{b}$$
$$A^{-1}A\vec{x} = A^{-1}\vec{b}$$
$$\vec{x} = A^{-1}\vec{b} \quad \text{which is the unique solution.}$$

$(\Rightarrow)$: If $A\vec{x} = \vec{b}$ has unique solution, then we write the system as $[A \,\vdots\, \vec{b}]$.

When we perform Gaussian elimination, we have $[\underbrace{E_m E_{m-1} \cdots E_1 A}_{= I} \,\vdots\, \underbrace{E_m E_{m-1} \cdots E_1 \vec{b}}_{\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}]$

$E_m E_{m-1} \cdots E_1 A = I$

i.e. A^{-1} exists and $A^{-1} = E_m E_{m-1} \cdots E_1$ i.e. $x_i = y_i$, for $1 \le i \le n$

Therefore, whether $A\vec{x} = \vec{b}$ has unique solution depends on A only but not $\vec{b}$!

Furthermore, if A is invertible, then $A\vec{x}=\vec{b}$ has unique solution $\vec{x}=A^{-1}\vec{b}$.

Let $\vec{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\vec{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ and recall that $A^{-1}=\frac{1}{\det A}\operatorname{adj}(A)$

$$=\frac{1}{\det A}\begin{bmatrix} M_{11} & -M_{21} & \cdots & (-1)^{n+1}M_{n1} \\ -M_{12} & M_{22} & \cdots & (-1)^{n+2}M_{n2} \\ \vdots & \vdots & & \vdots \\ (-1)^{1+n}M_{1n} & (-1)^{2+n}M_{2n} & \cdots & M_{nn} \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\frac{1}{\det A}\begin{bmatrix} M_{11} & -M_{21} & \cdots & (-1)^{n+1}M_{n1} \\ -M_{12} & M_{22} & \cdots & (-1)^{n+2}M_{n2} \\ \vdots & \vdots & & \vdots \\ (-1)^{1+n}M_{1n} & (-1)^{2+n}M_{2n} & \cdots & M_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\therefore x_i=\frac{1}{\det A}\left((-1)^{1+i}M_{1i}b_1+(-1)^{2+i}M_{2i}b_2+\cdots+(-1)^{n+i}M_{ni}b_n\right)=\frac{1}{\det A}\Big(\underbrace{\sum_{r=1}^{n}(-1)^{r+i}M_{ri}b_r}_{\text{What is it ?}}\Big)$$

Let A_i be the matrix obtained from A by replacing the i-th column by $\vec{b}$.

$\sum_{r=1}^{n}(-1)^{r+i}M_{ri}b_r=\det A_i$ (Check!)

Proposition 4.2 (Cramer's Rule)

Let $A\in M_n(\mathbb{R})$. If $A\vec{x}=\vec{b}$ has unique solution, then $x_i=\frac{\det A_i}{\det A}$.

Proposition 3.10 + 4.1 : The following are equivalent (TFAE).

1) A is invertible

2) $\det A\neq 0$

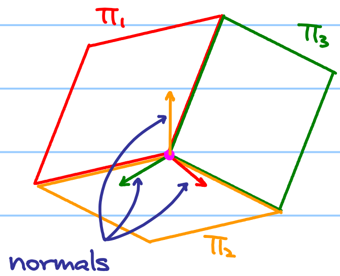
3) $A\vec{x}=\vec{b}$ has unique solution

Remark: If $\det A=0$, it implies $A\vec{x}=\vec{b}$ does NOT have unique solution.

It does NOT mean $A\vec{x}=\vec{b}$ has no solution, it may happen that $A\vec{x}=\vec{b}$ has infinity many solution!

More geometrical point of view:

$A\vec{x} = \vec{b}$ has unique solution $\Leftrightarrow$ $\det(A) \neq 0$



$$(S)\quad \begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 & - \Pi_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 & - \Pi_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 & - \Pi_3 \end{cases}$$

(S) has unique solution $\Leftrightarrow$ $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$

i.e. signed volume of the parallelepiped spanned by normals of Π_1, Π_2 and $\Pi_3 \neq 0$

Example 4.7

Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \\ 2 & 1 & 3 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 8 \\ 7 \\ 13 \end{bmatrix}$ and let (S) be a system of linear equations given by $[A \vdots \vec{b}]$.

Solve (S) if possible.

$\det A = \begin{vmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \\ 2 & 1 & 3 \end{vmatrix} = -2 \neq 0$ $\Rightarrow$ (S) has unique solution

By using Cramer's rule,

let $A_1 = \begin{bmatrix} 8 & 2 & 1 \\ 7 & 3 & 0 \\ 13 & 1 & 3 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 8 & 1 \\ 1 & 7 & 0 \\ 2 & 13 & 3 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 2 & 8 \\ 1 & 3 & 7 \\ 2 & 1 & 13 \end{bmatrix}$

$\det A_1 = -2$, $\det A_2 = -4$, $\det A_3 = -6$

$\therefore$ $x = \dfrac{\det A_1}{\det A} = 1$, $y = \dfrac{\det A_2}{\det A} = 2$, $z = \dfrac{\det A_3}{\det A} = 3$.

Remark: The solution can be found by using Gaussian elimination as well.

Example 4.8

Let $a, b \in \mathbb{R}$, $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 0 & a & 2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 2 \\ 0 \\ b \end{bmatrix}$ and let (S) be a system of linear equations given by $[A \vdots \vec{b}]$

a) Show that (S) has unique solution if and only if $a \neq 4$.

Find the unique solution in terms of a and b in this case.

b) When $a=4$, what is the value of b such that S has solution?

Solve (S) in this case.

a) $\det A = 4 - a$

(S) has unique solution $\Leftrightarrow \det A \neq 0$

$\Leftrightarrow a \neq 4$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 1 & 3 & 2 & 0 \\ 0 & a & 2 & b \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 2 & 1 & -2 \\ 0 & a & 2 & b \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{2} & -1 \\ 0 & a & 2 & b \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & \frac{4-a}{2} & a+b \end{array}\right]$$

$$\therefore \begin{cases} x + y + z = 2 & \Rightarrow x = 2 - x_2 - x_3 = \frac{4-2a-b}{4-a} \\ y + \frac{1}{2} z = 1 & \Rightarrow y = 1 - \frac{1}{2} x_3 = \frac{4-2a-b}{4-a} \\ \frac{4-a}{2} z = b + a & \Rightarrow z = \frac{2(a+b)}{4-a} \end{cases}$$

Note : $4 - a \neq 0$

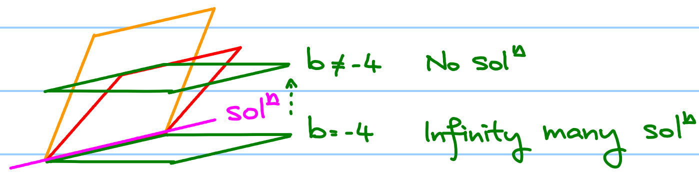
b) When $a = 4$,

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 1 & 3 & 2 & 0 \\ 0 & 4 & 2 & b \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & 4+b \end{array}\right]$$

The last equation is $0x + 0y + 0z = 4 + b$ which is consistent if $4 + b = 0$, i.e. $b = -4$.

When $b = -4$, we have

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{2} & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & 1 \\ 0 & 1 & \frac{1}{2} & 0 \end{array}\right]$$



Let $z = t \in \mathbb{R}$, $y = -\frac{z}{2} = -\frac{t}{2}$, $x = 1 - \frac{x_3}{2} = 1 - \frac{t}{2}$

$(x, y, z) = (1, 0, 0) + t(-\frac{1}{2}, -\frac{1}{2}, 1)$

Let $A\vec{x}=\vec{b}$ be a system of linear equations, where $A\in M_{m\times n}(\mathbb{R})$, $\vec{x}\in M_{n\times 1}(\mathbb{R})$, $\vec{b}\in M_{m\times 1}(\mathbb{R})$.

If $\vec{b}=\vec{0}$, then the system of linear equations is said to be homogeneous.

Note that $\vec{x}=\vec{0}$ is always a solution to $A\vec{x}=\vec{0}$, which is said to be the trivial solution.

(All affine hyperplanes are containing the origin.)

Example 49

Let $A=\begin{bmatrix}1-\lambda & 2 & -1\\ 0 & 1+\lambda & 1\\ 1 & 1 & 4-\lambda\end{bmatrix}$, where $\lambda\in\mathbb{R}$.

Find the value(s) of λ such that $A\vec{x}=\vec{0}$ has non-trivial solution.

Note: $A\vec{x}=\vec{0}$ always has a solution.

$\therefore$ $A\vec{x}=\vec{0}$ has non-trivial solution $\Leftrightarrow$ $\det A=0$

$$\begin{vmatrix}1-\lambda & 2 & -1\\ 0 & 1+\lambda & 1\\ 1 & 1 & 4-\lambda\end{vmatrix}=0$$

$$(1-\lambda)\begin{vmatrix}1+\lambda & 1\\ 1 & 4-\lambda\end{vmatrix}+(1)\begin{vmatrix}2 & -1\\ 1+\lambda & 1\end{vmatrix}=0$$

$$\lambda^3-4\lambda^2+\lambda+6=0$$

$$(\lambda+1)(\lambda-2)(\lambda-3)=0$$

$$\lambda=-1,\ 2 \text{ or } 3$$

§ 5 Linear Independence and Bases

Definition 5.1

Let $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k$ be vectors in $\mathbb{R}^n$.

If $\vec{u} = \sum_{r=1}^{k} c_r\vec{v}_r = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ for some $c_1, \cdots, c_k \in \mathbb{R}$,

then $\vec{u}$ is said to be a linear combination of $\vec{v}_1, \cdots, \vec{v}_k$.

$\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ = the set of all linear combinations of $\vec{v}_1, \cdots, \vec{v}_k$

$= \{\vec{u} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \in \mathbb{R}^n : c_1, \cdots, c_k \in \mathbb{R}\}$.

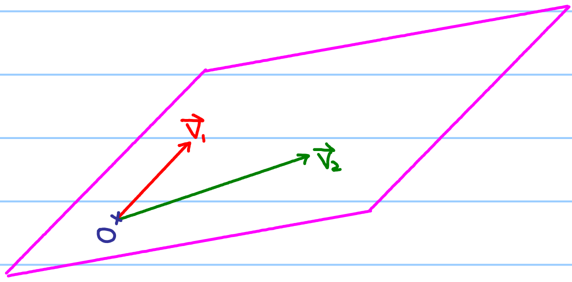
Example 5.1

Let $\vec{v}_1 = (1,0,1)$, $\vec{v}_2 = (1,1,0)$, $\vec{v}_3 = (3,1,2) \in \mathbb{R}^3$

$\text{span}(\{\vec{v}_1, \vec{v}_2\}) = \{\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$

= plane spanned by $\vec{v}_1$ and $\vec{v}_2$.

(Ex: Show that the equation of the plane is $x-y-z=0$.)



How about adding one more vector $\vec{v}_3$?

Note that $\vec{v}_3 = (3,1,2) = 2\vec{v}_1 + 1\vec{v}_2$, i.e. $\vec{v}_3$ itself is a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

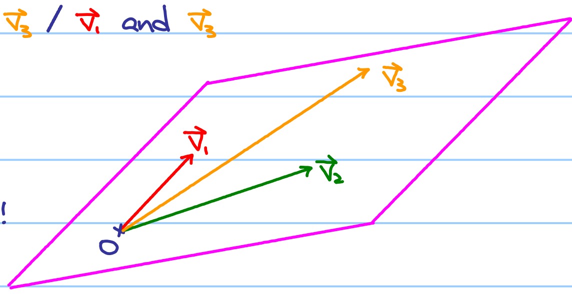
$\text{span}(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}) = \{\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 : c_1, c_2, c_3 \in \mathbb{R}\}$

$$\left(\begin{array}{c|c|c} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3(2\vec{v}_1 + 1\vec{v}_2) & = c_1(-\tfrac{1}{2}\vec{v}_2 + \tfrac{1}{2}\vec{v}_3) + c_2\vec{v}_2 + c_3\vec{v}_3 & = c_1\vec{v}_1 + c_2(\vec{v}_3 - 2\vec{v}_1) + c_3\vec{v}_3 \\ = (c_1 + 2c_3)\vec{v}_1 + (c_2 + c_3)\vec{v}_2 & = (-\tfrac{c_1}{2} + c_2)\vec{v}_2 + (\tfrac{c_1}{2} + c_3)\vec{v}_3 & = (c_1 - 2c_2)\vec{v}_1 + (c_2 + c_3)\vec{v}_3 \end{array} \right)$$

= plane spanned by $\vec{v}_1$ and $\vec{v}_2$ / $\vec{v}_2$ and $\vec{v}_3$ / $\vec{v}_1$ and $\vec{v}_3$

One vector is redundant (Which one ?).

This gives a motivation of the following definition!



Definition 5.2

Let $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k \in \mathbb{R}^n$.

$\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ is said to be linearly dependent if

there exist $c_1, c_2, \cdots, c_k \in \mathbb{R}$, but not all zero, such that

$\sum_{r=1}^{k} c_r\vec{v}_r = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$.

$\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ is said to be linearly independent if

when $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$, we must have $c_1 = c_2 = \cdots = c_k = 0$.

What is the meaning of the above definition?

Suppose there exist $c_1, c_2, \cdots, c_k$ with some $c_j \neq 0$ such that $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$.

$$\vec{v}_j = -\frac{c_1}{c_j}\vec{v}_1 - \cdots - \frac{c_{j-1}}{c_j}\vec{v}_{j-1} - \frac{c_{j+1}}{c_j}\vec{v}_{j+1} - \cdots - \frac{c_k}{c_j}\vec{v}_k$$

i.e. $\vec{v}_j$ is a linear combination of the other vectors!

Therefore, for a linearly independent set of vectors $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$, each vector cannot be expressed as a linear combination of the others.

Let $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k \in \mathbb{R}^n$.

How to determine if $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ is a linearly independent set?

Suppose that $\vec{v}_r = \begin{bmatrix} a_{1r} \\ a_{2r} \\ \vdots \\ a_{nr} \end{bmatrix}$

Find $c_1, c_2, \cdots, c_k$ such that $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$

$$c_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + c_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + c_k\begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

j-th column vector is $\vec{v}_j \in \mathbb{R}^n$

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1k} & 0 \\ a_{21} & a_{22} & \cdots & a_{2k} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & 0 \end{array}\right]$$

In $\mathbb{R}^n$, given k vectors, it associates a system of linear equations with n linear equations, k unknowns.

Example 5.2

Let $\vec{v}_1 = (1,0,1)$, $\vec{v}_2 = (1,1,0)$, $\vec{v}_3 = (3,1,2) \in \mathbb{R}^3$

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\therefore (c_1, c_2, c_3) = (-2t, -t, t)$, $t \in \mathbb{R}$.

i.e. $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$ has non-trivial solution, i.e. $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a linearly dependent set.

In particular, take $t=1$, $(c_1, c_2, c_3) = (-2, -1, 1)$ and we have $-2\vec{v}_1 - \vec{v}_2 + \vec{v}_3 = \vec{0}$.

Example 5.3

Let $\vec{v}_1 = (1,0,2,1)$, $\vec{v}_2 = (2,2,3,2)$, $\vec{v}_3 = (0,2,-1,0)$ and $\vec{v}_4 = (1,2,3,4) \in \mathbb{R}^4$

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 2 & 2 & 0 \\ 2 & 3 & -1 & 3 & 0 \\ 1 & 2 & 0 & 4 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$\therefore (c_1, c_2, c_3, c_4) = (2t, -t, t, 0)$, $t \in \mathbb{R}$

i.e. there exists $(c_1, c_2, c_3, c_4) \neq (0,0,0,0)$ such that $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + c_4\vec{v}_4 = \vec{0}$.

Therefore, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is a linearly dependent set.

In particular, take $t=1$, we have $2\vec{v}_1 - \vec{v}_2 + \vec{v}_3 + 0\vec{v}_4 = \vec{0}$

$$\vec{v}_3 = -2\vec{v}_1 + \vec{v}_2$$

Think:

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 2 & 2 & 0 \\ 2 & 3 & -1 & 3 & 0 \\ 1 & 2 & 0 & 4 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 2 \\ 2 & 3 & 3 & -1 \\ 1 & 2 & 4 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$ which gives $c_1 = -2$, $c_2 = 1$ and $c_4 = 0$

How about removing $\vec{v}_3$? Is $\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ a linear independent set?

If $c_1\vec{v}_1 + c_2\vec{v}_2 + c_4\vec{v}_4 = \vec{0}$, then

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 2 & 3 & 3 & 0 \\ 1 & 2 & 4 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad (c_1, c_2, c_4) = (0,0,0)$$

After transforming to reduced row echelon form,

- the vectors ($\vec{v}_1, \vec{v}_2, \vec{v}_4$ in the example) corresponding to columns with leading 1's form a linearly independent set;
- the other vectors ($\vec{v}_3$ in the example) are "redundant".

Example 5.4

$$\begin{array}{c} \left[\begin{array}{ccccc} 1 & 2 & 1 & 0 & 2 \\ 2 & 4 & 3 & 1 & 0 \\ 3 & 6 & 4 & 1 & 5 \end{array}\right] \\ \begin{array}{ccccc} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \end{array} \end{array} \to \cdots \to \left[\begin{array}{ccccc} 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$\therefore \vec{v}_2 = 2\vec{v}_1$, $\vec{v}_4 = -\vec{v}_1 + \vec{v}_3$ and $\{\vec{v}_1, \vec{v}_3, \vec{v}_5\}$ is a linear independent set of vectors in $\mathbb{R}^3$.

Proposition 5.1

Let $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ be a set of k vectors in $\mathbb{R}^n$.

If $k > n$, then the given set must be linearly dependent.

That means if $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, then $k \le n$.

proof.

Find $c_1, c_2, \cdots, c_k$ such that $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \longrightarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1k} & 0 \\ a_{21} & a_{22} & \cdots & a_{2k} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & 0 \end{array}\right]$$

unknowns $= k > n =$ # linear equations,

so it must have non-trivial solution.

j-th column vector is $\vec{v}_j \in \mathbb{R}^n$

Definition 5.3

Let $k \le n$ and let $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ be a linearly independent set of vectors in $\mathbb{R}^n$.

$\text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$ is said to be a k-dimensional subspace in $\mathbb{R}^n$.

In particular, if $k = n$, we have:

Proposition 5.2

Given a set of n vectors in $\mathbb{R}^n$ $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$.

The given set is linearly independent if and only if $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \ne 0$

j-th column vector is $\vec{v}_j$

proof.

The given set is linearly independent

$\Leftrightarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right]$ has trivial solution as the unique solution

$\Leftrightarrow \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \ne 0$

Example 5.2 (Cont.)

Let $\vec{v}_1 = (1, 0, 1)$, $\vec{v}_2 = (1, 1, 0)$, $\vec{v}_3 = (3, 1, 2) \in \mathbb{R}^3$

$\begin{vmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{vmatrix} = 0$ and so $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a linearly dependent set.

Recall: If $A \in M_n(\mathbb{R})$, $\det(A) = \det(A^T)$. Therefore,

Signed volume of n-parallelotope spanned by $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$

$$= \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} \leftarrow \text{j-th row vector is } \vec{v}_j \quad = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

j-th column vector is $\vec{v}_j$

The above proposition can be interpreted in a more geometrical way:

$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is linearly independent $\Leftrightarrow$ Signed volume of n-parallelotope spanned by $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is nonzero.

Proposition 3.10 + 4.1 + 5.2: The following are equivalent (TFAE).

Let $A \in M_n(\mathbb{R})$.

1) A is invertible

2) $\det(A) \neq 0$

3) $A\vec{x} = \vec{b}$ has unique solution

4) Column vectors (Row vectors) of A forms a linearly independent set of vectors in $\mathbb{R}^n$

Proposition 5.3

Let $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ be a linearly independent set of vectors in $\mathbb{R}^n$.

Then, every vector can be expressed as a unique linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, it also follows that $\text{span}(\{\vec{v}_1, \dots, \vec{v}_n\}) = \mathbb{R}^n$

proof:

Let $\vec{b} = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$.

Find $r_1, r_2, \dots, r_n$ such that $r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_n\vec{v}_n = \vec{b}$

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right] \quad \text{which has a unique solution since } \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$$

j-th column vector is $\vec{v}_j \in \mathbb{R}^n$

Definition 5.4

If $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, it is said to be a basis of $\mathbb{R}^n$.

An ordered basis of $\mathbb{R}^n$ is a basis of $\mathbb{R}^n$ equipped with a specified order.

Example 5.3

Let $\vec{e}_j = (0, \cdots, 0, \overset{j\text{-th}}{1}, 0, \cdots, 0) \in \mathbb{R}^n$.

Then $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ is an ordered basis of $\mathbb{R}^n$ (since $\det I_n = 1 \neq 0$), which is called the standard ordered basis.

If $\beta = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ is an ordered basis of $\mathbb{R}^n$, then every vector $b \in \mathbb{R}^n$ can be expressed as $r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_n\vec{v}_n$ uniquely. In this case, $r_1, r_2, \cdots, r_n$ are said to be coordinates of $\vec{b}$ with respect to β, which is denoted by $\vec{b} = (r_1, r_2, \cdots, r_n)_\beta$ or $[\vec{b}]_\beta = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}_\beta$.

When we simply write $\vec{b} = (b_1, b_2, \cdots, b_n)$ or $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, it means the coordinates of $\vec{b}$ with respect to the standard ordered basis.

Example 5.4

Let $\vec{v}_1 = (1, 2, 1)$, $\vec{v}_2 = (0, 1, 4)$, $\vec{v}_3 = (0, 2, 1) \in \mathbb{R}^3$.

Since $\begin{vmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \\ 1 & 4 & 1 \end{vmatrix} = -7 \neq 0$, $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an ordered basis of $\mathbb{R}^3$.

Let $\vec{b} = (3, 6, -4) \in \mathbb{R}^3$, we are going to find r_1, r_2, r_3 such that

$$r_1\vec{v}_1 + r_2\vec{v}_2 + r_3\vec{v}_3 = \vec{b} \rightsquigarrow \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \\ 1 & 4 & 1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ -4 \end{bmatrix} \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 2 & 1 & 2 & 6 \\ 1 & 4 & 1 & -4 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$\therefore \vec{b} = (3, -2, 1)_\beta$

Given a basis of $\mathbb{R}^n$ and a vector in $\mathbb{R}^n$, we have to solve a system of linear equations if we would like to express the given vector as a linear combination of vectors of the basis (which is complicated).

Is there any better choice of basis?

Definition 5.5

A subset $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$ is said to be orthogonal if each pair of $\vec{v}_i$ and $\vec{v}_j$ are orthogonal, i.e. $\vec{v}_i \cdot \vec{v}_j = 0$ for all $i \neq j$.

Furthermore, an orthogonal subset $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$ is orthonormal if $|\vec{v}_i| = 1$ for all i.

Proposition 5.4

If $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then it must be linearly independent.

In particular, if $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then it must be linearly independent, and so it is a (an orthogonal) basis of $\mathbb{R}^n$.

proof:

Let $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$. Then,

$$(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) \cdot \vec{v}_j = \vec{0} \cdot \vec{v}_j$$

$$c_j |\vec{v}_j|^2 = 0$$

$$c_j = 0 \qquad (\because |\vec{v}_j| > 0)$$

Now, suppose that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is an orthogonal basis of $\mathbb{R}^n$ and let $\vec{b} \in \mathbb{R}^n$.

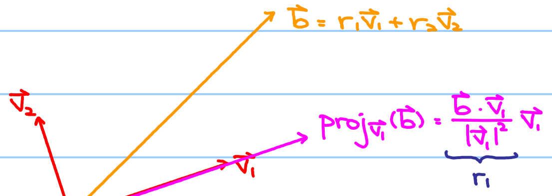
Then there exist unique $r_1, r_2, \dots, r_n \in \mathbb{R}$ such that $r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_n\vec{v}_n = \vec{b}$.

How to find $r_1, r_2, \dots, r_n$?

$$(r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_n\vec{v}_n) \cdot \vec{v}_j = \vec{b} \cdot \vec{v}_j$$

$$r_j \, \vec{v}_j \cdot \vec{v}_j = \vec{b} \cdot \vec{v}_j$$

$$r_j = \frac{\vec{b} \cdot \vec{v}_j}{|\vec{v}_j|^2}$$



$(= \vec{b} \cdot \vec{v}_j$ if $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is an orthonormal basis$)$

Example 5.4

Let $\vec{v}_1 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $\vec{v}_2 = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$, $\vec{v}_3 = (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}) \in \mathbb{R}^3$.

Check. $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ forms an orthonormal ordered basis of $\mathbb{R}^3$, i.e. $\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$

Let $\vec{b} = (2, 1, 3) \in \mathbb{R}^3$, then $\vec{b} = r_1\vec{v}_1 + r_2\vec{v}_2 + r_3\vec{v}_3$

where $r_1 = \vec{b} \cdot \vec{v}_1 = \frac{6}{\sqrt{3}}$, $r_2 = \vec{b} \cdot \vec{v}_2 = \frac{1}{\sqrt{2}}$, $r_3 = \vec{b} \cdot \vec{v}_3 = -\frac{3}{\sqrt{6}}$

Remark:

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$

j-th column vector is $\vec{v}_j \in \mathbb{R}^n$

Note that $[A^TA]_{ij} = \vec{v}_i^T\vec{v}_j$ (or $\vec{v}_i\ \vec{v}_j$)

$\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ is an orthonormal basis $\Leftrightarrow$ $\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j. \end{cases}$ $\Leftrightarrow$ $A^TA = I$ i.e. A is orthogonal.

Gram-Schmidt Process:

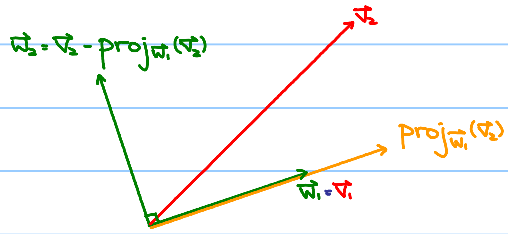
Let $\{\vec{v}_1, \vec{v}_2\}$ be a basis of $\mathbb{R}^2$.

1) Let $\vec{w}_1 = \vec{v}_1$.

2) Let $\vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{w}_1}(\vec{v}_2) = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{|\vec{w}_1|^2}\vec{w}_1$

Then $\{\vec{w}_1, \vec{w}_2\}$ is an orthogonal basis of $\mathbb{R}^2$.

Furthermore, $\{\hat{w}_1, \hat{w}_2\}$ is an orthonormal basis of $\mathbb{R}^2$.



The above method for producing an orthogonal / orthonormal basis from a basis is called the Gram-Schmidt process. The general statement is:

If $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ be a basis of $\mathbb{R}^n$,

let $\vec{w}_1 = \vec{v}_1$ and for $k = 2, \cdots, n$, let $\vec{w}_k = \vec{v}_k - \sum_{j=1}^{k-1} \text{proj}_{\vec{w}_j}(\vec{v}_k) = \vec{v}_k - \sum_{j=1}^{k-1} \frac{\vec{v}_k \cdot \vec{w}_j}{|\vec{w}_j|^2}\vec{w}_j$

then $\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_n\}$ is an orthogonal basis of $\mathbb{R}^n$ and $\{w_1, w_2, \cdots, w_n\}$ is an orthonormal basis of $\mathbb{R}^n$.

Exercise 5.1

Let $\vec{v}_1 = (1,2,1)$, $\vec{v}_2 = (0,1,4)$, $\vec{v}_3 = (0,2,1) \in \mathbb{R}^3$.

Then $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an ordered basis of $\mathbb{R}^3$ (see example 5.4).

By using the Gram-Schmidt process, construct an ordered orthonormal basis from β.

Change of Coordinates:

Let $\beta = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ and $\gamma = \{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_n\}$ be two ordered bases of $\mathbb{R}^n$.

Suppose that $\vec{b} \in \mathbb{R}^n$, $[\vec{b}]_\beta = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}_\beta$ and $[\vec{b}]_\gamma = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}_\gamma$ are coordinates of $\vec{b}$ with respect to β and γ.

What is the relation between $[\vec{b}]_\beta$ and $[\vec{b}]_\gamma$?

Firstly, note that β is an ordered basis, so we have

$$\vec{v}_j = a_{1j}\vec{w}_1 + a_{2j}\vec{w}_2 + \cdots + a_{nj}\vec{w}_n = \sum_{k=1}^{n} a_{kj}\vec{w}_k \text{ for some } a_{kj} \in \mathbb{R}, \text{ i.e. } [\vec{v}_j]_\gamma = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}_\gamma$$

then $\vec{b} = r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_n\vec{v}_n = \sum_{j=1}^{n} r_j\vec{v}_j = \sum_{j=1}^{n} r_j\left(\sum_{k=1}^{n} a_{kj}\vec{w}_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} r_j a_{kj}\right)\vec{w}_k$

therefore $s_k = \sum_{j=1}^{n} r_j a_{kj} = r_1 a_{k1} + r_2 a_{k2} + \cdots + r_n a_{kn}$, i.e. $\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}_\gamma = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}_\beta$

$[\vec{b}]_\gamma = A[\vec{b}]_\beta$

$[\vec{b}]_\beta = A^{-1}[\vec{b}]_\gamma$ where $A \in M_n(\mathbb{R})$ and the j-th column of A is $[\vec{v}_j]_\gamma$.

Example 5.5

Let $\vec{v}_1 = (1,2,1)$, $\vec{v}_2 = (0,1,4)$, $\vec{v}_3 = (0,2,1) \in \mathbb{R}^3$ and

let $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ be an ordered basis of $\mathbb{R}^3$.

Suppose that $\vec{b} = (3,7,0) \in \mathbb{R}^3$. Find $[\vec{b}]_\beta$.

By putting γ to be the standard ordered basis,

$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \\ 1 & 4 & 1 \end{bmatrix}$ and $[\vec{b}]_\beta = A^{-1}[\vec{b}]_\gamma = \frac{1}{7}\begin{bmatrix} 7 & 0 & 0 \\ 0 & -1 & 2 \\ -7 & 4 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \\ 0 \end{bmatrix}_\gamma = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}_\beta$ i.e. $\vec{b} = 3\vec{v}_1 - \vec{v}_2 + \vec{v}_3$

§6 Linear Transformation

Definition 6.1

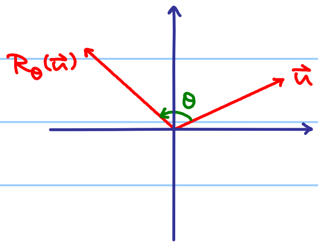
A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $T: \mathbb{R}^n \to \mathbb{R}^m$ which satisfies

1) $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v} \in \mathbb{R}^n$

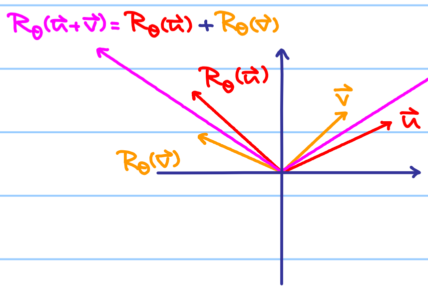
2) $T(c\vec{u}) = cT(\vec{u})$ for all $c \in \mathbb{R}$, $\vec{u} \in \mathbb{R}^n$

Example 6.1

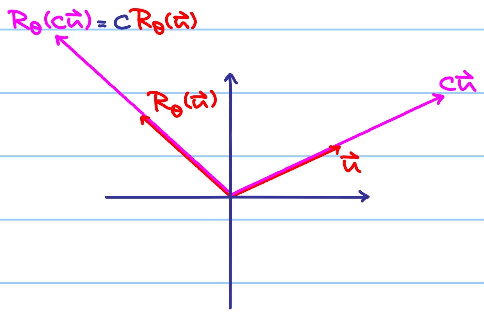
Let $R_\theta: \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation defined by rotation about the origin by θ in anticlockwise direction.



$R_\theta(\vec{u}+\vec{v}) = R_\theta(\vec{u}) + R_\theta(\vec{v})$



$R_\theta(c\vec{u}) = cR_\theta(\vec{u})$



Therefore, R_θ is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$

Question: Let $\theta = \frac{\pi}{3}$, given $\vec{u} = (3,2)$, $R_\theta(\vec{u}) = ?$

Question 1: How do we obtain linear transformations?

Proposition 6.1

Given $A \in M_{m\times n}(\mathbb{R})$, a linear transfromation $L_A: \mathbb{R}^n \to \mathbb{R}^m$ can be associated which is defined by $L_A(\vec{u}) = A\vec{u}$.

proof:

Let $\vec{u} \in \mathbb{R}^n$ which can be regarded as an element of $M_{n\times 1}(\mathbb{R})$.

Then $L_A(\vec{u}) = A\vec{u} \in M_{m\times 1}(\mathbb{R})$ which can be regarded as an element of $\mathbb{R}^m$.

$\therefore L_A$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$

Also $L_A(\vec{u}+\vec{v}) = A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v} = L_A(\vec{u}) + L_A(\vec{v})$ for all $\vec{u}, \vec{v} \in \mathbb{R}^n$

$L_A(c\vec{u}) = A(c\vec{u}) = c(A\vec{u}) = cL_A(\vec{u})$ for all $\vec{u} \in \mathbb{R}^n$, $c \in \mathbb{R}$.

$\therefore L_A$ is a linear transformation

Question 2: Why are linear transformations interesting?

Let $\beta = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ be an ordered basis of $\mathbb{R}^n$ and $\gamma = \{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_m\}$ be an ordered basis of $\mathbb{R}^m$.

Suppose that $T(\vec{v}_1), T(\vec{v}_2), \cdots, T(\vec{v}_n)$ are known.

$$T(\vec{v}_j) = \sum_{r=1}^{m} a_{rj}\vec{w}_r = a_{1j}\vec{w}_1 + a_{2j}\vec{w}_2 + \cdots + a_{mj}\vec{w}_m \text{, i.e. } [T(\vec{v}_j)]_\gamma = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}_\gamma \text{, for } 1 \le j \le n.$$

Let $\vec{u} = \sum_{i=1}^{n} u_i\vec{v}_i = u_1\vec{v}_1 + u_2\vec{v}_2 + \cdots + u_n\vec{v}_n$, i.e. $[\vec{u}]_\beta = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}_\beta$.

Then
$$\begin{aligned}
T(\vec{u}) &= T(u_1\vec{v}_1 + u_2\vec{v}_2 + \cdots + u_n\vec{v}_n) \\
&= u_1T(\vec{v}_1) + u_2T(\vec{v}_2) + \cdots + u_nT(\vec{v}_n) \qquad (\because T \text{ is a linear transformation!})
\end{aligned}$$

$$\begin{aligned}
[T(\vec{u})]_\gamma &= u_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}_\gamma + u_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}_\gamma + \cdots + u_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}_\gamma \\
&= \begin{bmatrix} u_1a_{11} + u_2a_{12} + \cdots + u_na_{1n} \\ u_1a_{21} + u_2a_{22} + \cdots + u_na_{2n} \\ \vdots \\ u_1a_{21} + u_2a_{22} + \cdots + u_na_{2n} \end{bmatrix}_\gamma \\
&= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}_\beta \\
&= A[\vec{u}]_\beta \text{ where } A \in M_{m\times n}(\mathbb{R}) \text{ and the } j\text{-th column vector is } [T(\vec{v}_j)]_\gamma
\end{aligned}$$

Understand T completely if we know $T(\vec{v}_1), T(\vec{v}_2), \cdots, T(\vec{v}_n)$!

Furthermore, for each linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, $T = L_A$ for some $A \in M_{m\times n}(\mathbb{R})$.

A is called the matrix representation of T.

Sometimes we write $[T]_\beta^\gamma$ instead of A to emphasize that the matrix representation depends on the choice of β and γ.

Example 6.2

Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation such that

By taking $\beta = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, $\gamma = \{\vec{e}_1, \vec{e}_2\}$.

$T(\vec{e}_1) = 2\vec{e}_1 + 3\vec{e}_2$, $T(\vec{e}_2) = 3\vec{e}_1 - \vec{e}_2$, $T(\vec{e}_3) = \vec{e}_1 + 2\vec{e}_2$

Matrix representation of T. $[T]_\beta^\gamma = \begin{bmatrix} 2 & 3 & 1 \\ 3 & -1 & 2 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$

$$T(\vec{e}_1 + 2\vec{e}_2 + 3\vec{e}_3) = T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 2 & 3 & 1 \\ 3 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 1 \end{bmatrix}$$

Example 6.3

Suppose that $\beta = \{\vec{v}_1, \vec{v}_2\}$ and $\gamma = \{\vec{w}_1, \vec{w}_2\}$ are ordered bases of $\mathbb{R}^2$.

If $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is a linear transformation such that $T(\vec{v}_1) = 2\vec{w}_1 + 4\vec{w}_2$ and $T(\vec{v}_2) = 3\vec{w}_1 + 1\vec{w}_2$,

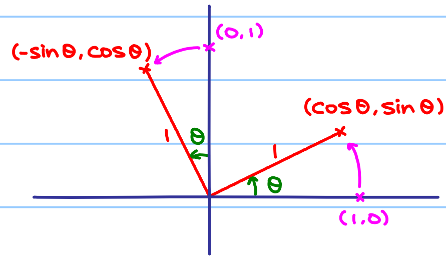
then the matrix representation $[T]_\beta^\gamma = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$

$$T(\vec{v}_1 + 2\vec{v}_2) = T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}_\beta\right) = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}_\gamma = 8\vec{w}_1 + 6\vec{w}_2.$$

Example 6.4

Let $R_\theta: \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation defined by rotation about the origin by θ in anticlockwise direction.

Note that $R_\theta(\vec{e}_1) = R_\theta\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ $\quad R_\theta(\vec{e}_2) = R_\theta\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$



Matrix representation of R_θ : $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

If there is no confusion, we denote the above matrix by R_θ again.

Let $\theta = \frac{\pi}{3}$, given $\vec{u} = (3,2)$, $R_\theta(\vec{u}) = \begin{bmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} - \frac{\sqrt{3}}{4} \\ \frac{3\sqrt{3}}{2} + 1 \end{bmatrix}$

Furthermore, given $\vec{u} \in \mathbb{R}^2$,

$(R_\alpha \circ R_\beta)(\vec{u}) = R_\alpha(R_\beta(\vec{u}))$ which rotates $\vec{u}$ in anticlockwise direction by α and then β, which equals to rotate $\vec{u}$ in anticlockwise direction by $\alpha + \beta$.

Therefore
$$\begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -(\sin\alpha\cos\beta + \cos\alpha\sin\beta) \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{bmatrix}$$

and hence obtain the compound angle formula:

$$\begin{cases} \cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \end{cases}$$

Exercise 6.1

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

Show that $T(\vec{0}) = \vec{0}$.

Direct consequence: Translation is not a linear transformation.

Exercise 6.2

1) Show that the following transformations on $\mathbb{R}^2$ are linear:

(a) scaling about the origin;

(b) reflection along any straight line passing through the origin;

(c) projection on any straight line passing through the origin.

2) For each of the above linear transformation, find the matrix representation (with respect to the standard ordered basis).