MATH 152 Study Notes
Linear Systems

Yecheng Liang

Contents

1. Scalars

A scalar x has

$$x \in \mathbb{R}. \tag{1}$$

In this course,

$$x \in \mathbb{C} \text{ (complex numbers)} \tag{2}$$

is also a scalar.

You can also say a scalar is a 1x1 matrix.

2. Vectors

A vector is 2 or more scalars arranged in a predetermined order.

When written, an arrow is placed above the variable to indicate that it is a vector.

$$x \text{ is a normal variable,} \tag{3.1}$$

$$\vec{x} \text{ is a vector.} \tag{3.2}$$

In printed media, vectors are often written in boldface.

$$\boldsymbol{x} \text{ is a vector.} \tag{4}$$

2.1. Vector Dimensions

The number of scalars in a vector is called the dimension of the vector. For example,

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ is a } \mathbb{R}^2 \tag{5.1}$$

$$\boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \text{ is a } \mathbb{R}^4 \tag{5.2}$$

2.2. Linear Combinations

A linear combination of vectors is the sum of the vectors multiplied by scalars. Each vector is a orthogonal basis vector.

For example, $\boldsymbol{a} \in \mathbb{R}^3$ can be written as

$$\boldsymbol{a} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} \tag{6}$$

where $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are the orthogonal basis vectors for a 3 dimensional space.

2.3. Vector Operations

2.3.1. Vector & Scalar

Yeah just do it.

2.4. Addition & Subtraction

For addition and subtraction, simply add or subtract the corresponding scalars. Commutate, associate and distribute them.
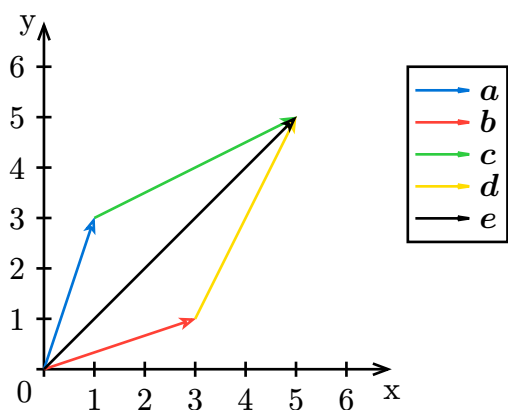
$$a + b = c + d = e \tag{7.1}$$
$$e - a = b \tag{7.2}$$
$$e - c = d \tag{7.3}$$

By subtracting vectors, we can find the vector that connects the two vectors. The length of the connecting vector is the distance between the two vectors.

2.4.1. Dot Product

The two vectors must have the same size/dimension. The dot product of two vectors is the sum of the products of the corresponding scalars.

$$a \cdot b = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{8}$$

Note that the result is a scalar.

Dot product is commutative (even with scalars) and distributive.

2.4.2. Angle Between Vectors

The angle between two vectors can be found using the dot product.

$$\cos(\theta) = \frac{a \cdot b}{\|a\|\|b\|} \tag{9}$$

where θ is the smallest angle between the two vectors.
(This can be proven using Pythagorean theorem.)

If neither of the vectors is the zero vector, and the dot product is 0, then the vectors are orthogonal (perpendicular) to each other.

Given $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, the angle between them is

$$\cos(\theta) = \frac{1 \times 1 + 1 \times -1 + 1 \times 1}{\sqrt{1^2 + 1^2 + 1^2} \times \sqrt{1^2 + -1^2 + 1^2}} = \frac{1}{3} \tag{10.1}$$

$$\theta \approx 1.23\,\text{rad} = 70.53°. \tag{10.2}$$

Note: if the dot product of two vectors $a \cdot b = 0$, then vectors a and b are **perpendicular**.

2.4.3. Vector Length

The length of a vector is the square root of the sum of the squares of the scalars in the vector, which is also the square root of the dot product of the vector with itself. It can be notated as $\|a\|$.

$$\|a\| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}} = \sqrt{a_1^2 + a_2^2 + a_3^2}. \tag{11}$$

For example, $\|\boldsymbol{x} - \boldsymbol{c}\| = 1$ means all points which are 1 unit away from $\boldsymbol{c}$, a circle.

2.4.4. Projection

The projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$ is the vector that is parallel to $\boldsymbol{b}$ and has the same length as the projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$.

$$\text{proj}_{\boldsymbol{b}}\boldsymbol{a} = \text{'shadow' length of } \boldsymbol{a} \text{ on } \boldsymbol{b} \times \text{direction of } \boldsymbol{b} \tag{12.1}$$

$$= (\boldsymbol{a} \cdot \boldsymbol{b})\frac{\boldsymbol{b}}{\|b\|^2} \tag{12.2}$$

$$= \left(\boldsymbol{a} \cdot \hat{\boldsymbol{b}}\right)\hat{\boldsymbol{b}} \tag{12.3}$$

Here's a proof:

The projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$ can be represented as a scalar multipled by $\boldsymbol{b}$, where $\boldsymbol{b} \neq 0$.

$$\text{proj}_{\boldsymbol{b}}\boldsymbol{a} = s\boldsymbol{b} \tag{13}$$

We know that the perpendicular vector, connecting point on $\boldsymbol{b}$ to the end of $\boldsymbol{a}$, can be denoted as $\boldsymbol{a} - s\boldsymbol{b}$.

Hence, we know that

$$(\boldsymbol{a} - s\boldsymbol{b}) \cdot \boldsymbol{b} = 0 \tag{14}$$

$$\boldsymbol{a} \cdot \boldsymbol{b} - s(\boldsymbol{b} \cdot \boldsymbol{b}) = 0 \tag{15}$$

$$s = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{b} \cdot \boldsymbol{b}} \tag{16}$$

Plugging back into Equation 13, we hence get:

$$\text{proj}_{\boldsymbol{b}}\boldsymbol{a} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|b\|^2}\boldsymbol{b} \tag{17}$$

Further, a vector $\boldsymbol{a}$ has $\frac{a}{\|a\|} = \hat{\boldsymbol{a}}$, so

$$\frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|b\|^2}\boldsymbol{b} = \left(\boldsymbol{a} \cdot \hat{\boldsymbol{b}}\right)\hat{\boldsymbol{b}}. \tag{18}$$

2.4.5. Parallelogram

A parallelogram is a quadrilateral with opposite sides parallel. Given two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, the area of the parallelogram formed by them is

$$A = \text{base} \times \text{height} \tag{19.1}$$

$$= \|a\|\|b\| \sin(\theta) \tag{19.2}$$

$$= \|a\|\|b\| \cos\left(\frac{\pi}{2} - \theta\right) \tag{19.3}$$

$$= \boldsymbol{a}_\perp \cdot \boldsymbol{b} \tag{19.4}$$

where $\boldsymbol{a}_\perp$ is the vector perpendicular to $\boldsymbol{a}$.

2.5. Matrices

A matrix is a rectangular array of scalars. It can also be thought of as rows of vectors.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{20}$$

2.5.1. Determinant of 2x2 Matrix

Given a matrix $\boldsymbol{a} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant of a 2x2 matrix is

$$\det \boldsymbol{a} \equiv \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{21}$$

2.5.2. Determinant of 3x3 Matrix

Let a matrix be defined as:

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \tag{22}$$

The minor matrix for an element is the 2x2 matrix obtained from deleting the row and column that element is in.

e.g.

$$\text{minor matrix of } f = \begin{bmatrix} a_1 & a_2 \\ c_1 & c_2 \end{bmatrix} \tag{23}$$

One way to find the determinant is by row expansion, which is taking each of the elements in a row OR column and multiplying it by the determinant of the minor matrix. Take note of the alternating signs. (This is knkown as "Laplace Expansion" for anyone curious).

The signs for computing the sum of each element x its minor is:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix} \tag{24}$$

If you pick row 1:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}. \tag{25}$$

If you pick row 2:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -b_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}. \tag{26}$$

If you pick column 2:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}. \tag{27}$$

Tip: pick any row / column with the **most zeroes**.

The determinant of a 3x3 matrix is

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh \tag{28.1}$$

$$\text{(try find a pattern here!)} \tag{28.2}$$

This can be visualised as putting copying the first two columns onto the right, then for each 3 number diagonal in the shape of a backslash, take the product of the three numbers and sum them, and for each 3 number diagonal in the shape, subtract the each product from the previous sum.

For a parallelogram formed by two vectors a and b, the area is the absolute value of the determinant of the matrix formed by the vectors:

$$A_{\text{parallelogram}} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}. \tag{29}$$

For example, for $a = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $b = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, the area of the parallelogram is

$$A_{\text{parallelogram}} = \begin{vmatrix} 1 & 3 \\ -2 & 5 \end{vmatrix} = 11. \tag{30}$$

2.5.3. Linear Dependence

2.5.4. Cross Product

The cross product of two vectors is a vector that is orthogonal to the two vectors,
a.k.a. the normal vector of the plane formed by the two vectors,
a.k.a. the vector that is perpendicular to the two vectors.

Given $a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, the cross product is

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{31.1}$$

$$= \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}. \tag{31.2}$$

In a Right Hand (RH) coordinate system, the cross product is the vector that points in the direction of the thumb when the fingers of the right hand curl from $\boldsymbol{a}$ to $\boldsymbol{b}$. In other words, the cross product is the vector that is orthogonal/perpendicular to the two vectors.

$$\boldsymbol{a} \cdot (\boldsymbol{a} \times \boldsymbol{b}) = 0 \tag{32}$$

2.5.5. Volume of Parallelepiped

The volume of a parallelepiped formed by three vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ is the absolute value of the determinant of the matrix formed by the vectors.

$$V_{\text{parallelepiped}} = \text{base area} \times \text{height} \tag{33.1}$$

$$= A_{\text{parallelogram}} \times \cos(\theta)\|c\| \tag{33.2}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \tag{33.3}$$

$$= |\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})|. \tag{33.4}$$

3. Lines, Curves and Planes in Vector Form

3.1. Lines in 2D Space

There is a line L. Take a point on the line, $\boldsymbol{p} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, then take a vector that is parallel to the line, $\boldsymbol{l} = \begin{bmatrix} i \\ j \end{bmatrix}$. The line can be represented as

$$\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{l} \tag{34}$$

where $t \in \mathbb{R}$.

Alternatively,

$$\begin{cases} x = it + x_0 \\ y = jt + y_0 \end{cases} \tag{35}$$

The directional vector $\boldsymbol{l}$ is can be compared to other vectors to determine if they are parallel or perpendicular, or neither.

- If $\boldsymbol{l}_1 \cdot \boldsymbol{l}_2 = 0$, then the two vectors are perpendicular.
- Else, if $\boldsymbol{l}_1 = c\boldsymbol{l}_2$ where c is a scalar constant, then the two vectors are parallel.

Where there is a line, there is a normal vector to the line, $\boldsymbol{n}$. Thus the line can also be represented as

$$\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0 \tag{36.1}$$
$$\boldsymbol{n} \cdot \boldsymbol{x} = \boldsymbol{n} \cdot \boldsymbol{p}. \tag{36.2}$$

If we have a line $x - 2y + 3z = 0$, we know that the LHS is not 0, while the RHS $\boldsymbol{n} \cdot \boldsymbol{p} = 0$, $\boldsymbol{p} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, so the line must pass through the origin. Additionally, the normal vector is $\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$.

3.2. Planes in 3D Space

There is a plane S in $\mathbb{R}^3$. Take a point on the plane, $\boldsymbol{p} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$. If $\boldsymbol{x}$ is a point on the plane, $\boldsymbol{x} - \boldsymbol{p}$ must be perpendicular to the normal vector of the plane from origin, $\boldsymbol{n}$. Thus the plane can be represented as

$$\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0 \quad \text{or} \quad \boldsymbol{n} \cdot \boldsymbol{x} = \boldsymbol{n} \cdot \boldsymbol{p}. \tag{37}$$

Similarly, if we have a plane $x - 2y + 3z = 0$, we know that the LHS is not 0, while the RHS $\boldsymbol{n} \cdot \boldsymbol{p} = 0$, $\boldsymbol{p} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, so the plane must pass through the origin. Additionally, the normal vector is $\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$.

The plane can also be seen as:

$$ax + by + cz = d. \tag{38}$$

where $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \boldsymbol{n}$ and $d = \boldsymbol{n} \cdot \boldsymbol{p}$. We can see that this is an expansion of the $\boldsymbol{n} \cdot \boldsymbol{x} = \boldsymbol{n} \cdot \boldsymbol{p}$ form.

Alternatively, we can use the parametric form:

$$\boldsymbol{x} - \boldsymbol{p} = s\boldsymbol{u} + t\boldsymbol{v} \tag{39}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors on the plane and $s, t \in \mathbb{R}$.

For example, a plane S has parametric form

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}. \tag{40}$$

To find the normal representation, we can do a cross product of the two vectors on the plane to get the normal vector.

$$\boldsymbol{n} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \tag{41.1}$$

$$= \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} \tag{41.2}$$

$$= \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}. \tag{41.3}$$

Another example, given plane S's equation $x + 2y + 3z = 6$, we can find its parametric form by locating three points on the plane, then finding two vectors connecting the dots.

However, a better way is to restrict two of the variables to 0 and the third to 1.

$$x + 2y + 3z = 6 \tag{42.1}$$
$$y = s \tag{42.2}$$
$$z = t \tag{42.3}$$
$$x = 6 - 2y - 3z \tag{42.4}$$
$$= 6 - 2s - 3t. \tag{42.5}$$

$$\boldsymbol{x} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}. \tag{43}$$

3.2.1. Distance from Point to Plane

The distance from a point $P(x_0, y_0, z_0)$ to a plane $ax + by + cz = d$ is

$$d_{\text{distance}} = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \tag{44}$$

3.3. Lines in 3D Space

3D line L can be written as

$$x = p + tl, \tag{45}$$

the same as in 2D space.

The line can also be represented as the intersection of two planes.

$$\begin{cases} n_1 \cdot x = n_1 \cdot p_1 \\ n_2 \cdot x = n_2 \cdot p_2 \end{cases} \tag{46.1}$$

$$l = n_1 \times n_2. \tag{46.2}$$

Example: given a line L with

$$\begin{cases} x + y + z = 3 \\ x - y + 2z = -7 \end{cases} \tag{47}$$

Firstly, the planes' normals are obviously

$$n_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{48.1}$$

$$n_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}. \tag{48.2}$$

So we can have l

$$l = n_1 \times n_2 = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 1 & 1 & 1 \\ 1 & -1 & 2 \end{vmatrix} = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}. \tag{49}$$

Let $z = 0$,

$$\begin{cases} x + y = 3 \\ x - y = -7 \end{cases} \tag{50.1}$$

$$\begin{cases} x = -2 \\ y = 5 \end{cases}. \tag{50.2}$$

Thus, $p = \begin{bmatrix} -2 \\ 5 \\ 0 \end{bmatrix}$ is on L.

$$x = \begin{bmatrix} -2 \\ 5 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}. \tag{51}$$

Inversely, given a line L with

$$x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \tag{52}$$

we can find two planes that intersect at the line. Look for two combinations of $t = x, y, z$ that when plugged in, will remove t from the equation. That will give us two planes.

3.3.1. Projection of a Line onto A Plane

Let $\boldsymbol{a}$ be a line and S be a plane in 3D space.

In case that S is one of the coordinate planes, the projection is simple enough (taught in PHYS 170).

$$\boldsymbol{a}_{\text{proj on xy-plane}} = \begin{bmatrix} a_1 \\ a_2 \\ 0 \end{bmatrix} \tag{53}$$

and so for the other two planes.

In general, the projection in 3D space resembles the projection in 2D space.

$$\boldsymbol{a}_\parallel = \boldsymbol{a} - \boldsymbol{a}_\perp \tag{54.1}$$

$$= \boldsymbol{a}_{\text{proj}} \tag{54.2}$$

4. Systems of Linear Equations

A system of linear equations is a set of equations that can be written in the parametric form

$$\begin{cases} a_{11}x + b_{12}y + c_{13}z = d_1 \\ a_{21}x + b_{22}y + c_{23}z = d_2 \\ a_{31}x + b_{32}y + c_{33}z = d_3 \end{cases} \tag{55}$$

where $a_{ij}, b_i, (i, j \in (1, 2, 3)) \in \mathbb{R}$, which i is for row and j is for column.

Notice that each equation defines a plane in 3D space. The planes can intersect at a point, form a line, or (have 2 or more of them) be parallel.

4.1. Linear Dependence and Independence

Linear Dependence (LD) Non-zero vectors which are parallel to each other, or can be expressed as a scalar multiple of each other, or most generally, $s_1 a_1 + s_2 a_2 + \ldots = 0$ can be true when *not all s's* are 0.

In 2D geometry, the parallelogram formed by two LD vectors would have 0 area.

Linear Independence (LI) Non-zero vectors which satisfy none of the LD requirements.

Linear Combination A vector expressed as a sum of scalar multiples of other vectors.

Basis A set of LI vectors that can be used to express any vector in the space. In an $\mathbb{R}^n$ space, the basis has n vectors. Hence, a set of vectors are definitely LD if there are more than n vectors in the set.

> If a set of vectors are LD, any subset of them *is possible to be* LI.

For example,

$$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \text{ and } \begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix} \text{ are LD} \tag{56.1}$$

$$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ are LI.} \tag{56.2}$$

To check if a set of vectors are LI, we can form a matrix with the vectors as *rows* and find the determinant. If the determinant is not 0, then the vectors are LI.

For example,

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{vmatrix} = 1 \neq 0, \tag{57}$$

so $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ are LI.

For 3 3D vectors, the other way is to check the volume of the parallelepiped formed by the vectors, $|a \times (b \cdot c)|$. If the volume is not 0, then the vectors are LI.

4.2. Solving Systems of Linear Equations

Previously, we solve systems of linear equations by substitution or elimination. (No way I'm explaining elimination here.)

Two equations and two unknowns are good, how about three for each? A hundred for each? This is when matrices come in handy.

Using elimination, we are essentially multiplying the equations by a scalar and adding them together. Recall that this is a linear combination of the equations. If the system is solvable, we must be able to express one equation as a linear combination of the others.

To solve the system, look at the following example:

$$\begin{cases} x_1 + x_2 + x_3 = 4 \\ x_1 + 2x_2 + 3x_3 = 9 \\ 2x_1 + 3x_2 + x_3 = 7 \end{cases} \tag{58.1}$$

$$\begin{cases} (2) = (2) - (1) \\ (3) = (3) - 2(1) \end{cases} \quad \begin{cases} x_1 + x_2 + x_3 = 4 \\ x_2 + 2x_3 = 5 \\ x_2 - x_3 = -1 \end{cases} \tag{58.2}$$

$$\dots \tag{58.3}$$

How about we stop copying the x's and write the coefficients and right-hand-side in a matrix? Put coefficients in each row into rows to form a matrix, and the right-hand-side into an augmented column.

$$\begin{bmatrix} 2 & 1 & 3 & | & 1 \\ 4 & 5 & 7 & | & 1 \\ 2 & -5 & 5 & | & 7 \end{bmatrix} \xrightarrow[\;(3)=(3)-(1)\;]{(2)=(2)-2(1)} \begin{bmatrix} 2 & 1 & 3 & | & 1 \\ 0 & 3 & 1 & | & 5 \\ 0 & -6 & 2 & | & -8 \end{bmatrix} \tag{59.1}$$

$$\xrightarrow[\text{"row echelon form"}]{(3)=(3)+2(2)} \begin{bmatrix} 2 & 1 & 3 & | & 1 \\ 0 & 3 & 1 & | & 5 \\ 0 & 0 & 4 & | & 2 \end{bmatrix} \tag{59.2}$$

$$\xrightarrow{(3)=(3)/4} \begin{bmatrix} 2 & 1 & 3 & | & 1 \\ 0 & 3 & 1 & | & 5 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{bmatrix} \tag{59.3}$$

$$\xrightarrow[x_3 = 1/2]{\text{substitute back}} \begin{cases} x_1 = -1 \\ x_2 = \frac{3}{2} \\ x_3 = \frac{1}{2} \end{cases} \tag{59.4}$$

Voila! This is very similar to the Gaussian elimination we learned probably in high school, but there are reasons why we do it this way.

Also, the matrix with the weird lines is called the augmented matrix.

Row Echelon Form (REF) A matrix where the leading element of each row is to the right of the leading element of the row above it.

To write a matrix in row echelon form, we need to make sure that:

- The all-zero rows are at the bottom.
- There are no identical LHS rows.
- There are increasing number of zeroes from upper to lower rows.

When the last row is all zeros, we cannot give a unique solution to the system. Instead, we assign a parameter, like t, to the last variable and solve for the others.

For example,

$$
\begin{bmatrix} 1 & 1 & 1 & | & 1 \\ 1 & 2 & 2 & | & 2 \\ 2 & 3 & 3 & | & 4 \end{bmatrix} \xrightarrow[(3)=(3)-2(1)]{(2)=(2)-(1)} \begin{bmatrix} 1 & 1 & 1 & | & 1 \\ 0 & 1 & 1 & | & 1 \\ 0 & 1 & 1 & | & 2 \end{bmatrix} \tag{60.1}
$$

$$
\xrightarrow{(3)=(3)-(2)} \begin{bmatrix} 1 & 1 & 1 & | & 1 \\ 0 & 1 & 1 & | & 1 \\ 0 & 0 & 0 & | & 1 \end{bmatrix} \tag{60.2}
$$

$$
0x_1 + 0x_2 + 0x_3 = 1 \text{ (impossible)} \tag{60.3}
$$

$$
\text{No solution.} \tag{60.4}
$$

4.2.1. Checking for Linear Dependence by REF

To check if a set of vectors are LD, we can form a matrix with the vectors as ***columns*** and find the REF.

If there is a row of all zeros, then the vectors are LD.

4.3. Distance from Line to Line

Given two lines L_1 and L_2 in 3D space. The distance between the two lines is the length of the vector that connects the two lines and is perpendicular to both lines.

Thus, a cross product of the two directional vectors of the lines gives the direction of the connecting vector.

For example, given $L_1 : \boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $L_2 : \boldsymbol{x} = s \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, one can first get the direction of the connecting vector by taking the cross product of the two directional vectors. Then find a point on each line and the connecting vector.

Alternatively, we can find two arbitrary points on each line, connect them, and find the projection of the connecting vector onto the direction of the perpendicular line from one of the points.

4.4. Rank and Solution Structure

Rank The number of LI (non-zero) rows in the REF of matrix, including augmentation.

- If $\text{rank}[A] = \text{rank}[A|\boldsymbol{b}]$ and $\text{rank}[A] = n$, then the system has a unique solution.
- If $\text{rank}[A] < \text{rank}[A|\boldsymbol{b}]$, then the system has no solution.
- If $\text{rank}[A] = \text{rank}[A|\boldsymbol{b}] < n$, then the system has infinite solutions.

Homogeneous System A system where the RHS is all zeros. 0 is always a solution.

Reduced Row Echelon Form (RREF) A matrix where the leading element (pivot) of each row is 1 and is to the right of the pivot of the row above it. All elements above a pivot are 0. Any column without a pivot is a free variable.

To form a matrix in RREF, we can start from a REF matrix and divide each row by the pivot to make it 1, then eliminate the elements above the pivot.

Recall that LD vectors can be expressed as a linear combination of each other. So to verify if a set of vectors are LD, we can form a matrix with the vectors as columns and find the RREF. If all columns have pivots, then the vectors are LI.

4.4.1. Application in Resistor Networks

We know Ohm's Law, $V = IR$. (If you don't, you should.)

Independent loop A loop with at least one edge that is not part of any other loop.

Each independent loop implies a variable. Dependent loops are represented by the independent loops.

By Kirchhoff's Laws, the sum of the voltage drops in a loop is 0, and the current entering a node is equal to the current leaving the node. Use these to form a system of linear equations.

This process can be troublesome for large networks, so below is another way:

Loop only method:
1. Identify a number of independent loops that is enough for all unknowns.
   The loops can be arbitrary, current should conform to the direction of the loop.
2. For each loop, write the sum of the currents times the resistance of the edge as 0.
3. Form a system of linear equations.
4. Move the constants to the RHS and form a matrix.
5. Solve the system using RREF.

4.4.2. Application in Chemical Reactions

Consider chemical reaction $x_1 C_6 H_{12} O_6 + x_2 O_2 \Rightarrow x_3 CO_2 + x_4 H_2 O$. Because atoms are conserved, we can form a system of linear equations. Then solve it using RREF.

4.4.3. Application in Traffic Flow

Consider four nodes connected by roads to form a rectangle. Given incoming traffic for two nodes and out-coming traffic for two other nodes, we can form a system of linear equations to find the traffic flow between the nodes.

You may not find a unique solution out right, but take reality into account, e.g. solutions must be positive.

4.5. Matrix Multiplication

Given two matrices A and B where A has m rows and n columns and B has n rows and p columns, the product AB is a matrix with m rows and p columns.

In principle, the number of *columns* (*size of row*) in the first matrix must equal the number of *rows* (*size of column*) in the second matrix for the product to be defined.

For an element in the product matrix C,

$$c_{ij} = \left(i^{\text{th}} \text{ row of } A\right) \cdot \left(j^{\text{th}} \text{ column of } B\right) \tag{61.1}$$

$$= \sum_{k=1}^{n} a_{ik}b_{kj}. \tag{61.2}$$

For example,

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}_{2\times3} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{3\times1} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \end{bmatrix}_{2\times1} \tag{62.1}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{3\times1} \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}_{2\times3} = \text{undefined}, 1 \neq 2. \tag{62.2}$$

Even if $A_{n\times n}$ and $B_{n\times n}$, $AB \neq BA$ in general. Only in the special case of diagonal matrices:

$$A = \begin{bmatrix} a & 0 & \cdots & 0 \\ 0 & b & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{bmatrix} \tag{63.1}$$

$$AB = BA. \tag{63.2}$$

There is a type of matrix called the identity matrix, I, where

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n\times n}. \tag{64}$$

We would explore more wonders of the diagonal, identity and triangular matrices later.

Now, we can write linear systems in matrix multiplication form.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{23}x_3 = b_2 \\ \vdots \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{33}x_3 = b_3 \end{cases} \tag{65.1}$$

$$\text{is equivilant to} \tag{65.2}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{13} \\ a_{21} & a_{22} & \cdots & a_{23} \\ \vdots & \vdots & \vdots & \vdots \\ a_{31} & a_{32} & \cdots & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_3 \end{bmatrix}. \tag{65.3}$$

Such form has a simple notation:

$$A\boldsymbol{x} = \boldsymbol{b}. \tag{66}$$

which A is called the coefficient matrix.

4.5.1. Properties of Matrix Multiplication

- Associative: $A(BC) = (AB)C$.
- Distributive: $A(B + C) = AB + AC$.
- Not commutative: $AB \neq BA$ in general.
- Identity: $IA = AI = A$ where I is the identity matrix.
- Inverse: $AA^{-1} = A^{-1}A = I$ where A^{-1} is the inverse of A.
- Transpose: $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$.

4.5.2. Powers of Matrices

One numb enough to calculate M^{999} by $M \times M \times \ldots$ would be dumb, we have a smarter way.

We just learned that diagonal matrices behave like scalars in multiplication. So, if we can make M a diagonal matrix, then we can easily calculate M^n, even when n is very large.

How? See Section 6.9.

4.6. Linear Transformation

Linear transformation / map / function a rule that assigns one output to each input.

For example, given $\boldsymbol{x} \in \mathbb{R}^n$

$$A\boldsymbol{x} = \boldsymbol{y} \in \mathbb{R}^m \tag{67}$$

indicates that A is a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$:

$$A : \mathbb{R}^n \to \mathbb{R}^m. \tag{68}$$

Linearity A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear iff

$$T(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) = \alpha T(\boldsymbol{x}_1) + \beta T(\boldsymbol{x}_2) \tag{69.1}$$

$$\text{or} \tag{69.2}$$

$$T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2) \text{ and } T(\alpha \boldsymbol{x}) = \alpha T(\boldsymbol{x}). \tag{69.3}$$

Thus, all matrix multiplications are linear.

For example,

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}. \tag{70}$$

In the case above, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ transforms 3, $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ transforms 4:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \left[\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right] \begin{bmatrix} 3 \\ 4 \end{bmatrix} \tag{71.1}$$

$$= 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} -1 \\ 0 \end{bmatrix} \tag{71.2}$$

$$= \begin{bmatrix} -4 \\ 3 \end{bmatrix}. \tag{71.3}$$

Looks like $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ was rotated by 90° in counterclockwise!

In the example above, we also notice that the transforming matrix was written in forms of vector columns.

4.6.1. Column Space of Matrix

Matrix can be expressed as columns space of vectors.

$$\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} = \left[\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \ \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right]. \tag{72}$$

Matrix multiplication would be the same as individual "vector multiplication", provided that such operation is defined.

So, for an $M_{m \times n}$, its column space will be a row of n vectors. Recall that matrix multiplication is only defined when the number of columns of the first matrix is the same as the number of rows of the second matrix!

$$\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 x_1 & \boldsymbol{a}_2 x_2 & \cdots & \boldsymbol{a}_n a_n \end{bmatrix}. \tag{73}$$

Write $\boldsymbol{a}$ as a transformation, $T(e)$, each column of the result becomes an image/transformation of x.

> **Sanity check**: *projection* is not a one-to-one operation, so the matrix you use must have a determinant of 0.
>
> Inverse of a reflection matrix is itself.
>
> $$AA = AA^{-1} = I. \tag{74}$$

4.6.2. Finding Transformations in 2D Space

Say, we are projecting some vector to a vector with $\boldsymbol{u}$ at angle φ to the x axis, Of course, we can use the projection formula, but here we want to find the transformation matrix that does the job for us (and it's easier). Let $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ be vectors on the x and y that projects to $\boldsymbol{u}$.

$$M_{\text{proj}} = [T(\boldsymbol{e}_1) \ T(\boldsymbol{e}_2)] \tag{75.1}$$

$$T(\boldsymbol{e}_1) = \text{Proj}_\varphi(\boldsymbol{e}_1) \tag{75.2}$$

$$= \cos(\varphi)\boldsymbol{u} \tag{75.3}$$

$$= \cos(\varphi) \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} \tag{75.4}$$

$$= \begin{bmatrix} \cos^2(\varphi) \\ \sin(\varphi)\cos(\varphi) \end{bmatrix} \tag{75.5}$$

$$T(\boldsymbol{e}_2) = \mathrm{Proj}_\varphi(\boldsymbol{e}_2) \tag{75.6}$$

$$= \sin(\varphi)\boldsymbol{u} \tag{75.7}$$

$$= \sin(\varphi)\begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} \tag{75.8}$$

$$= \begin{bmatrix} \sin(\varphi)\cos(\varphi) \\ \sin^2(\varphi) \end{bmatrix} \tag{75.9}$$

$$M_{\mathrm{proj}} = \begin{bmatrix} \cos^2(\varphi) & \sin(\varphi)\cos(\varphi) \\ \sin(\varphi)\cos(\varphi) & \sin^2(\varphi) \end{bmatrix}. \tag{75.10}$$

Then, to project any vector $\boldsymbol{a}$ to $\boldsymbol{u}$, we can simply put

$$\mathrm{Proj}_{\boldsymbol{u}}\boldsymbol{a} = M_{\mathrm{proj}}\boldsymbol{a}. \tag{76}$$

To find other transformations, similarly find transformation of each orthogonal vector/basis, then combine them to a column space.

Here is another example: rotation. Let's say we are to find a rotation matrix of φ counter-clockwise.

$$T(\boldsymbol{e}_1) = \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} \tag{77.1}$$

$$T(\boldsymbol{e}_2) = \begin{bmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{bmatrix} \tag{77.2}$$

$$M_{\mathrm{rot}} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}. \tag{77.3}$$

Yet another one, reflection by function $y = mx$.

$$\varphi = \mathrm{acos}\left(\frac{\boldsymbol{x}\cdot\boldsymbol{y}}{\|x\|\|y\|}\right) \tag{78.1}$$

$$\mathrm{Ref}_{\varphi(\boldsymbol{x})} = \boldsymbol{x}_\| - \boldsymbol{x}_\perp \tag{78.2}$$

$$= \boldsymbol{x}_\| - \left(\boldsymbol{x}_\| - \boldsymbol{x}\right) \tag{78.3}$$

$$= 2\boldsymbol{x}_\| - \boldsymbol{x} \tag{78.4}$$

$$= 2\mathrm{Proj}_{\varphi(\boldsymbol{x})} - \boldsymbol{x} \tag{78.5}$$

$$= 2M_{\mathrm{proj}_{\varphi(\boldsymbol{x})}} - I\boldsymbol{x} \tag{78.6}$$

$$= \left(2M_{\mathrm{proj}_\varphi} - I\right)\boldsymbol{x} \tag{78.7}$$

$$M_{\mathrm{ref}_\varphi} = 2M_{\mathrm{proj}_\varphi} - I \tag{78.8}$$

$$= 2\begin{bmatrix} \cos^2(\varphi) & \sin(\varphi)\cos(\varphi) \\ \sin(\varphi)\cos(\varphi) & \sin^2(\varphi) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{78.9}$$

$$= \begin{bmatrix} 2\cos(2\varphi) + 1 & \sin(2\varphi) \\ \sin(2\varphi) & -\cos(2\varphi) + 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{78.10}$$

$$= \begin{bmatrix} \cos(2\varphi) & \sin(2\varphi) \\ \sin(2\varphi) & -\cos(2\varphi) \end{bmatrix}. \tag{78.11}$$

4.6.3. Finding Transformations in 3D Space

Firstly, reflection. Given a plane $P : x + y + z = 0$ and a vector $[x]$, find the reflection of x about the plane.

$$n = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{79.1}$$

$$\hat{n} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{79.2}$$

$$\mathrm{Ref}_{P(\boldsymbol{x})} = \boldsymbol{x}_{\parallel} - \boldsymbol{x}_{\perp} \tag{79.3}$$

$$= (\boldsymbol{x} - \boldsymbol{x}_{\perp}) - \perp \tag{79.4}$$

$$= \boldsymbol{x} - 2\boldsymbol{x}_{\perp} \tag{79.5}$$

$$= (I - 2\mathrm{Proj}_{\hat{n}}(x))\boldsymbol{x} \tag{79.6}$$

$$M_{\mathrm{ref}_P} = I - 2M_{T_{\mathrm{proj}_{\hat{n}}}}(x) \tag{79.7}$$

$$\mathrm{Proj}_{\hat{n}}(\boldsymbol{e}) = (\boldsymbol{e} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}} \tag{79.8}$$

$$= \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{79.9}$$

$$M_{\mathrm{proj}_{\hat{n}}} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{79.10}$$

$$M_{\mathrm{proj}_P} = I - M_{\mathrm{proj}_{\hat{n}}} \tag{79.11}$$

$$= \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \tag{79.12}$$

$$M_{\mathrm{ref}_P} = I - 2M_{\mathrm{proj}_{\hat{n}}} \tag{79.13}$$

$$= \frac{1}{3} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}. \tag{79.14}$$

4.6.4. Checking for Linearity

To ensure a one-to-n-one relationship, a transformation matrix's determinant should *not* be 0. To be precise, the determinant must be 1 or −1: when multiplying a matrix to a vector, the resultant vector will have its length be det × original length.

Given 2 linear transformations S, T, then their composition $S \circ T$ is also linear.

4.6.5. Solving for Orthogonal Transformation

Given

$$\boldsymbol{x}_1 = \begin{bmatrix} x \\ y \end{bmatrix} = c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2 \tag{80.1}$$

$$\boldsymbol{x}_2 = \begin{bmatrix} z \\ w \end{bmatrix} = c_3 \boldsymbol{e}_1 + c_4 \boldsymbol{e}_2. \tag{80.2}$$

We can first find out how to eliminate one of the $\boldsymbol{e}$ by solving

$$\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{81.1}$$

$$\gamma \boldsymbol{x}_1 + \delta \boldsymbol{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{81.2}$$

then solve the resulting linear system.

4.7. Transpose of Matrices

Transpose For a matrix, turning its rows to columns, and columns to rows, notated as A^T.
In other words, flipping the matrix along its diagonal axis.

For example,

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \tag{82.1}$$

$$A^\mathsf{T} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \tag{82.2}$$

Some nice conclusions:
- $(AB)^\mathsf{T} = B^\mathsf{T} A^\mathsf{T}$.
- If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and are columns, then $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^\mathsf{T} \boldsymbol{y} = \boldsymbol{y}^\mathsf{T} \boldsymbol{x}$.
- If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, A = A_{m \times n}$, then $\boldsymbol{y} \cdot (A\boldsymbol{x}) = (A^\mathsf{T} \boldsymbol{y}) \cdot \boldsymbol{x}$.

Remember again, matrices $AB \neq BA$.

4.8. Probability Matrices

When we have a system with multiple states, each state being not so stable and has probabilities to switch to other states. We can represent each state as a probability matrix, P_{ij}, for i is the current state, and j is the state that it can switch to.

4.8.1. Markov process

Consider a particle that jumps among 3 distinct states at predetermined probabilities in discrete time steps.

Let $x_{i(n)}$ be the probability of being in state i at time n where $i \in [1, 2, 3], n \in \mathbb{Z}$.

As time goes to $n+1$, $x_{i(n+1)} = P_{i1}x_1(n) + P_{i2}x_2(n) + P_{i3}x_3(n)$. Let $\boldsymbol{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \\ x_3(n) \end{bmatrix}$, then

$$\boldsymbol{x}_{n+1} = \begin{bmatrix} x_1(n+1) \\ x_2(n+1) \\ x_3(n+1) \end{bmatrix} \tag{83.1}$$

$$= \begin{bmatrix} P_{11}x_1(n) + P_{12}x_2(n) + P_{13}x_3(n) \\ P_{21}x_1(n) + P_{22}x_2(n) + P_{23}x_3(n) \\ P_{31}x_1(n) + P_{32}x_2(n) + P_{33}x_3(n) \end{bmatrix} \tag{83.2}$$

$$= \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \\ x_3(n) \end{bmatrix} \tag{83.3}$$

$$= P\boldsymbol{x}_n. \tag{83.4}$$

The P is called a transition matrix. The sum of each of its columns is 1.

This Probability $\times$ State action can be iterated for as many times as pleased, but that is kind of inefficient... We will talk about that later.

4.8.2. Stationary Probability

For a state $\boldsymbol{x}$, if $P\boldsymbol{x} = \boldsymbol{x}$, the state is called a stationary probability, $\boldsymbol{x}_s$.

To find out $\boldsymbol{x}_s$ from P,

$$P\boldsymbol{x}_s = \boldsymbol{x}_s \tag{84.1}$$

$$I\boldsymbol{x}_s - P\boldsymbol{x}_s = \boldsymbol{0} \tag{84.2}$$

$$(I - P)\boldsymbol{x}_s = \boldsymbol{0}. \tag{84.3}$$

Then, form a linear system and solve. The elements of $\boldsymbol{x}_s$ solved are *unlikely to add up to 1*! Hence, we should scale them to have a sum of 1.

4.9. Matrix Inverse

One-to-one linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ is one-to-one if

$$T(\boldsymbol{x}_1) = T(\boldsymbol{x}_2) \text{ iff } \boldsymbol{x}_1 = \boldsymbol{x}_2. \tag{85}$$

It is not to be confused with many-to-one, which that is valid for functions.

Matrix inverse Matrix A is invertible if $\det A \neq 0$. Matrix B is the inverse of A iff $AB = BA = I$, and B is then notated as A^{-1}.

For a matrix to be invertible, its determinant has to *not be 0.*

A matrix's inverse A^{-1} is the augmented part of the RREF of $[A|I]$,

$$A : \boldsymbol{x} \to \boldsymbol{y} \quad \Leftrightarrow \quad A^{-1} : \boldsymbol{y} \to \boldsymbol{x} \tag{86.1}$$

$$A\boldsymbol{x} = I\boldsymbol{y} \quad \Leftrightarrow \quad I\boldsymbol{x} = A^{-1}\boldsymbol{y}. \tag{86.2}$$

$$\text{rref}[A|I] = [I|A^{-1}]. \tag{86.3}$$

For 2×2 matrices, there is the other formula:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{87.1}$$

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{87.2}$$

An invertible matrix $A_{n \times n}$ implies that
- $A : \mathbb{R}^n \to \mathbb{R}^n$ is a one-to-one transformation.
- $\boldsymbol{x} = \boldsymbol{0}$ is the unique solution of $A\boldsymbol{x} = \boldsymbol{0}$.
- $\text{rank}[A] = n$.
- $\text{rref}[A] = I$.
- Columns or rows of A are LI and forms a basis of $\mathbb{R}^n$.

4.10. Determinant

We have used determinant for a long time, but what exactly is a determinant?

Determinant   The 'magnitude' or 'size' of *square* matrices.

For an $n \times n$ matrix, the determinant is the sum of all products of elements at ij where j is permutations of $j \in (\mathbb{Z}^+ \leq n)$, meaning that there are $n!$ terms. Keeping i and permute i also works.

For example,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{88.1}$$

$$\text{has } 2! = 2 \text{ terms} \tag{88.2}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \dots \tag{88.3}$$

$$\text{has } 3! = 6 \text{ terms.} \tag{88.4}$$

Number of permutation   For each permutation $(j_1 j_2 \dots j_n)$ of integers $(1, 2, \dots, n)$, $\sigma(j_1 j_2 \dots j_n)$ is the minimal number of pairwise exchanges between the numbers to recover the order $(1\,2\dots n)$.

Even/odd permutation   The permutations which their permutation number is even/odd.

An odd permutation number adds a negative sign to the term.

But it would be troublesome to calculate the permutation number of each arrangement...

Now recall the row expansion formula like Equation 25°. How do we know which term to add a minus sign to?

$$\det(A) = a_{i1}(-1)^{i+1} \left| A_{i1} \right| \tag{89.1}$$

$$+a_{i2}(-1)^{i+2} \left| A_{i2} \right| \tag{89.2}$$

$$\vdots \tag{89.3}$$

$$+a_{ij}(-1)^{i+j}\left|A_{ij}\right| \tag{89.4}$$

where A_{ij} is the (i, j)th minor matrix of A. The factor of a_{ij} is called the (i, j)th cofactor.

$$\det(A) = \sum_{j=1}^{n} a_{ij}c_{ij}. \tag{90}$$

Cofactor A number that is obtained by multiplying the minor of the element of any given matrix with –1 raised to the power of the sum of the row and column number to which that element belongs.

$$c_{ij} = (-1)^{i+j}\left|A_{ij}\right| \tag{91}$$

Fixing j and progressing i *also works.* Using this formula, we can get determinants of square matrices of any size! A reminder that picking the *row or column* with the most 0s would make the computation easier.

...but, but what if there are not enough 0s in any row or column? I don't want to do math!

In that case, you can instead compute the determinant of an alternate matrix, which is the original matrix, but with a row (a) being itself subtracted by a multiple of another row, $c(b)$, which gives you $(a) = (a) - c(b)$. The determinant will stay the same. Use this rule to construct a row with many 0s to your advantage. *If you happen to time (a) by a coefficient, all its terms should also multiplied with the same coefficient.*

This is because determinant is linear in each row or column.

Important properties of determinants:
- $\det(A) = \det(A^{\mathsf{T}})$.
- $\det(A_{k\leftrightarrow l}) = -\det(A)$ where $A_{k\leftrightarrow l}$ is A with the kth and the lth row/column swapped.
- $\det(\alpha A) = a^n \det(A)$ where A is $n \times n$.

$$\begin{vmatrix} cx_1 & cx_2 \\ cy_1 & cy_2 \end{vmatrix} = c^2 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}. \tag{92}$$

- $\det(A_{k=l}) = 0$ where $A_{k=l}$ is a matrix with the kth and the lth row/column identical. This can assist in proving the next property by

$$\begin{vmatrix} x_1 & x_2 \\ y_1 + a & y_2 + b \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_1 & x_2 \\ a & b \end{vmatrix}. \tag{93}$$

- $\det(A_{k=k+\alpha l}) = \det(A)$.
- $\det(ABCD) = \det(A)\det(B)\det(C)\det(D)$.

Fun fact,

$$\begin{vmatrix} cx_1 & cx_2 \\ y_1 & y_2 \end{vmatrix} = c \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \tag{94}$$

4.11. Matrix Inverse With Determinant

Finally, the ultimate formula for any matrix inverse:

$$A^{-1} = \left(\frac{1}{\det(A)} \right) C^{\mathsf{T}} \tag{95}$$

where C is the cofactor matrix of A.

5. Complex Numbers

Starting with the infamous square root. We know that

$$\sqrt{a} \text{ has } \begin{cases} 2 \text{ real roots if } a > 0 \\ 1 \text{ real root if } a = 0 \\ \text{no real root if } a < 0 \end{cases}. \tag{96}$$

So, we define

$$\sqrt{-1} \equiv i \tag{97.1}$$

$$i^2 = -1. \tag{97.2}$$

The i is the "1" of *imaginary numbers*. Every imaginary number is a real-valued multiple of i, and are usually represented by z.
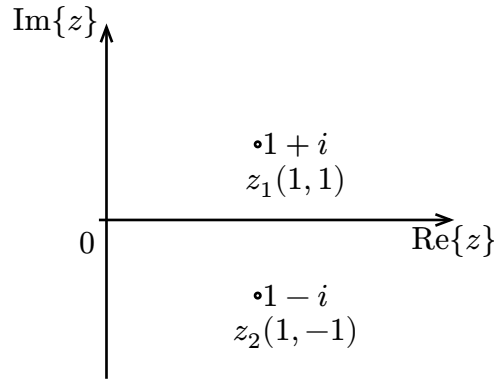
Figure 1: Complex number

We call z_2 the complex conjugate of z_1, or $z_2 = \overline{z}_1$.

5.1. Complex Number Operations

5.1.1. With Scalars

Just do it as usual, with i, of course.

5.1.2. Multiplication

Same.

5.1.3. Length

For $z = a + bi$,

$$|z| = \sqrt{a^2 + b^2} \tag{98.1}$$

$$= \sqrt{a^2 - (-1)b^2} \tag{98.2}$$

$$= \sqrt{a^2 - (bi)^2} \tag{98.3}$$

$$= \sqrt{(a + bi)(a - bi)} \tag{98.4}$$

$$= \sqrt{z\overline{z}}. \tag{98.5}$$

5.1.4. Division

For $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$,

$$\frac{z_1}{z_2} = \frac{a_1 + b_1 i}{a_2 + b_2 i} \tag{99.1}$$

$$= \frac{z_1 \overline{z_2}}{z_2 \overline{z_2}} \tag{99.2}$$

$$= \frac{z_1 \overline{z_2}}{|z_2|^2} \tag{99.3}$$

$$= \frac{(a_1 b_1 + a_2 b_2) + (a_2 b_1 - a_1 b_2)i}{a_2^2 + b_2^2}. \tag{99.4}$$

5.2. Euler's Formula

$$e^{it} = \cos(t) + i \sin(t). \tag{100}$$

Through this, we have a way of obtaining sin and cos values:

$$e^{it} = \cos(t) + i \sin(t) \tag{101.1}$$

$$e^{-it} = \cos(t) - i \sin(t) \tag{101.2}$$

$$\cos(t) = \frac{e^{it} + e^{-it}}{2} \tag{101.3}$$

$$\sin(t) = \frac{e^{it} - e^{-it}}{2i}. \tag{101.4}$$

Now, substitute $t = \pi$:

$$e^{i\pi} + 1 = 0 \tag{102}$$

because for some reason people love it.

5.3. Polar Representation of Complex Numbers

Put an imaginary number into a $x - y$ coordinate system, let it be $z = x + iy$.

Thus, it has a polar representation

$$z = \rho e^{i\theta} \tag{103}$$

where ρ is its length and $\theta \in [0, 2\pi]$ is its angle to the positive x-axis. Why? Use Euler's formula.

To rotate an imaginary number, simply manipulate θ, to change its length, modify ρ.

More specifically, for rotation,

$$e^{i\varphi} \rho e^{i\theta} = \rho e^{i(\theta + \varphi)} \tag{104}$$

You should be able to convert between the two forms. Hint: compute length using Pythagoras theorem, angle using atan, also, Euler's formula.

5.4. Finding Imaginary Roots

Multiples of 2π would round back to the original angle, hence a root can have multiple solutions.

For example,

$$z = \sqrt[3]{1} \tag{105.1}$$

$$= e^{i\theta} \tag{105.2}$$

$$z^3 = 1 \tag{105.3}$$

$$= \left(e^{i\theta}\right)^3 \tag{105.4}$$

$$\theta = 0 \text{ or } \pm \frac{2\pi}{3} \text{ or } \pm \frac{4\pi}{3}. \tag{105.5}$$

6. Eigenvalues and Eigenvectors

Eigenvalue and Eigenvector For a matrix A that defines $\mathbb{R}^n \to \mathbb{R}^n$, a scalar λ and a vector $\boldsymbol{v}$ which

$$\boldsymbol{v} \neq \boldsymbol{0} \tag{106.1}$$

$$A\boldsymbol{v} = \lambda\boldsymbol{v} \tag{106.2}$$

are called an eigen-pair, where λ is the eigenvalue, and $\boldsymbol{v}$ is the eigenvector.

- An eigenvector is never $\boldsymbol{0}$.
- The image of an eigenvector is to the same direction as itself, with length factored by λ.
- $\boldsymbol{v}$ can be replaced by $\hat{\boldsymbol{v}}$ and the equality still holds true.
- An $n \times n$ matrix often has n eigen-pairs.

6.1. Solving for Eigen-pairs

$$A\boldsymbol{v} = \lambda\boldsymbol{v} \tag{107.1}$$

$$A\boldsymbol{v} - \lambda\boldsymbol{v} = \boldsymbol{0} \tag{107.2}$$

$$(A - \lambda I)\boldsymbol{v} = \boldsymbol{0} \tag{107.3}$$

$\det(A - \lambda I)$ must be 0, else $\boldsymbol{v}$ would be $\boldsymbol{0}$, which is not an eigenvector.

$$\det(A - \lambda I) = \boldsymbol{0} \tag{108}$$

is called the characteristic equation.

Algebraic multiplicity (AM) The number of times the *same* eigenvalue occurs as a root of the characteristic equation.

Geometric multiplicity (GM) The number of LI eigenvectors corresponding to an *repeated* eigenvalue.

For an eigenvalue λ, if $\mathrm{AM}(\lambda) = \mathrm{GM}(\lambda)$, then we found all its eigenvectors! This also implies that the eigenvectors are LI, hence the basis of all possible eigenvectors.

6.1.1. 2D Matrices

$$\det(A - \lambda I) = \boldsymbol{0} \tag{109.1}$$

$$= \left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| \tag{109.2}$$

$$= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \tag{109.3}$$

$$= \lambda^2 - \mathrm{Tr}\,\lambda + \det(A) \tag{109.4}$$

where $\mathrm{Tr} = \mathrm{trace}(A) = a_{11} + a_{22}$.

Trace Only for $A_{n \times n}$, the sum of all main diagonal entries, also the sum of all eigenvalues.

Also, determinant is the product of all eigenvalues.

Solving this equation gives us λ.

Then, plug λ back in, $(A - \lambda I)v = 0$ to form a linear system,. Solving this linear system (using rref($[A-\lambda I \mid 0]$)) will give you v. Note that your λ should depend on an unbounded variable, t, as there would be infinite number of solutions, i.e. would solve to a row of 0. We take the simplest pair, which usually involves an 1.

> When solving a homogeneous system, the augmented 0 can be omitted. Also, multiplying a row by a coefficient would not change the result, as the RHS is all 0.

6.1.2. 3D Matrices

$$\det(A - \lambda I) = 0 \tag{110.1}$$

$$= \lambda^3 - \operatorname{Tr} \lambda^2 + c_2 \lambda + \det(A) \tag{110.2}$$

where $c_2 = a_{11}a_{22} + a_{22}a_{33} + a_{33}a_{11} - a_{12}a_{21} - a_{13}a_{31} - a_{23}a_{32}$.

6.2. Complex-valued Eigen-pairs

Eigenvalue or eigenvector can contain imaginary numbers.

For example,

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \tag{111.1}$$

$$\operatorname{trace}(A) = 4 \tag{111.2}$$

$$\det(A) = 5 \tag{111.3}$$

$$\lambda^2 - 4\lambda + 1 = 0 \tag{111.4}$$

$$(\lambda - 2)^2 = -1 \tag{111.5}$$

$$\lambda = 2 \pm i. \tag{111.6}$$

Thus, we have

$$\begin{cases} \lambda_+ = 2 + i \\ \lambda_- = 2 - i \end{cases} \tag{112}$$

Solve the corresponding eigenvectors:

$$(A - \lambda_+ I)v_+ = 0 \tag{113.1}$$

$$\operatorname{rref}(A - \lambda_+ I) = \begin{bmatrix} -1 - i & 2 \\ 0 & 0 \end{bmatrix} \tag{113.2}$$

Let $v_{+_2} = t$,

$$(-1 - i)v_{+_1} + 2t = 0 \tag{114.1}$$

$$v_{+_1} = \frac{2t}{i + 1} \tag{114.2}$$

$$= (1 - i)t \tag{114.3}$$

$$\boldsymbol{v}_+ = \begin{bmatrix} (1-i)t \\ t \end{bmatrix} \tag{114.4}$$

$$= \begin{bmatrix} 1-i \\ 1 \end{bmatrix}. \tag{114.5}$$

Since $\lambda_- = \overline{\lambda_+}$,

$$v_- = \overline{v_+} = \begin{bmatrix} 1+i \\ 1 \end{bmatrix}. \tag{115}$$

> For conjugate eigenvalues, their eigenvectors are also conjugate.

6.3. Solving for Eigen-pairs by Linear Relations

Null space / kernel For all matrix $A_{n\times n}$, its kernel is the set of all vectors that are mapped to zero by A.

$$\ker(A) := \mathcal{N}(A) := \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{0}\} \tag{116}$$

We are supposed to solve

$$(A - \lambda_i I)\boldsymbol{v} = \boldsymbol{0}. \tag{117}$$

...which is

$$\ker(A - \lambda_i I). \tag{118}$$

Linear relation (LR) Given a set of n vectors $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, ...\}$, a linear relation is an *ordered* set of scalars $\{a_1, a_2, a_3, ...\}$ that satisfies

$$a_1\boldsymbol{c}_1 + a_2\boldsymbol{c}_2 + a_3\boldsymbol{c}_3 + ... = \boldsymbol{0}. \tag{119}$$

Thus, we define $\boldsymbol{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^\mathsf{T}$ as a LR of vectors above. If $\boldsymbol{a} \equiv \boldsymbol{0}$, then the LR is trivial.

Now, look:

For all $A_{n\times n}$,

$$\ker(A) := \mathcal{N}(A) := \{\text{LRs of columns of } A\}. \tag{120}$$

A non-trivial kernel is the eigenvector. Just "notice" that you can cancel out the columns... Can't notice? RREF, you may even spot the LR midway.

For example,

$$\begin{bmatrix} -1-i & 2 \\ -1 & 1-i \end{bmatrix} \to (-1-i)\underline{(1-i)} = -2 \tag{121.1}$$

$$(-1)\underline{(1-i)} = i - 1 \tag{121.2}$$

$$\ker \begin{bmatrix} -1-i & 2 \\ -1 & 1-i \end{bmatrix} = \begin{bmatrix} 1-i \\ 1-i \end{bmatrix}. \tag{121.3}$$

6.4. Special Eigen Cases

If the jth column of A is $a e_j$, then $\{a, e_j\}$ is an eigen-pair.

Let x be arbitrary values. For

$$A = \begin{bmatrix} a & x & 0 \\ 0 & x & 0 \\ 0 & x & b \end{bmatrix}, \tag{122}$$

$\{a, e_1\}, \{b, e_3\}$ are eigen-pairs.

For

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}, \tag{123}$$

$\{a, e_1\}, \{b, e_2\}, \{c, e_3\}$ are eigen-pairs.

For

$$A = \begin{bmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{bmatrix}, \tag{124}$$

only $\{b, e_2\}$ is an eigen-pair.

For upper/lower triangular matrices, the head/tail of each row are eigenvalues, and the column of $a e_n$ makes $\{a, e_n\}$ an eigen-pair.

6.5. Application: Probability

Using probability matrix to evaluate the probabilities after a few iteration, we would have to compute

$$A^k v \tag{125}$$

where A is the probability matrix.

If $\{\lambda, v\}$ is an eigen-pair of A, then

$$A^k v = \lambda^k v. \tag{126}$$

But this is kind of useless, as the coincidence can hardly happen.

Remember that, any n-length vector can be expressed as a linear combination of n other vectors with the same length (refer to LI and LD). Oh, if an $n \times n$ matrix has n eigenvectors...

If two eigenvalues are different, their eigenvectors are linearly independent, which means they can be used to construct any vector in the same space.

For example, $A \in \mathbb{R}^2, \boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \lambda_1 \neq \lambda_2$,

$$\boldsymbol{a} = \alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2 \tag{127.1}$$

$$A^k \boldsymbol{a} = A^k(\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2) \tag{127.2}$$

$$= \alpha A^k \boldsymbol{v}_1 + \beta A^k \boldsymbol{v}_2 \tag{127.3}$$

$$= \alpha \lambda_1^k \boldsymbol{v}_1 + \beta \lambda_2^k \boldsymbol{v}_2. \tag{127.4}$$

6.6. Application: Differential Equation

For instance, $m \frac{\mathrm{d}^2 \boldsymbol{x}}{\mathrm{d}t^2}(t) + c \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t) + k\boldsymbol{x}(t) = 0$ describes a motion.

It can be written as two 1st-degree differentials:

$$\begin{cases} x'(t) = v(t) \\ v'(t) = -\frac{k}{m}x(t) - \frac{c}{m}v(t) \end{cases} \tag{128}$$

Let $\boldsymbol{x} = \begin{bmatrix} x(t) \\ v(t) \end{bmatrix}$,

$$\boldsymbol{x}' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} \tag{129.1}$$

$$x_1'(t) = c_{11}x_1(t) + c_{12}x_2(t) \tag{129.2}$$

$$x_2'(t) = c_{21}x_1(t) + c_{22}x_2(t) \tag{129.3}$$

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \tag{129.4}$$

Let A be the matrix of a's. If A has eigen pairs $\{\lambda_1, \boldsymbol{v}_1\}, \{\lambda_2, \boldsymbol{v}_2\}$ and $\lambda_1 \neq \lambda_2$ or $(\lambda_1 = \lambda_2 = \lambda) \wedge (\mathrm{AM}(\lambda) = \mathrm{GM}(\lambda))$, there exists a general solution to the differential equations

$$\begin{cases} x_1 = e^{\lambda_1 t} \boldsymbol{v}_1 \\ x_2 = e^{\lambda_2 t} \boldsymbol{v}_2 \end{cases} \tag{130.1}$$

$$\boldsymbol{x}(t) = c_1 e^{\lambda_1 t} \boldsymbol{v}_1 + c_2 e^{\lambda_2 t} \boldsymbol{v}_2 \tag{130.2}$$

where c_1, c_2 are arbitrary constants.

Given initial value $\boldsymbol{x}(0)$, the constants can be solved.

If the λ's found are conjugates, then the solutions are real:

$$\boldsymbol{x}_1 = \mathrm{Re}\{\boldsymbol{z}_1(t)\} = \frac{\boldsymbol{z}_1 + \boldsymbol{z}_2}{2} \tag{131.1}$$

$$\boldsymbol{x}_2 = \mathrm{Im}\{\boldsymbol{z}_1(t)\} = \frac{\boldsymbol{z}_1 - \boldsymbol{z}_2}{2i} \tag{131.2}$$

where z are the imaginary numbers formed by $e^{\lambda t} \boldsymbol{v}$. Note that the "imaginary" solution is *real*, no i!

In short, if $\boldsymbol{x}' = A\boldsymbol{x}, \lambda_{1,2} = \alpha \pm \beta i, (\alpha, \beta \in \mathbb{R})$,

$$\boldsymbol{x}(t) = c_1 e^{\lambda_1 t} \boldsymbol{v}_1 + c_2 \overline{e^{\lambda_2 t} \boldsymbol{v}_2} \tag{132.1}$$

$$= c_1 \operatorname{Re}\{e^{\lambda_1 t} \boldsymbol{v}_1\} + c_2 \operatorname{Im}\{e^{\lambda_1 t} \boldsymbol{v}_1\}. \tag{132.2}$$

Then, you will compute each value, distribute them to two vectors of the either the real or imaginary the part.

6.7. Application: Vector Transformation

Let A be a matrix that reflects vectors in $\mathbb{R}^2$ in the line at angle $\frac{\pi}{3}$. This is not the easiest, but learn.

$$\varphi = \frac{\pi}{3} \tag{133.1}$$

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix} \tag{133.2}$$

$$\boldsymbol{u}_\perp = \begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}. \tag{133.3}$$

Let $\boldsymbol{v}_1 = \boldsymbol{u}, \boldsymbol{v}_2 = \boldsymbol{u}_\perp$, notice that

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix} + \frac{\sqrt{3}}{4}\begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix} = \frac{1}{4}\boldsymbol{v}_1 + \frac{\sqrt{3}}{4}\boldsymbol{v}_2 \tag{134}$$

Notice the same for $\boldsymbol{e}_2$. Then compute $A\boldsymbol{e}$:

$$A\boldsymbol{e}_1 = A\left(\frac{1}{4}\boldsymbol{v}_1 + \frac{\sqrt{3}}{4}\boldsymbol{v}_2\right) \tag{135.1}$$

$$= \frac{1}{4}\lambda_1 \boldsymbol{v}_1 + \frac{\sqrt{3}}{4}\lambda_2 \boldsymbol{v}_2. \tag{135.2}$$

Same for $A\boldsymbol{e}_2$.

6.8. Application: Probability

For every probability matrix, it always has an eigenvalue $\lambda = 1$. The equilibrium is the eigenvector to the eigenvalue that is 1.

6.9. Matrix Similarity and Diagonalization

Similar matrices Matrices A, B are similar if there exists an invertible matrix T that

$$B = TAT^{-1}. \tag{136}$$

Similar matrices have
- $\operatorname{rank}(A) = \operatorname{rank}(B)$
- $\det(A) = \det(B)$
- $\operatorname{trace}(A) = \operatorname{trace}(B)$
- Identical eigenvalues, but *not necessarily* eigenvectors.

Using similarity, we can finally rewrite matrix powers as

$$A^k = \left(TDT^{-1}\right)^k \tag{137.1}$$

$$= TDT^{-1} \tag{137.2}$$

$$= T \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} T^{-1} \tag{137.3}$$

where D is the diagonal matrix made of A's eigenvalues, and T is the matrix made of eigenvector columns, *in order* of the eigenvalues in D.

Diagonalizable matrix For an $n \times n$ matrix, if there exist n LI eigenvectors, i.e. when all eigenvalues ate unique or each repeated eigenvalue has $\mathrm{AM}(\lambda) = \mathrm{GM}(\lambda)$, then the matrix is diagonalizable.

6.10. Matrix Exponential

So... what the heck is that?

$$e^A \tag{138}$$

Recall our previous differential euqation:

$$\begin{cases} x'(t) = ax(t) \\ x(0) = x_0 \end{cases} \tag{139.1}$$

$$x(t) = e^{at}x_0 \tag{139.2}$$

Instead of scalar a, we now give it a matrix "coefficient":

$$\begin{cases} \boldsymbol{x}'(t) = A\boldsymbol{x}(t) \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases} \tag{140.1}$$

$$\boldsymbol{x}(t) = e^{At}\boldsymbol{x}_0 = ? \tag{140.2}$$

Matrix exponential (e^A) For an $n \times n$ matrix A,

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} \tag{141.1}$$

$$= I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \dots \tag{141.2}$$

The result is a matrix of the same size as A. Note that the same applies to A replaced with a scalar a.

There are two cases:
- A is diagonalizable,
- A is not diagonalizable.

Regardless, we can get the exponential, but the approaches vary.

6.10.1. Diagonalizable Matrix as Power

Let A be a diagonalizable 3×3 matrix.

$$A = TDT^{-1} \tag{142.1}$$

$$T = [\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \boldsymbol{v}_3] \tag{142.2}$$

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \tag{142.3}$$

$$e^{At} = \sum_{k=0}^{\infty} \frac{(TDT^{-1})^k t^k}{k!} \tag{142.4}$$

$$= I + (TDT^{-1})t + \frac{1}{2!}(TDT^{-1})^2 t^2 + \frac{1}{3!}(TDT^{-1})^3 t^3 + \dots \tag{142.5}$$

Notice that

$$\sum_{k=0}^{\infty} \frac{D^k t^k}{k!} = \sum_{k=0}^{\infty} \begin{bmatrix} \frac{(\lambda_1 t)^k}{k!} & 0 & 0 \\ 0 & \frac{(\lambda_2 t)^k}{k!} & 0 \\ 0 & 0 & \frac{(\lambda_3 t)^k}{k!} \end{bmatrix} \tag{143.1}$$

$$= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{(\lambda_1 t)^k}{k!} & 0 & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{(\lambda_2 t)^k}{k!} & 0 \\ 0 & 0 & \sum_{k=0}^{\infty} \frac{(\lambda_3 t)^k}{k!} \end{bmatrix} \tag{143.2}$$

$$= \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{bmatrix}. \tag{143.3}$$

By the same token we used in prior power of matrices, all the in-between $T^{-1}T$ cancel out,

$$e^{At} = T \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{bmatrix} T^{-1} \tag{144.1}$$

$$= Te^{Dt}T^{-1}. \tag{144.2}$$

Now, given a linear system,

$$\begin{cases} x_1'(x) = a_{11}x_1(t) + a_{12}x_2(t) \\ x_2'(x) = a_{21}x_1(t) + a_{22}x_2(t) \end{cases}, \tag{145}$$

we can write is as

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{146.1}$$

$$\boldsymbol{x}'(t) = A\boldsymbol{x}(t) \tag{146.2}$$

$$\boldsymbol{x}(t) = e^{At}\boldsymbol{x}_0 \tag{146.3}$$

$$= Te^{Dt}T^{-1}\boldsymbol{x}_0. \tag{146.4}$$

Elegantly solved.

6.10.2. Non-diagonalizable Matrix as Power

See textbook lol.

Definitions

Linear Dependence (LD)° Non-zero vectors which are parallel to each other, or can be expressed as a scalar multiple of each other, or most generally, $s_1\boldsymbol{a}_1 + s_2\boldsymbol{a}_2 + ... = \boldsymbol{0}$ can be true when *not all s*'s are 0.

In 2D geometry, the parallelogram formed by two LD vectors would have 0 area.

Linear Independence (LI)° Non-zero vectors which satisfy none of the LD requirements.

Linear Combination° A vector expressed as a sum of scalar multiples of other vectors.

Basis° A set of LI vectors that can be used to express any vector in the space. In an $\mathbb{R}^n$ space, the basis has n vectors. Hence, a set of vectors are definitely LD if there are more than n vectors in the set.

Row Echelon Form (REF)° A matrix where the leading element of each row is to the right of the leading element of the row above it.

Rank° The number of LI (non-zero) rows in the REF of matrix, including augmentation.

Homogeneous System° A system where the RHS is all zeros. 0 is always a solution.

Reduced Row Echelon Form (RREF)° A matrix where the leading element (pivot) of each row is 1 and is to the right of the pivot of the row above it. All elements above a pivot are 0. Any column without a pivot is a free variable.

Independent loop° A loop with at least one edge that is not part of any other loop.

Linear transformation / map / function° a rule that assigns one output to each input.

Linearity° A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear iff

$$T(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) = \alpha T(\boldsymbol{x}_1) + \beta T(\boldsymbol{x}_2) \tag{147.1}$$

$$\text{or} \tag{147.2}$$

$$T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2) \text{ and } T(\alpha\boldsymbol{x}) = \alpha T(\boldsymbol{x}). \tag{147.3}$$

Transpose° For a matrix, turning its rows to columns, and columns to rows, notated as A^T. In other words, flipping the matrix along its diagonal axis.

One-to-one linear transformation° $T : \mathbb{R}^n \to \mathbb{R}^m$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ is one-to-one if

$$T(\boldsymbol{x}_1) = T(\boldsymbol{x}_2) \text{ iff } \boldsymbol{x}_1 = \boldsymbol{x}_2. \tag{148}$$

Matrix inverse° Matrix A is invertible if $\det A \neq 0$. Matrix B is the inverse of A iff $AB = BA = I$, and B is then notated as A^{-1}.

Determinant° The 'magnitude' or 'size' of *square* matrices.

Number of permutation° For each permutation $(j_1 j_2 ... j_n)$ of integers $(1, 2, ..., n)$, $\sigma(j_1 j_2 ... j_n)$ is the minimal number of pairwise exchanges between the numbers to recover the order $(1\,2...n)$.

Even/odd permutation° The permutations which their permutation number is even/odd.

Cofactor° A number that is obtained by multiplying the minor of the element of any given matrix with −1 raised to the power of the sum of the row and column number to which that element belongs.

$$c_{ij} = (-1)^{i+j} \left| A_{ij} \right| \tag{149}$$

Eigenvalue and Eigenvector° For a matrix A that defines $\mathbb{R}^n \to \mathbb{R}^n$, a scalar λ and a vector $\boldsymbol{v}$ which

$$\boldsymbol{v} \neq \boldsymbol{0} \tag{150.1}$$

$$A\boldsymbol{v} = \lambda\boldsymbol{v} \tag{150.2}$$

are called an eigen-pair, where λ is the eigenvalue, and $\boldsymbol{v}$ is the eigenvector.

Algebraic multiplicity (AM)° The number of times the *same* eigenvalue occurs as a root of the characteristic equation.

Geometric multiplicity (GM)° The number of LI eigenvectors corresponding to an *repeated* eigenvalue.

Trace° Only for $A_{n \times n}$, the sum of all main diagonal entries, also the sum of all eigenvalues.

Null space / kernel° For all matrix $A_{n \times n}$, its kernel is the set of all vectors that are mapped to zero by A.

$$\ker(A) := \mathcal{N}(A) := \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{0}\} \tag{151}$$

Linear relation (LR)° Given a set of n vectors $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, ...\}$, a linear relation is an *ordered* set of scalars $\{a_1, a_2, a_3, ...\}$ that satisfies

$$a_1\boldsymbol{c}_1 + a_2\boldsymbol{c}_2 + a_3\boldsymbol{c}_3 + ... = \boldsymbol{0}. \tag{152}$$

Thus, we define $\boldsymbol{a} = \begin{bmatrix} a_1 & a_2 & ... & a_n \end{bmatrix}^\mathsf{T}$ as a LR of vectors above. If $\boldsymbol{a} \equiv \boldsymbol{0}$, then the LR is trivial.

Similar matrices° Matrices A, B are similar if there exists an invertible matrix T that

$$B = TAT^{-1}. \tag{153}$$

Similar matrices have
- $\operatorname{rank}(A) = \operatorname{rank}(B)$
- $\det(A) = \det(B)$
- $\operatorname{trace}(A) = \operatorname{trace}(B)$
- Identical eigenvalues, but *not necessarily* eigenvectors.

Diagonalizable matrix° For an $n \times n$ matrix, if there exist n LI eigenvectors, i.e. when all eigenvalues ate unique or each repeated eigenvalue has $\operatorname{AM}(\lambda) = \operatorname{GM}(\lambda)$, then the matrix is diagonalizable.

Matrix exponential (e^A)° For an $n \times n$ matrix A,

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} \tag{154.1}$$

$$= I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \dots \tag{154.2}$$

The result is a matrix of the same size as A. Note that the same applies to A replaced with a scalar a.