MAT1341-LECTURE NOTES-BY ERIC HUA

§ 2.5.1, 2.5.2, 2.5.4 Complex numbers

Complex numbers:

A complex number is of the form *a*+*b*i, where *a* and *b* are real numbers and i is imaginary,

$$i = \sqrt{-1}, \; i^2 = -1.$$

If we have a complex number z, where z=*a*+*bi* then *a* would be the real component (denoted: Re z) and b would represent the imaginary component of z (denoted Im z). Thus the real component of z=4+3i is 4 and the imaginary component would be 3. From this, it is obvious that two complex numbers *(a+bi)* = *(c+di)* if and only if *a*=*c* and *b*=*d*, that is, the real and imaginary components are equal.

The complex number *(a+bi)* can also be represented by the ordered pair *(a,b)* and plotted on a special plane called the complex plane, the horizontal axis is called the real axis and the vertical axis is called the imaginary axis.

Operations

Addition: $(a+bi)+(c+di) = (a+c)+(b+d)i$.
Multiplication: $(a+bi)(c+di) = (ac-bd)+(ad+bc)i$.

Properties: commutative laws, associative laws, distributive laws, zero, unit, inverse.

Conjugate: If $z=a+bi$, then $\bar{z} = \overline{a+bi} = a - bi$.
Properties of conjugate:

(i) $\bar{\bar{z}} = z$;

(ii) $\overline{z+w} = \bar{z} + \bar{w}$;

(iii) $\overline{zw} = \bar{z}\,\bar{w}$;

(iv) If $z \neq 0$, $\overline{(w/z)} = \bar{w}/\bar{z}$;

(v) z is real if and only if $\bar{z} = z$.

Division: $(a+bi)/(c+di) = [(a+bi)(c-di)]/(c^2+d^2)$

Modulus (absolute value):

$|z| = 0$ if and only if $z = 0$;

$|\overline{z}| = |z|$;

$|zw| = |z|\,|w|$;

If $z \neq 0$, $|w/z| = |w|/|z|$;

$|z + w| \leq |z| + |w|$.

Examples. $(2 + 3i) + (-1 + i)$, $i - (2 - i)$; $(1 + i)(1 - i)$; $(2 + 3i)(-1 + i)$; $(1 - i) / (1 + i)$; $(1 + 2i) / (3 - 4i)$.

Quadratics

$ax^2+bx+c=0$, $\Delta = b^2 - 4ac$: two different real roots, when $\Delta > 0$; two same real roots, when $\Delta = 0$; two conjugate complex roots, when $\Delta < 0$.

Example. Find a real irreducible quadratic with 5+12i as a root.

Roots of a polynomial

Any polynomial of degree n has n roots.

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = a_n (x - z_1) \ldots (x - z_n).$$

§ 3.1. Geometric Vectors

Algebraic representation of vectors:

- In R^2: **v** (or $\vec{v}$) = $[a\ b]^T$, or (a,b), where a,b are components.
- In R^3: **v** (or $\vec{v}$) = $[a\ \ b\ \ c]^T$,where a,b,c are components.
- $[a\ \ b\ \ c]^T$ is any directed line segment from (x,y,z) to (x+a,y+b,z+c).
- The zero vector **0** = $[0\ \ 0\ \ 0]^T$.
- Position vectors: $[a\ b\ \ c]^T$ is called the **position vector** of the point (a,b,c).
- In R^n: $[a_1\ \ a_2\ \ \ldots\ a_n\ ]^T$, it is called n-vectors.

Length (magnitude) of **v** = $[a\ b\ \ c]^T$ is: $\|\mathbf{v}\| = \sqrt{a^2+b^2+c^2}$. $\|\mathbf{v}\| \geq 0$, and $\|\mathbf{v}\| = 0$ iff **v** = **0**.

Let A=(a_1, a_2, a_3) , B= (b_1, b_2, b_3). Then $\overrightarrow{AB}$ =$[\ b_1 - a_1\ \ \ b_2 - a_2\ \ \ b_3 - a_3]^T$. The point A is called the tail of $\overrightarrow{AB}$, the point B is called the tip of $\overrightarrow{AB}$.

Arithmetic:

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, then $\mathbf{x}+\mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}$, $c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ cx_3 \end{bmatrix}$

Properties:

$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$,
$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$,
$\mathbf{a} + \mathbf{0} = \mathbf{a}$,
$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$,
$c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$,
$(c + d)\mathbf{a} = c\mathbf{a} + d\mathbf{a}$,
$(cd)\mathbf{a} = c(d\mathbf{a})$,
$1\mathbf{a} = \mathbf{a}$,
a//**b** iff **a**=m**b**.

Standard basis of R^3: $\mathbf{i} = [1\ \ 0\ \ 0]^T$, $\mathbf{j} = [0\ \ 1\ \ 0]^T$, $\mathbf{k} = [0\ \ 0\ \ 1]^T$.

Position vectors can be expressed in terms of standard basis vectors: $[a\ \ b\ \ c]^T = a\mathbf{i}+b\mathbf{j}+c\mathbf{k}$.

Unit vectors: $\mathbf{v} / \|\mathbf{v}\|$ is a unit vector, if $\mathbf{v} \neq \mathbf{0}$.

§ 3.2. Dot Product and Projections

Dot product: Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, then $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$.

Properties:

$\mathbf{a} \bullet \mathbf{a} = \| \mathbf{a} \|^2$,

$\mathbf{a} \bullet \mathbf{b} = \mathbf{b} \bullet \mathbf{a}$,

$\mathbf{a} \bullet (\mathbf{b} + \mathbf{c}) = \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \bullet \mathbf{c}$,

$(c\mathbf{a}) \bullet \mathbf{b} = c(\mathbf{a} \bullet \mathbf{b}) = \mathbf{a} \bullet (c\mathbf{b})$,

$\mathbf{0} \bullet \mathbf{a} = 0$.

Angle between two vectors: $\cos \theta = \mathbf{a} \bullet \mathbf{b} / (\| \mathbf{a} \| \| \mathbf{b} \|)$, always assume $0 \le \theta \le \pi$.

Orthogonality: $\mathbf{a} \perp \mathbf{b}$ iff $\mathbf{a} \bullet \mathbf{b} = 0$.

Vector projection: vector projection of **b** onto **a**, $\text{proj}_{\mathbf{a}}\, \mathbf{b}$ is defined as:

$$\text{proj}_{\mathbf{a}}\, \mathbf{b} = \frac{\vec{a} \bullet \vec{b}}{\| \vec{a} \|^2} \vec{a}.$$

General dot product: $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$.

§ 3.3 Lines and Planes

Lines in space through the point $\mathbf{r}_0 = [x_0 \quad y_0 \quad z_0]^T$ and parallel to $\mathbf{v} = [a \ \ b \ \ c]^T$:

- Vector equation of a line: $\mathbf{r} = \mathbf{r}_0 + t\,\mathbf{v}$, $\mathbf{v} =: [a \ \ b \ \ c]^T$ is a *direction vector*, $t \in R$
- The parametric equation: $x = x_0 + at$, $y = y_0 + bt$, $z = z_0 + ct$.
- *Symmetric equations*: $\dfrac{x - x_0}{a} = \dfrac{y - y_0}{b} = \dfrac{z - z_0}{c}$.

Example. Find the line goes through $p_1 = (x_1, y_1, z_1)$ and $p_2 = (x_2, y_2, z_2)$.

Let $a = x_1 - x_2$, $b = y_1 - y_2$, and $c = z_1 - z_2$.

$\mathbf{r} = \mathbf{p}_2 + t\,[x_1 - y_1 \ \ x_2 - y_2 \ \ x_3 - y_3]^T$,

$\mathbf{r} = \mathbf{p}_2 + t\mathbf{p}_1 - t\mathbf{p}_2$,

$\mathbf{r} = t\mathbf{p}_1 + (1 - t)\mathbf{p}_2$, $t \in R$.

When $t = 0$, $\mathbf{r} = \mathbf{p}_2$; $t = 1$, $\mathbf{r} = \mathbf{p}_1$.

Example. Find the line segment between $p_1 = (x_1, y_1, z_1)$ and $p_2 = (x_2, y_2, z_2)$.

Solution: $\mathbf{r} = t\mathbf{p}_1 + (1 - t)\mathbf{p}_2$, $0 \le t \le 1$.

Determine whether two lines intersect:

$\mathbf{r} = \mathbf{r}_1 + t\mathbf{v}_1$, $\mathbf{r} = \mathbf{r}_2 + s\mathbf{v}_2$.

If they intersect, there exist $t = t^*$ and $s = s^*$ such that $\mathbf{r}_1 + t^*\,\mathbf{v}_1 = \mathbf{r}_2 + s^*\,\mathbf{v}_2$. This gives three equations with two unknown t^* and s^*. If this system of equations has a solution, these two lines intersect. Two non-parallel lines that do not interest are *skew lines*.

Example 1. Determine whether L_1: $x = 1 + t$, $y = -2 + 3t$, $z = 4 - t$ and L_2 : $x = 2s$, $y = 3 + s$, $z = -3 + 4s$ intersect.

$1 + t = 2s$, $-2 + 3t = 3 + s$, $4 - t = -3 + 4s$.

$t = 2s - 1$: $-2 + 6s - 3 = 3 + s$ -> $5s = 8$, $s = 8 / 5$.

$4 - 2s + 1 = -3 + 4s$, $6s = 8$, $s = 8 / 6$.

These two lines do not intersect.

Example 2. L_1: $x = 1 + t$, $y = -2 + 3t$, $z = 4 - t$. L_2 : $x = 2s$, $y = 3 + s$, $z = -3 + 3s$.

$1 + t = 2s$, $-2 + 3t = 3 + s$, $4 - t = -3 + 4s$.

$t = 2s - 1$: $-2 + 6s - 3 = 3 + s$ -> $5s = 8$, $s = 8 / 5$.

$4 - 2s + 1 = -3 + 5s$, $6s = 8$, $s = 8 / 5$.

$t = 2s - 1 = 11 / 5$.

These two lines intersect at $(16 / 5, 23 / 5, 9 / 5)$.

Distance from a point to a line

Distance from a point p to a line L: Take a point q on L. $\mathbf{r} = p - q$ is the vector from q to p. $\mathbf{u} = \text{proj}_{\mathbf{v}}\,\mathbf{r}$ is the projection of $\mathbf{r}$ onto the direction vector $\mathbf{v}$ of L. The distance from p to L is the length of $\mathbf{r} - \mathbf{u}$.

Example. Find the distance from p = (3, −2, 5) to L: $[0\ \ 0\ \ 1]^T + t[1\ \ 1\ \ 1]^T$.

Let q = (0, 0, 1). $\mathbf{r} = p - q = [3\ \ -2\ \ 4]^T$. $\mathbf{v} = [1\ \ 1\ \ 1]^T$. $\mathbf{u} = (\mathbf{r}\cdot\mathbf{v} / \| \mathbf{v} \|^2)\mathbf{v} = [5 / 3\ \ \ 5 / 3\ \ \ 5 / 3]^T$.
Thus $\mathbf{r} - \mathbf{u} = [4 / 3\ \ \ -11 / 3\ \ \ 7 / 3]^T$.
The distance is $d = \| \mathbf{r} - \mathbf{u} \| = (1 / 3)\sqrt{186}$.

Planes

The normal form (*vector equation*): A plane passing through $\mathbf{r}_0$ with normal $\mathbf{n}$ has equation: $\mathbf{n}\cdot(\mathbf{r} - \mathbf{r}_0) = 0$, or $\mathbf{n}\cdot\mathbf{r} = \mathbf{n}\cdot\mathbf{r}_0$.

The scalar equation:

$$[a\ \ b\ \ c]^T \cdot [x - x_0\ \ \ y - y_0\ \ \ z - z_0]^T = 0,$$

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0, \text{ i.e.,} \qquad ax + by + cz = ax_0 + by_0 + cz_0.$$

If we have the scalar form of the equation, $ax + by + cz = a$, then $\mathbf{n} = [a\ \ b\ \ c]^T$.

Examples

1. Equation: $2x - y + 3z = 5$, a normal vector is $[2\ \ -1\ \ 3]^T$.

2. Normal vector $[-1\ \ 3\ \ 2]^T$, point P = (4, 0, −2). Equation: $-x + 3y + 2z = -8$.

The angle between two planes equals the angle between their normal vectors. Two planes are parallel if the normal vectors are parallel. Two planes are perpendicular if their normal vectors are perpendicular.

Distance from a point to a plane:

- **Formula:** The distance from a point $P=(x_0, y_0, z_0)$ to a plane $ax+by+cz+d=0$ is given by $D=\dfrac{|ax_0+by_0+cz_0+d|}{\sqrt{a^2+b^2+c^2}}$.
- Method 2: Finding the distance from a point **p*** to a plane Π

Let the vector form of the plane be **n**· (**p** – $\mathbf{p}_0$) = 0. Then $\mathbf{p}_0$ is on the plane. Let vector **u** = **p*** – $\mathbf{p}_0$. Then the length of the projection **r** = proj $_\mathbf{n}$ **u** is the distance from **p*** to *Π*. The point on *Π* that is closest to **p*** is **u** – **r** + $\mathbf{p}_0$ = **p*** – **r**.

- Method 3: Let **q** be the point on the plane that is closest to **p***. Then the equation of the line joining **q** and **p*** has the equation **p** = **p*** + *t***n**. Hence **q** = **p*** + *t***n** for some *t*. If the equation of the plane is **rn** = *d*, then **qn** = *d*. Hence, **qn** = **p*n** + $t\,\|\mathbf{n}\|^2$, $t = (d - \mathbf{p}^*\mathbf{n}) / \|\mathbf{n}\|^2$.

Examples. **p*** = (1, –1, 2), *Π*: $3x - y - z = 5$.

$\mathbf{p}_0$ = (1, –1, –1). **u** = **p*** – $\mathbf{p}_0$ = (0, 0, 3). **n** = (3, –1, –1). **r** = proj $_\mathbf{n}$ **u** = (**un** / $\|\mathbf{n}\|^2$) **n** = $-\dfrac{3}{11}\begin{bmatrix}3\\-1\\-1\end{bmatrix}$. The distance is ‖ **r** ‖ = $3/\sqrt{11}$. The point on *P* that is closest to **q** = **p*** – **r** =

$$\begin{bmatrix}1\\-1\\2\end{bmatrix}+\frac{3}{11}\begin{bmatrix}3\\-1\\-1\end{bmatrix}=\frac{1}{11}\begin{bmatrix}20\\-14\\19\end{bmatrix}.$$

An alternative solution:

$t = (5-2)/11 = 3/11$. **q** = **p*** + (3 / 11)**n** = $\begin{bmatrix}1\\-1\\2\end{bmatrix}+\dfrac{3}{11}\begin{bmatrix}3\\-1\\-1\end{bmatrix}=\dfrac{1}{11}\begin{bmatrix}20\\-14\\19\end{bmatrix}$. The distance from **p*** to this plane is the distance between **p*** and **q**.

The Cross Product

Definition: If **a** = $[a_1 \quad a_2 \quad a_3]^T$, **b** = $[b_1 \quad b_2 \quad b_3]^T$, then

$$\mathbf{a}\times\mathbf{b} = [a_2b_3 - a_3b_2 \quad a_3b_1 - a_1b_3 \quad a_1b_2 - a_2b_1]^T = \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{vmatrix}.$$

Example. $[1 \quad 3 \quad 4]^T \times [2 \quad 7 \quad -5]^T = [-43 \quad 13 \quad 1]^T$.

Properties:

1. $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.
2. $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{a}$, $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{b}$:

$$a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1) = 0.$$

Finding the equation of a plane containing three given points:

$$A = (x_1, y_1, z_1),\ B = (x_2, y_2, z_2),\ C = (x_3, y_3, z_3).$$

Note that $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are perpendicular to the normal vectors. A normal vector of the plane is $\overrightarrow{AB} \times \overrightarrow{AC}$.

Examples.

1. Find an equation of the plane that passes through the three points P(1,2,3), Q(1,3,2) and R(2,4,3).

Solution. $\overrightarrow{PQ} = [0\ 1\ -1]^T$, $\overrightarrow{PR} = [1\ 2\ 0]^T$. Hence

$$\vec{n} = \overrightarrow{PQ} \times \overrightarrow{PR} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{vmatrix} = 2\vec{i} - \vec{j} - \vec{k}$$

The equation is:

2(x-1) + -1(y-2) + -1(z-3) = 0, i.e., 2x-y-z = -3.

2. Find the symmetric equation of the intersection of $x + y + z = 1$, $x - 2y + 3z = 1$.

The direction vector is perpendicular to both normal vectors:

$$\mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 1 \\ 1 & -2 & 3 \end{vmatrix} = 5\mathbf{i} - 2\mathbf{j} - 3\mathbf{k}$$

To find a point on the line, let $z = 0$. $x + y = 1$, $x - 2y = 1$. $x = 1$, $y = 0$. Hence, (1, 0, 0) is on this line. The equation is $\dfrac{x-1}{5} = \dfrac{y}{-2} = \dfrac{z}{-3}$.

§ 3.5 The Cross Products

In this section we only consider 3-dimensional vectors.

More Properties:

1. $\mathbf{a} \times \mathbf{0} = \mathbf{0}.$
2. $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.
3. $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$.
4. $(c\mathbf{a}) \times \mathbf{b} = c(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (c\mathbf{b})$.
5. $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$.
6. $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$.
7. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$.

8. Lagrange Identity: $\| \mathbf{x} \times \mathbf{y} \|^2 = \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 - (\mathbf{x}\cdot\mathbf{y})^2$.

9. $\| \mathbf{a} \times \mathbf{b} \| = \| \mathbf{a} \| \| \mathbf{b} \| \sin \theta$, the area of the parallelogram.

Since $\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$, $\| \mathbf{x} \times \mathbf{y} \|^2 = \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 - (\mathbf{x}\cdot\mathbf{y})^2 = \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 - \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \cos^2 \theta = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2 \theta$.

10. $\mathbf{a} \,//\, \mathbf{b} \leftrightarrow \mathbf{a} \times \mathbf{b} = \mathbf{0}$. $\mathbf{a}$, $\mathbf{b}$ non-zero.

11. $\mathbf{u}\cdot(\mathbf{v} \times \mathbf{w}) = \mathbf{w}\cdot(\mathbf{u} \times \mathbf{v}) = \mathbf{v}\cdot(\mathbf{w} \times \mathbf{u})$.

12. $| \mathbf{u}\cdot(\mathbf{v} \times \mathbf{w}) |$ is the volume of the parallelepiped spanned by **u**, **v**, and **w**. If the mixed product is 0, then these three vectors are in the same plane.

If $\mathbf{u} = [x_1 \ \ y_1 \ \ z_1]^T$, $\mathbf{v} = [x_2 \ \ y_2 \ \ z_2]^T$, $\mathbf{w} = [x_3 \ \ y_3 \ \ z_3]^T$, then $\mathbf{u}\cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$.

13. $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

Examples.

1. Find the area of the triangle with three vertices P(1,2,3), Q(1,3,2) and R(2,4,3).
Solution. $\overrightarrow{PQ} = (0,1,-1), \overrightarrow{PR} = (1,2,0)$. Hence

$$\overrightarrow{PQ} \times \overrightarrow{PR} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{vmatrix} = 2\vec{i} - \vec{j} - \vec{k}$$

The area of the triangle = $\frac{1}{2} \| \overrightarrow{PQ} \times \overrightarrow{PR} \| = \frac{1}{2} \| 2\vec{i} - \vec{j} - \vec{k} \| = \frac{\sqrt{6}}{2}$.

2. Find the volume of the parallelepiped spanned by

u= $[1 \ \ 2 \ \ 3]^T$, **v=** $[1 \ \ 3 \ \ 2]^T$, and **w=** $[1 \ \ 2 \ \ 2]^T$

Solution:

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{vmatrix} = -1.$$ The volume is $|\,\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\,| = 1$.

§ 1.1 Matrices

Terminology

A matrix is a rectangular array of numbers. Each number is called an entry.

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}], \text{ or } [a_{ij}]_{\text{mxn}}.$$

M is called a matrix with size $m \times n$. The (i,j)-entry is a_{ij}

$M = [C_1\ C_2\ \dots\ C_n] = [R_1\ R_2\ \dots\ R_m]^T$.

Special matrices: matrices with a single row or a single column, square matrices, zero matrices.

Equality of matrices: Same size and same corresponding entries: $a_{ij} = b_{ij}$ for all i and j.

Matrix Addition and Scalar Multiplication

Addition: $A+B = [a_{ij}+b_{ij}]$, Scalar multiplication: $cA = [ca_{ij}]$.

Properties:

$A + B = B + A$,
$A + (B + C) = (A + B) + C$,
$0 + A = A$,
$A + (-A) = 0$,
$c(A + B) = cA + cB$,
$(c + d)A = cA + dA$,
$(cd)A = c(dA)$,
$1A = A$,
$(-1)A = -A$,
$cA = 0$ if and only if $c = 0$ or $A = 0$.

Example. If

$$A = \begin{bmatrix} 0 & 9 \\ 2 & -3 \\ -1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 8 & 1 \\ -7 & 0 \\ 4 & -1 \end{bmatrix} \qquad C = \begin{bmatrix} 2 & 3 \\ -2 & 5 \\ 10 & -6 \end{bmatrix}$$

Then

$$3A+2B-\frac{1}{2}C=\begin{bmatrix}0 & 27\\ 6 & -9\\ -3 & 3\end{bmatrix}+\begin{bmatrix}16 & 2\\ -14 & 0\\ 8 & -2\end{bmatrix}-\begin{bmatrix}1 & \frac{3}{2}\\ -1 & \frac{5}{2}\\ 5 & -3\end{bmatrix}=\begin{bmatrix}15 & \frac{55}{2}\\ -7 & -\frac{23}{2}\\ 0 & 4\end{bmatrix}$$

Transposition

Definition. If A=$[a_{ij}]_{mxn}$, then A^T = =$[a_{ii}]_{nxm}$

Properties:

$(A^T)^T = A$,
$(A + B)^T = A^T + B^T$,
$(cA)^T = cA^T$.

Symmetric matrices: Matrix A is symmetric if $A^T = A$.

If A and B are symmetric, then $A + B$ is symmetric.

Identity matrices: All entries on the main diagonal are 1, other entries are 0.

§5.1 Vector Spaces and Subspaces

Definition. Let V be a *set* of vectors. If operation *addition* and operation *scalar multiplication* are defined in V satisfying the following *axioms*:

A1. If $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} \in V$.
A2. If $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
A3. If $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, then $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
A4. There exists an element, denoted by $\mathbf{0}$, in V, such that $\mathbf{0} + \mathbf{u} = \mathbf{u}$ for every $\mathbf{u}$.
A5. For every $\mathbf{u} \in V$, there exists an element, denoted by $-\mathbf{u}$, in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

S1. Operation *scalar multiplication* is defined for every number c and every $\mathbf{u}$ in V, and $c\mathbf{u} \in V$.
S2. Operation scalar multiplication satisfies the distributive law: $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
S3. Operation scalar multiplication satisfies the second distributive law: $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
S4. Operation scalar multiplication satisfies the associative law: $(cd)\mathbf{u} = c(d\mathbf{u})$.
S5. For every element $\mathbf{u} \in V$, $1\mathbf{u} = \mathbf{u}$.

Then V is called a vector space.

Examples

(i) R^n = {all n-vectors $(a_1, a_2, \ldots, a_n)$}.
(ii) $\mathbf{P}$=the set of all polynomials, $\mathbf{P}_n$=the set of all polynomials of degree at most n.
(iii) $\mathbf{M}_{m,n}$ = the set of all m by n matrices.

Properties

(i) $\mathbf{w} + \mathbf{v} = \mathbf{u} + \mathbf{v}$ implies $\mathbf{w} = \mathbf{u}$.
(ii) $0\mathbf{v} = \mathbf{0}$.
(iii) $a\mathbf{0} = \mathbf{0}$.
(iv) $a\mathbf{v} = \mathbf{0}$ implies $a = 0$ or $\mathbf{v} = \mathbf{0}$.
(v) $(-1)\mathbf{v} = -\mathbf{v}$.
(vi) $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$.

Subspaces

A set U is a subspace of a vector space V if U is a vector space with respect to the operations of V. U is a subspace of V if

(i) the zero vector is in U;
(ii) if $\mathbf{x}$ is in U, then $a\mathbf{x}$ is in U for any scalar a, and
(iii) if $\mathbf{x}$, $\mathbf{y}$ are in U, then $\mathbf{x} + \mathbf{y}$ is in U.

Examples

(i) {**0**} is a subspace. A subspace that is not {**0**} is a *proper subspace*.

(ii) A line through the origin in the space is a subspace; A plane through the origin in the space is a subspace.

(iii) $S_1 = \{[s\ \ 2s\ \ 1]^T;\ s \in \mathrm{R}\}$ is not a subspace.

(iv) $S_2 = \{[s\ \ s^2]^T;\ s \in \mathrm{R}\}$ is not a subspace. It does not satisfy (ii).

(v) $S_3 = \{[x\ \ y]^T;\ x^2 = y^2\}$ is not a subspace. It does not satisfy (iii).

(vi) $S_4 = \{[s+t\ \ \ 2s-t\ \ 3t]^T;\ s, t \in \mathrm{R}\}$ is a subspace.

(vii) R^n is a subspace of itself.

(viii) $\mathbf{P}_n$ is a subspace of $\mathbf{P}$. $\{1, x, x^2, \ldots, x^n\}$ is a spanning set of $\mathbf{P}_n$.

(ix) $\{p \in \mathbf{P}_2;\ p(3) = 0\}$ is a subspace of $\mathbf{P}_2$.

(x) $\{p \in \mathbf{P}_2;\ p(3) = 1\}$ is NOT a subspace of $\mathbf{P}_2$.

(xi) Let U be the set of 2 by 2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $a + b + c + d = 0$. U is a subspace of $\mathbf{M}_{2,2}$:

$\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ satisfies the condition. $\mathbf{0} \in U$.

If $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in U$, and k is a scalar, then $kM = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}$.

Since $ka + kb + kc + kd = k(a + b + c + d) = 0$, $kM \in U$.

If $M_1 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$, $M_2 = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \in U$, then $M_1 + M_2 = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{bmatrix}$.
Since $(a_1 + a_2) + (b_1 + b_2) + (c_1 + c_2) + (d_1 + d_2) = (a_1 + b_1 + c_1 + d_1) + (a_2 + b_2 + c_2 + d_2) = 0$, $M_1 + M_2 \in U$. U is a subspace of $\mathrm{M}_{2,2}$.

(xii) Let $\boldsymbol{M}_{2,2}$ be the vector space that contains all 2×2 matrices, and let $\vec{v}$ be a non-zero vector in R^2. Consider the set S of all matrices A in $\boldsymbol{M}_{2,2}$ such that $\vec{v}$ is an eigenvector of A. Show that S is a subspace of $M_{2,2}$.

Solution. Let A_0 be the 2 by 2 zero matrix. Since $A_0\vec{v} = 0\vec{v}$, the zero matrix is in S.

Assume matrix A is in S, i.e., $A\vec{v} = k\vec{v}$ for some number k, where k is an eigenvalue of A. Then $(cA)\vec{v} = c(A\vec{v}) = ck\vec{v} = (ck)\vec{v}$. Vector $\vec{v}$ is an eigenvector of cA corresponding to eigenvalue ck. Hence, cA is also in S.

Assume matrix A and B are in S, i.e., $A\vec{v} = k_1\vec{v}$, $B\vec{v} = k_2\vec{v}$ for some k_1 and k_2. Then

$(A + B)\vec{v} = A\vec{v} + B\vec{v} = k_1\vec{v} + k_2\vec{v} = (k_1 + k_2)\vec{v}$.

Hence, $\vec{v}$ is an eigenvector of matrix $A + B$ corresponding to eigenvalue $k_1 + k_2$, and $A + B$ is in S.

(xiii) Let A be an m by n matrix. The set of matrices in $M_{m,n}$ that are commutative with A is a subspace of $M_{m,n}$.

(xiv) Let $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$. The set of matrices in $M_{2,2}$ that are commutative with A is a subspace of $M_{2,2}$.

Proof. For any $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $M_{2,2}$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a+3b & 2a-b \\ c+3d & 2c-d \end{bmatrix} = \begin{bmatrix} a+2c & b+2d \\ 3a-c & 3b-d \end{bmatrix}.$$

Hence, $a + 3b = a + 2c$, $c + 3d = 3a - c$, $2a - b = b + 2d$, $2c - d = 3b - d$.

$3b - 2c = 0$, $3a - 2c - 3d = 0$, $2a - 2b - 2d = 0$, $3b - 2c = 0$.

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

$c = s$, $d = t$, $b = (2/3)s$, $a = b + d = (2/3)s + t$.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (2/3)s+t & (2/3)s \\ s & t \end{bmatrix} = \begin{bmatrix} 2/3 & 2/3 \\ 1 & 0 \end{bmatrix}s + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}t.$$

§ 4.1 Subspaces and Spanning

The set of all linear combinations of a set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subseteq$ Vector space V is a subspace of V, called the span of S, denoted by **span S**.

$$\textbf{span } S = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k \mid c_1, \ldots, c_k \in \mathrm{R}\}.$$

Examples

(i) The span of a single non-zero vector in the space is a line through the origin.

(ii) The span of a two non-parallel non-zero vectors **u** and **v** in the space is a plane through the origin with normal vector $\mathbf{u} \times \mathbf{v}$.

Let A be a subspace, and let S be a subset of A. If span $S = A$, then S is a *spanning set* of A. In particular, A itself is a spanning set of A. A subspace generally has more than one spanning set. For example, for a non-zero vector **u**, every subset that contains a non-zero member of the subspace $A = \{k\mathbf{u};\ k \in \mathrm{R}\}$ is a spanning set.

Properties

(i) If $\mathbf{x} \in$ S, then $\mathbf{x} \in$ span S.

(ii) If a subspace W *contains every vector in* S, then W contains span$S$.

As an example of using the second property, span $\{X + Y, X - Y\}$ = span $\{X, Y\}$.

(iii) If **b** is a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$, then

$$\text{span}\{\mathbf{b}, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\} = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}.$$

(iv) $\mathrm{R}^n = \text{span}\{E_1, E_2, \ldots, E_n\}$.

(v) null A = $\{\mathbf{x}: A\mathbf{x}=\mathbf{0}\}$ = the span of the basic solutions of $A\mathbf{x}=\mathbf{0}$.

(vi) im A = the span of the columns of A.

Examples

(i) Verify that $[1\ 2\ 0\ {-1}]^T$ is in span$\{[2\ \ 1\ \ 2\ \ 0]^T, [0\ \ {-3}\ \ 2\ \ 2]^T\}$.

Solution: $[1\ 2\ 0\ {-1}]^T = 0.5[2\ \ 1\ \ 2\ \ 0]^T - 0.5\,[0\ \ {-3}\ \ 2\ \ 2]^T$.

(ii) Verify that the set of vectors $S = \{[1\ \ 2\ \ 3]^T, [-1\ \ 0\ \ 1]^T, [2\ \ 1\ \ {-1}]^T\}$ spans R^3.

Solution: For any $[a\ \ b\ \ c]^T$ in R^3,

$[a\ \ b\ \ c]^T = (3.5a - 6.5b + 3.5c)[1\ \ 2\ \ 3]^T + (0.5a-2.5b+1.5c)\ [-1\ \ 0\ \ 1]^T + (-a+2b-c)\ [2\ \ 1\ \ -1]^T$.

Remark. Later we will study the method to get coefficients as follows :

$$\begin{bmatrix} 1 & -1 & 2 & | & a \\ 2 & 0 & 1 & | & b \\ 3 & 1 & -1 & | & c \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 2 & | & a \\ 0 & 2 & -3 & | & b-2a \\ 0 & 0 & -1 & | & a-2b+c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & | & 3.5a-6.5b+3.5c \\ 0 & 1 & 0 & | & 0.5a-2.5b+1.5c \\ 0 & 0 & 1 & | & -a+2b-c \end{bmatrix}.$$

(iii) Find a,b such that $X = [a\ b\ a{+}b\ a{-}b]^T$ is in span$\{X_1, X_2, X_3\}$, where $X_1 = [1\ \ 1\ \ 1\ \ -1]^T$, $X_2 = [1\ \ 0\ \ 1\ \ 2]^T$, $X_3 = [-1\ \ 0\ \ 1\ \ 0]^T$

(iv) **P** does not have a finite spanning set.

(v) Let $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$. The set of matrices in $M_{2,2}$ that are commutative with A is a subspace of $M_{2,2}$, which is spanned by $\left\{\begin{bmatrix} 2 & 2 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$.

§ 4.2.1, 4.2.3 Linear Independence

Motivation. Reduce the number of vectors in a panning set of a subspace.

Independence Sets of Vectors:

Let $S = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}$ be a set of vectors. An $\mathbf{a}_i$ is a linear combination of the other vectors if and only if there exist $c_1, c_2, \ldots, c_m$, not all zero, such that $\boldsymbol{c}_1\mathbf{a}_1 + \boldsymbol{c}_2\mathbf{a}_2 \ldots + \boldsymbol{c}_m\mathbf{a}_m = \mathbf{0}.$

Definition. A set of vectors $S = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}$ is *linearly independent* (or simply independent) if

$$\boldsymbol{c}_1\mathbf{a}_1 + \boldsymbol{c}_2\mathbf{a}_2 \ldots + \boldsymbol{c}_m\mathbf{a}_m = \mathbf{0}$$

implies $c_1 = c_2 = \ldots = c_m = 0$. Otherwise, S is *linearly dependent*.

Alternative definition. A set of vectors $S = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}$ is *linearly dependent* if one of the vectors in S can be expressed as a linear combination of the other vectors in S. Otherwise, S is linearly independent.

Properties

(i) If $\mathbf{0} \in S$, then S is linearly dependent.

(ii) If $S=\{\mathbf{u}\}$, and $\mathbf{u} \neq \mathbf{0}$, then S is linearly independent.

(iii) $S =\{\mathbf{u}, \mathbf{v}\}$ is linearly independent if and only if **u** and **v** are not parallel.

(iv) If S has two parallel vectors, then it is linearly dependent.

(v) If S is linearly independent, then any subset of S is linearly independent.

(vi) If S is linearly dependent, then any superset of S is linearly dependent.

(vii) A linearly independent set has the minimum number of vectors among all spanning sets of its span.

(viii) Every vector in its span has a unique representation as a linear combination of the vectors in the spanning set.

Independence Test

(i) Construct a linear combination with variable coefficients, and let the combination be the zero vector.

(ii) This set is linearly independent if and only if all the coefficients are 0.

Examples. (i) Test whether the following set of vectors is linearly independent:

$S = \{[1\ \ -2\ \ 3]^T, [5\ \ 6\ \ -1]^T, [3\ \ 2\ \ 1]^T\}$.

$x[1\ \ -2\ \ 3]^T + y[5\ \ 6\ \ -1]^T + z[3\ \ 2\ \ 1]^T = [0\ \ 0\ \ 0]^T$.

This set is linearly dependent, since x=-1, y=-1, z=2 satisfies the equation above.

(ii) If **a**, **b** are linearly independent, show that **a** + 2**b**, and 2**a** − **b** are linearly independent.

$x(\mathbf{a} + 2\mathbf{b}) + y(2\mathbf{a} - \mathbf{b}) = \mathbf{0},$

$(x + 2y)\mathbf{a} + (2x - y)\mathbf{b} = \mathbf{0},$

$x - 2y = 0,\ 2x - y = 0,$

$x = 0,\ y = 0.$

(iii) Let $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ be a set of four vectors in R^3. This set is linearly dependent.

(iv) If a set of vectors contains zero vector, then this set is linearly dependent.

§ 4.3 and 5.2 Independence and Dimension

Basis

Let $S = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_m}\}$ be a set of vectors in a subspace V. S is called a basis of V if

(1) span S =V;

(2) S is linearly independent set.

Dimension of a subspace is the number of vectors in a basis of the subspace.

Examples

(i) $\{1, x, \ldots, x^n\}$ is linearly independent, which is called the standard basis of $\mathbf{P}_n$. Thus dim $\mathbf{P}_n = n + 1$.

(ii) dim $\mathbf{P} = \infty$.

(iii) Let U be the set of 2 by 2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $a + b + c + d = 0$. U is a subspace of $M_{2,2}$. Dimension of U is 3. Show the spanning set is linearly independent.

(iv) Let $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$. The set of matrices in M_{22} that are commutative with A is a subspace of M_{22}. Dimension of this subspace is 2. Show the spanning set is linearly independent.

(v) $\{p \in \mathbf{P}_2;\ p(3) = 0\}$ is a subspace of $\mathbf{P}_2$. Let $p = ax^2 + bx + c$. Then $9a + 3b + c = 0$, $c = -9a - 3b$. Hence, $p = ax^2 + bx - 9a - 3b = a(x^2 - 9) + b(x - 3)$. This subspace is spanned by $x^2 - 9$ and $x - 3$. Since $\{x^2 - 9, x - 3\}$ is linearly independent, the dimension of this subspace is 2.

(vi) The set

$$\mathbf{v}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is a basis for M_{22} and is called the **standard basis** of M_{22}. Thus dim M_{22}=4.

(vii) The set of all m by n matrices with exactly one non-zero entry equal to 1 is a basis for M_{mn} and is called the **standard basis** of M_{mn}. Thus dim M_{mn}=mn.

Remark. Let U be a subspace of R^n. **A *basis* of *U* is a linearly independent spanning set of *U*.**

Fundamental Theorem: If a subspace U is spanned by m vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$, and $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ is a subset of U, with $k > m$, then S is linearly dependent. This theorem may also be expressed as: If a subspace U is spanned by m vectors, and S is an independent subset of U with k vectors, then $k \leq m$.

This is a homogeneous system with m equations and k variables. Since $k > m$, this system has non-trivial solutions. S is linearly dependent.

Invariance Theorem. All bases of a subspace U have the same number of vectors. This number is called the *dimension* of U, denoted by dim U.

Proof. Let B_1 and B_2 be two bases of U. $| B_1 | = m$, $| B_2 | = n$, $m > n$. Since U is spanned by B_2 with n vectors, and B_1 has $m > n$ vectors, B_1 is dependent, a contradiction.

Corollaries

(i) If U has a finite dimension, any independent subset of a subspace U can be extended to a basis of U.

(ii) Every finite spanning set of a subspace U contains a basis of U.

(iii) If U and V are subspace of $\mathbf{R}^n$, $U \subseteq V$, then $\dim U \leq \dim V$. If $\dim U = \dim V$, then $U = V$.

(iv) If U is a subspace with $\dim U = d$, then an independent set of d vectors is a basis, and a spanning set of d vectors is a basis.

Examples

(i) $\dim \{\mathbf{0}\} = 0$.

(ii) The dimension of a line through the origin in the space is 1. A plane through the origin in the space is a subspace of dimension 2.

Bases and Dimension of R^n

Let E_i be the i-th column of the identity matrix I_n. Then $\{\mathrm{E}_1, \mathrm{E}_2, \ldots, \mathrm{E}_n\}$ is a basis of R^n, called the *standard basis* of R^n.

$\dim \mathrm{R}^n = n$.

A spanning set of R^n with n vectors is a basis of R^n, and an independent set of n vectors is a basis of R^n.

§ 1.2 Linear Equations

1. Terminology and Notation

General form of a linear equation with *n variables* (*unknowns* $x_1, x_2, \ldots, x_n$):

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ are called coefficients, b is the constant term.
A solution of the linear equation is a set of numbers $s_1, s_2, \ldots, s_n$, so that if we set $x_1=s_1$, $x_2=s_2$, …, $x_n=s_n$ then the linear equation will be satisfied.

System of linear equations. General form of a system of m linear equations with n variables.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

A solution of the system of linear equations is a set of numbers $s_1, s_2, \ldots, s_n$, so that if we set $x_1=s_1, x_2=s_2, \ldots, x_n=s_n$ then the system of linear equations will be satisfied.

Consistent and inconsistent systems: A system is inconsistent if it has no solution. Otherwise, the system is called consistent.

Three cases for a linear system: Unique solution, infinitely many solutions, no solution.

Examples.

(i) $2x + y = 3$, $x - y = 0$. Unique solution.

(ii) $2x + y = 3$, $4x + 2y = 6$. Infinitely many solutions.

(iii) $2x + y = 3$, $4x + 2y = 2$. No solution.

Geometric interpretation of the three cases with two variables.

(iv) $x_1 - 2x_2 + 3x_3 + x_4 = -3$, $2x_1 - x_2 + 3x_3 - x_4 = 0$.
For any values of s and t, the solution $x_1 = t - s + 1$, $x_2 = t + s + 2$, $x_3 = s$, $x_4 = t$, is a solution of the system, which is in Parametric Form and called General Solution of the system, s and t are *parameters.*

A system has a solution of the parametric form if and only if it has infinitely many solutions.

Coefficient matrix and augmented matrix:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

2. Elementary Operations

Equivalent systems: Two systems have the same set of solutions.

Elementary operations on a system:

(i) Interchanging two equations.

(ii) Multiply an equation by a non-zero number.

(iii) Add a multiple of an equation to another equation.

The matrix form of elementary operations:

(i) Interchanging two rows: $R_i \leftrightarrow R_j$;

(ii) Multiply a row by a non-zero constant: $R_i \to cR_i, c \neq 0$;

(iii) Add a multiple of a row to another row: $R_i \to R_i + cR_j$.

3. Gaussian Elimination

(1). Row Echelon Form and Reduced-Echelon form of a Matrix

(*Row-*)*Echelon form of the augmented matrix:*

(i) All zero rows are at the bottom;

(ii) Each row has a *leading* 1 (first non-zero entry in a row from left);

(iii) The leading 1 in a row is to the right of the leading 1 in rows above it.

If, in addition,

(iv) Each leading 1 is the only non-zero entry in its column,

then this matrix is in *reduced echelon form.*

Examples.

$$\begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The second and the third are in reduced-echelon form.

Non-echelon form:

$$\begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(2). Gaussian Algorithm

Every matrix can be carried to echelon form by the following steps:

Step 1. If the matrix has all zero rows, stop.

Step 2. Find the first column that has a non-zero entry. Move the row with a non-zero entry in this column to the top. This row is called the pivot row, and the non-zero entry is called the pivot entry.

Step 3. Divide the pivot row by the pivot entry to create a leading 1.

Step 4. For every row below the pivot row, add a multiple of the pivot row to make the entry below the leading 1 zero. (If we also make the entries above the leading 1 zero, then we will have the reduced echelon form.)

Step 5. Repeat steps 1 - 4 for the rows below the pivot row.

Example. Solving the following system with Gaussian elimination with back-substitution

$$\begin{aligned} 2x_1 - 2x_2 + 4x_3 + 6x_4 &= 8, \\ -4x_1 + 5x_2 - 2x_3 - 7x_4 &= -10, \\ 2x_1 + x_2 + 22x_3 + 26x_4 &= 36, \\ -3x_1 + 5x_2 - 4x_3 + 11x_4 &= 10. \end{aligned}$$

Solution: At first we reduce the augmented matrix to echelon form by Gaussian elimination:

$$\begin{bmatrix} 2 & -2 & 4 & 6 & 8 \\ -4 & 5 & -2 & -7 & -10 \\ 2 & 1 & 22 & 26 & 36 \\ -3 & 5 & -4 & 11 & 10 \end{bmatrix} \xrightarrow{R_1 \to \frac{1}{2}R_1} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ -4 & 5 & -2 & -7 & -10 \\ 2 & 1 & 22 & 26 & 36 \\ -3 & 5 & -4 & 11 & 10 \end{bmatrix} \xrightarrow[\substack{R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 + 3R_1}]{R_2 \to R_2 + 4R_1} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 3 & 18 & 20 & 28 \\ 0 & 2 & 2 & 20 & 22 \end{bmatrix}$$

$$\xrightarrow[R_4 \to R_4 - 2R_2]{R_3 \to R_3 - 3R_3} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & 0 & 5 & 10 \\ 0 & 0 & -10 & 10 & 10 \end{bmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & -10 & 10 & 10 \\ 0 & 0 & 0 & 5 & 10 \end{bmatrix}$$

$$\xrightarrow{R_3 \to \frac{1}{-10}R_3} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 5 & 10 \end{bmatrix} \xrightarrow{R_4 \to \frac{1}{5}R_4} \begin{bmatrix} 1 & -1 & 2 & 3 & 4 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Now we get an equivalent system

$$\begin{array}{rcrcrcrcr} x_1 & - & x_2 & + & 2x_3 & + & 3x_4 & = & 4, \\ & & x_2 & + & 6x_3 & + & 5x_4 & = & 6, \\ & & & & x_3 & - & x_4 & = & -1, \\ & & & & & & x_4 & = & 2. \end{array}$$

$x_4 = 2$, $x_3 = -1 + x_4 = 1$, $x_2 = 6 - 6x_3 - 5x_4 = -10$, $x_1 = 4 + x_2 - 2x_3 - 3x_4 = -14$.

We can reduce augmented matrix further to a reduced-echelon form:

$$\xrightarrow[\substack{R_3 \to R_3 - 3R_3 \\ R_4 \to R_4 - 2R_2}]{R_1 \to R_1 + R_2} \begin{bmatrix} 1 & 0 & 8 & 8 & 10 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & 0 & 5 & 10 \\ 0 & 0 & -10 & 10 & 10 \end{bmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{bmatrix} 1 & 0 & 8 & 8 & 10 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & -10 & 10 & 10 \\ 0 & 0 & 0 & 5 & 10 \end{bmatrix}$$

$$\xrightarrow{R_3 \to \frac{1}{-10}R_3} \begin{bmatrix} 1 & 0 & 8 & 8 & 10 \\ 0 & 1 & 6 & 5 & 6 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 5 & 10 \end{bmatrix} \xrightarrow{\substack{R_1 \to R_1 - 8R_3 \\ R_2 \to R_2 - 6R_3 \\ R_4 \to \frac{1}{5}R_4}} \begin{bmatrix} 1 & 0 & 0 & 16 & 18 \\ 0 & 1 & 0 & 11 & 12 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\xrightarrow[R_4 \to R_4 + R_3]{\substack{R_1 \to R_1 - 16R_3 \\ R_2 \to R_2 - 11R_3}} \begin{bmatrix} 1 & 0 & 0 & 0 & -14 \\ 0 & 1 & 0 & 0 & -10 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Then the solution can be obtained directly from the reduce-echelon form of the augmented matrix: $x_1 = -14$, $x_2 = -10$, $x_3 = 1$, $x_4 = 2$.

In this example, each of the first n columns has a leading 1, and the system has a unique solution.

(3). Gaussian Elimination

Step 1. Carry the augmented matrix of the system to reduced row echelon form.

Step 2. Assign the nonleading variables (free variables) as parameters. The variables corresponding the leading 1's are called the leading variables.

Step 3. Use the equations corresponding to the reduced row-echelon form to solve for the leading variables in terms of parameters.

Example. Find the solutions of a system if the echelon form of its augmented matrix is

$$\begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solution. Reduce to echelon form:

$$\begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The system becomes:

$$\begin{aligned} x_1 + 2x_2 + \quad x_4 &= 2, \\ x_3 + 3x_4 &= -2. \end{aligned}$$

Let $x_2 = s$, $x_4 = t$. We have $x_1 = 2 - 2s - t$, $x_2 = s$, $x_3 = -2 - 3t$, $x_4 = t$.

(4). Rank

When a matrix M is carried to an echelon form, the number of leading 1's is the *rank* of the matrix, denoted by rank M. If a matrix has p rows and q columns, then rank $(M) \le \min(p, q)$.

Example. $\text{rank}\begin{bmatrix} 1 & 2 & -1 & -2 & 4 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = 3.$

If the rank of the coefficient matrix of a system equals the rank of the augmented matrix that equals the number of variables, then the constant column does not have a leading 1, and this system has a unique solution.

Look at other cases:

Let r be the rank of the augmented matrix, and let r^* be the rank of the coefficient matrix of a system with m equations and n variables.

- Then $r \ge r^*$, $r \le \min(m, n + 1)$, $r^* \le \min(m, n)$.
- If $r > r^*$, i.e., the constant column is a leading column, then this system is inconsistent, i.e., it has no solution.
- Otherwise, i.e., $r = r^*$, i.e., the constant column does not have a leading 1. Then this system is consistent.

If a variable column has a leading 1, it is a leading *column*, and the corresponding variable is a *leading variable*; otherwise, the column is a *free column* and the variable is a *free variable*.

If $r = r^* = n$, every variable is a leading variable, and the system has a unique solution.

If $r = r^* < n$, then there are $n - r$ free variables, and the system has infinitely many solutions.

In this case, the general solution can be expressed in parametric form:

Let every free variable be a parameter. (Hence, there are $n - r$ parameters). Solve the system for leading variables, which are expressed in terms of parameters.

Examples

(i) Solve the system of linear equations:

$x + y - z = 4$, $2x + y + 3z = 0$, $y - 5z = 2$.

Solution. Reduce the augmented matrix of this system to echelon form:

$$\begin{bmatrix} 1 & 1 & -1 & 4 \\ 2 & 1 & 3 & 0 \\ 0 & 1 & -5 & 2 \end{bmatrix} \xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 1 & -1 & 4 \\ 0 & -1 & 5 & -8 \\ 0 & 1 & -5 & 2 \end{bmatrix} \xrightarrow[R_3 \to R_3 - R_2]{R_2 \to -R_2} \begin{bmatrix} 1 & 1 & -1 & 4 \\ 0 & 1 & -5 & 8 \\ 0 & 0 & 0 & -6 \end{bmatrix}.$$

We see that the constant column has a leading 1 (dividing it by –6). This system is inconsistent.

(ii) For which value(s) of p and q does the following system have a unique solution, infinitely many solutions, or no solution?
$2x - 2y + z = 1$,
$x + y - 2z = -3$,
$3x - y - pz = q$.

$$\begin{bmatrix} 2 & -2 & 1 & 1 \\ 1 & 1 & -2 & -3 \\ 3 & -1 & -p & q \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1/2 & 1/2 \\ 1 & 1 & -2 & -3 \\ 3 & -1 & -p & q \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1/2 & 1/2 \\ 0 & 2 & -5/2 & -7/2 \\ 0 & 2 & -p-3/2 & q-3/2 \end{bmatrix} \to$$
$$\begin{bmatrix} 1 & -1 & 1/2 & 1/2 \\ 0 & 1 & -5/4 & -7/4 \\ 0 & 2 & -p-3/2 & q-3/2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1/2 & 1/2 \\ 0 & 1 & -5/4 & -7/4 \\ 0 & 0 & -p+1 & q+2 \end{bmatrix}$$

The system has a unique solution if $-p + 1 \neq 0$, or $p \neq 1$. If $p = 1$ and $q = -2$, the system has infinitely many solutions. If $p = 1$, and $q \neq -2$, the system has no solution.

When $p = 1$, and $q = -2$, the echelon form of the augmented matrix is

$$\begin{bmatrix} 1 & -1 & 1/2 & 1/2 \\ 0 & 1 & -5/4 & -7/4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The solution is $x_3 = t$, $x_2 = 7/4 + (5/4)t$, $x_1 = 1/2 + x_2 - (1/2)x_3 = 1/2 + 7/4 + (5/4)t - (1/2)t$ $= 9/4 + (3/4)t$.

(iii) For which value(s) of a and b the following system has (a) a unique solution, (b) infinitely many solutions, or (c) no solution:

$x + y + z = 1$,
$x - y + z = 3$,
$x + ay + a^2z = b$.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 3 \\ 1 & a & a^2 & b \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & 2 \\ 0 & a-1 & a^2-1 & b-1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & a-1 & a^2-1 & b-1 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & a^2-1 & b+a-2 \end{bmatrix}.$$

This system has a unique solution if $a^2 \neq 1$, or $a \neq \pm 1$. When $a^2 = 1$, this system has infinitely many solutions if $b = 2 - a$. I.e., if $a = 1$, $b = 1$; if $a = -1$, $b = 3$. Otherwise, this system has no solution. In other words, this system has no solution if $a = 1$ and $b \neq 1$, or $a = -1$ and $b \neq 3$.

Example. Given a system of linear equations

$$\begin{aligned} x + 2y + (3-k)z &= 4, \\ 3x - y + 5z &= 2, \\ 4x + y + (k^2 - 12)z &= k + 2, \end{aligned}$$

determine the value(s) of k so that this system

(i) has a unique solution,

(ii) has no solution,

(iii) has infinitely many solutions.

Solution. Bring the augmented matrix of this system to an echelon form:

$$\begin{bmatrix} 1 & 2 & 3-k & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & k^2-12 & k+2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3-k & 4 \\ 0 & -7 & -4+3k & -10 \\ 0 & -7 & k^2+4k-24 & k-14 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 2 & 3-k & 4 \\ 0 & 1 & (4-3k)/7 & 10/7 \\ 0 & -7 & k^2+4k-24 & k-14 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3-k & 4 \\ 0 & 1 & (4-3k)/7 & -10 \\ 0 & 0 & k^2+k-20 & k-4 \end{bmatrix}$$

(i) This system has a unique solution if $k^2 + k - 20 \neq 0$. $k \neq 4$, $k \neq -5$.

(ii) This system has no solution if $k^2 + k - 20 = 0$ and $k - 4 \neq 0$. $k = -5$.

(iii) This system has infinitely many solution if $k^2 + k - 20 = 0$ and $k - 4 = 0$. $k = 4$.

When $k = 4$, the echelon form of the system becomes

$$\begin{bmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & -8/7 & -10 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 9/7 & 24 \\ 0 & 1 & -8/7 & -10 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The solutions are

$x_3 = t$, $x_2 = 10 + (8/7)t$, $x_1 = 24 - (9/7)t$.

Remarks:

(i) The reduced echelon form is unique, so is the leading 1s.

(ii) The general solution includes all particular solutions, but not more.

§ 1.3. Homogeneous Systems

Homogeneous equations and homogeneous systems: all constant terms are zero.

A homogeneous system is always consistent, since $x_1=0, x_2=0, \ldots, x_n=0$ is a solution, which is called trivial solution, other solutions are non-trivial solutions.

If n (number of variables) > m (number of equations), a homogeneous system has infinitely many solutions.

Addition and scale multiplication:

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, then $\mathbf{x}+\mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$, $c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$

Properties of solutions to a homogeneous system:

(i) If $\mathbf{x}$ is a solution, then, for any constant c, $c\mathbf{x}$ is a solution.

(ii) If $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions, then $\mathbf{x}_1 + \mathbf{x}_2$ is also a solution.

A linear combination of columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ is $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m$.

If $\mathbf{v}_i$ are solutions to a homogeneous system, then any linear combination is a solution to this system.

Determine whether a column is a linear combination of a set of columns:

Examples

(i) $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. $c_1 - 2c_2 = -1$, $-2c_1 + 3c_2 = 0$. $c_1 = 3c_2 / 2$, $-c_2 / 2 = -1$. $c_2 = 2$, $c_1 = 3$. $\mathbf{v} = 3\mathbf{v}_1 + 2\mathbf{v}_2$.

(ii) $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. $2c_2 = 1$, $c_1 = 1$, $c_1 + c_2 = 1$. This system has no solution. $\mathbf{v}$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Finding solutions of a homogeneous system with n variables:

If the coefficient matrix has rank n, then this system has only the trivial solution.

If the coefficient matrix has rank $r < n$, then this system has infinitely many non-trivial solutions. The general solution has $n - r$ parameter.

Example. Suppose the coefficient matrix of a homogeneous system is reduced to echelon form:

$$C = \begin{bmatrix} 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}.$$

$x_2 = s$,
$x_5 = t$,
$x_4 = 3x_5 = 3t$,
$x_3 = -4x_5 = -4t$,
$x_1 = 2x_2 - x_5 = 2s - t$.

The general solution can be expressed as a linear combination of two solutions:

$$\mathbf{x} = \begin{bmatrix} 2s-t \\ s \\ -4t \\ 3t \\ t \end{bmatrix} = \begin{bmatrix} 2s \\ s \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ 0 \\ -4t \\ 3t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} -1 \\ 0 \\ -4 \\ 3 \\ 1 \end{bmatrix} t = \mathbf{x}_1 s + \mathbf{x}_2 t.$$

These two solutions $\mathbf{x}_1$ and $\mathbf{x}_2$ are called the *basic solutions*. In general, the general solution of a homogeneous system can be expressed as a linear combination of a set of particular solutions, each corresponds to a parameter. This set of solutions is the set of basic solutions.

Number of basic solutions = number of parameters = number of variables − rank of coefficient matrix.

§ 1.4. Matrix multiplication

Dot product: $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$.

Matrix multiplication: Let A be m by n, B be n by p, then AB is defined to be a matrix whose (i,j)-entry is the dot product of the ith row of A and the jth column of B.

If A=$[a_{ij}]_{m\times n}$, B=$[b_{ij}]_{n\times p}$, then AB=$[c_{ij}]_{m\times p}$, where
$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$.

Examples.

(i) $\begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 2 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \left[\begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right] = \begin{bmatrix} -2 & 2 \\ 0 & -1 \end{bmatrix}$

(ii) $\begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 2 \end{bmatrix} = \left[\begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right] = \begin{bmatrix} -5 & 6 & 5 \\ -2 & 3 & 2 \\ 1 & 0 & -1 \end{bmatrix}$.

(iii) $\begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 2 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 0 & -1 \end{bmatrix}$.

(iv) $\begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 2 \end{bmatrix} = \begin{bmatrix} -5 & 6 & 5 \\ -2 & 3 & 2 \\ 1 & 0 & -1 \end{bmatrix}$.

Powers of square matrices: $M^0 = I$, $M^{n+1} = MM^n$.

Properties of matrix multiplication:

(i) Let A be an m be n matrix. $I_m A = A$, $AI_n = A$.

Suppose the dimensions of the matrices are compatible with operations performed.

(ii) $A(BC) = (AB)C$,

(iii) $A(B \pm C) = AB \pm AC$,

(iv) $(B \pm C)A = BA \pm CA$,

(v) $(cA)B = c(AB)$,

(vi) $(AB)^T = B^T A^T$.

(vii) $AB \neq BA$ even both are defined.

$$\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -2 & -1 \end{bmatrix}.$$

(viii) $AB = AC$ or $BA = CA$ does not imply $B = C$.

$$\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -2 & -2 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -2 & -2 \end{bmatrix}.$$

Example. Find all 2 by 2 matrices that are commutative with $B = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$.

Let $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Then $AB = \begin{bmatrix} a+2c & -c \\ b+2d & -d \end{bmatrix} = BA = \begin{bmatrix} a & c \\ 2a-b & 2c-d \end{bmatrix}$. Hence, $a + 2c = a$, and $c = 0$. $b + 2d = 2a - b$. $2b + 2d = 2a$. $a = b + d$. A matrix A that commutes with B if and only if A is of the type $A = \begin{bmatrix} b+d & 0 \\ b & d \end{bmatrix}$, where b and d can be any real number.

Matrix equation of a linear system:

The following system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

can be written as a matrix equation

AX=B, where

A= $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, X = $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is called the matrix of variables, B = $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ bm \end{bmatrix}$ is called the matrix of constants.

AX=0 is called the associated homogeneous system.

Theorem. If X_0 is a particular solution of AX=B, and X_1 is the general solution of AX=0, then $X_0 + X_1$ is the general solution of AX=B.

Block multiplication:

Let A=[A_1 A_2 … A_n], X =[x_1 x_2 … x_n]^T . Then AX = $x_1A_1 + x_2A_2 + \ldots + x_nA_n$.

Properties:

(i) $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$,

(ii) $(cA)\mathbf{x} = A(c\mathbf{x}) = c(A\mathbf{x})$,

(iii) $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$.

$$\begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{m1} & c_{m2} & \ldots & c_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \begin{bmatrix} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n \\ c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n \\ \ldots \\ c_{m1}x_1 + c_{m2}x_2 + \ldots + c_{mn}x_n \end{bmatrix}$$

$AB = A[B_1\ B_2\ \ldots\ B_n] = [AB_1\ \ AB_2\ \ \ldots\ \ AB_n]$.

§ 4.4 Rank

Row Space and Column Space

Recall that the *rank* of a matrix is the number of leading ones in its echelon form.

Definition. Let A be an m by n matrix.

row A = span{the rows of A} $\subseteq \mathbf{R}^m$,

col A = span{the columns of A} $\subseteq \mathbf{R}^n$.

null A (or ker(A)) = {$\mathbf{x}$: A$\mathbf{x} = \mathbf{0}$}

Properties:

(i) If N is obtained from A by a row elementary operation, then row N = row A.

If N is in echelon form, then the non-zero rows of N are linearly independent. They form a basis of row N = row A. Hence,

(ii) dim (row A) = rank A.

(iii) If N is obtained from A by a column elementary operation, then col N = col A.

(iv) col M = im M. A basis of col M is also a basis of im M.

Proof. Let $M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$. A vector $\mathbf{v}$ is in col M if and only if

$$\mathbf{v} = c_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{bmatrix} + \dots + c_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} c_1 a_{11} + c_2 a_{12} + \dots + c_n a_{1n} \\ c_1 a_{21} + c_2 a_{22} + \dots + c_n a_{2n} \\ \dots \\ c_1 a_{m1} + c_2 a_{m2} + \dots + c_n a_{mn} \end{bmatrix} = M \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix}.$$

Hence, col M = im M.

Rank Theorem, Null Space and Image

dim (col A) = dim (row A) = rank A. When A is reduced to echelon form, the non-zero rows form a basis of row A, and the columns in A where the echelon form have leading 1s form a basis of col A.

Example. Find a basis of the row space, a basis of the column space and ker(M) of the matrix

$$M = \begin{bmatrix} 1 & 3 & 1 & 2 \\ 1 & 2 & 0 & 3 \\ 2 & 5 & 1 & 5 \end{bmatrix}.$$

Carry M to echelon form:

$$\begin{bmatrix} 1 & 3 & 1 & 2 \\ 1 & 2 & 0 & 3 \\ 2 & 5 & 1 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 & 2 \\ 0 & -1 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A basis of row M is $\{(1, 3, 1, 2), (0, 1, 1, -1)\}$. A basis of col M is $\{[1\ 1\ 2]^T, [1\ 0\ 1]^T\}$. The rank of M is 2.

$$\text{Ker(M)} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} : x_1 - 2x_3 + 5x_4 = 0,\ x_2 + x_3 - x_4 = 0 \right\}.$$

Properties of the rank of a matrix:

(i) rank A = rank A^T.

(ii) If U and V are invertible, then rank A = rank UA = rank AV.

(iii) An n by n square matrix A is invertible if and only if rank $A = n$.

(iv) rank $AB \le \min\{\text{rank } A, \text{rank } B\}$.

(v) Let A be m by n matrix. Then

$$\dim(\text{null } A) = n - \text{rank } A.$$

$$\dim(\text{null } A) + \dim(\text{im } A) = n.$$

(vi) The basic solutions of a homogeneous system $A\mathbf{x} = \mathbf{0}$ form a basis of null A.

Example. Let $A = \begin{bmatrix} 1 & -2 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}$. The solution set of $A\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} 2s+2t \\ s \\ -4t \\ 3t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} 2 \\ 0 \\ -4 \\ 3 \\ 1 \end{bmatrix} t$$. The null space of A is $\text{null}\, A = \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} 2 \\ 0 \\ -4 \\ 3 \\ 1 \end{bmatrix} t;\ s, t \in \mathbb{R} \right\}$.

Since the basic solutions are linearly independent, they form a basis of the null space of A. The dimension of the null space is 2.

The dimension of the null space = the number of solutions in a set of basic solutions = the number of parameter variables = the number of variables – rank of the matrix.

The image space is

$$\textbf{im}\, \boldsymbol{A} = \left\{ \begin{bmatrix} 1 & -2 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}, a, b, c, d, e \in R \right\} = \left\{ \begin{bmatrix} a - 2b - d + 4e \\ c + 2d - 2e \\ d - 3e \end{bmatrix}, a, b, c, d, e \in R \right\}.$$

Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \\ -3 \end{bmatrix} \right\}$.

This set is linearly dependent because $\mathbf{v}_2 = -2\mathbf{v}_1$, and $\mathbf{v}_5 = -2\mathbf{v}_1 + 4\mathbf{v}_3 - 3\mathbf{v}_4$.
Since $\mathbf{v}_1$, $\mathbf{v}_3$, and $\mathbf{v}_4$ are linearly independent, this set is a basis of im A, and dim im A is 3.

im A = span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ = span$\{\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_4\}$.

Properties: The following statements are equivalent for an m by n matrix A ($m \geq n$):

(a) rank $A = n$;
(b) row $A = \mathbf{R}^n$;
(c) The columns of A are linearly independent;
(d) $A^T A$ is invertible;
(e) There exists an n by m matrix C, $CA = I_n$, $C = (A^T A)^{-1} A^T$
(f) If $AX = 0$, then $X = 0$.

Properties: The following are equivalent for an m by n matrix A ($m \leq n$):

(a) rank $A = m$;
(b) col $A = \mathbf{R}^m$;

(c) The rows are linearly independent in $\mathbf{R}^n$;
(d) AA^T is invertible;
(e) There exists an n by m matrix C, $AC = I_m$, $C= A^T (AA^T)^{-1}$.
(f) $AX = B$ is consistent for any B in $\mathbf{R}^m$.

§ 1.5 Matrix Inverses

Definition. Let A be a square matrix. If C is a square matrix such that $AC = CA = I$, then C is the *inverse* of A, denoted by A^{-1}. If A has an inverse, then A is *invertible.*

In fact, we need only $AC = I$ or $CA = I$. The inverse of a matrix is unique.

Properties of the inverse of a matrix:

(i) $I^{-1} = I$,
(ii) $(A^{-1})^{-1} = A$,
(iii) $(AB)^{-1} = B^{-1}A^{-1}$, (if both A and B are invertible square matrices)
(iv) If A is invertible, then A^k is invertible and $(A^k)^{-1} = (A^{-1})^k$,
(v) If A is invertible, then A^T is invertible, and $(A^T)^{-1} = (A^{-1})^T$,
(vi) If a is invertible, then cA is invertible, and $(cA)^{-1} = (1 / c)A^{-1}$.

2 by 2 matrix: Let A=$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We define the determinant det$A$ and the adjugate adj$A$ as:

$$\det A = ad-bc, \quad \text{adj}A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Then $$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{adj\, A}{\det A}.$$

Matrix inversion algorithm:

$$[A \mid I] \to [I \mid A^{-1}],$$

where the row operations on A and I are carried out simultaneously.

Example. Show that $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 2 & -1 & 3 \end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 2 & 1 \\ -5 & 1 & 2 \\ -1 & -1 & 1 \end{bmatrix}$.

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

$$\to \left[\begin{array}{cccccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & -2 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1/3 & -1/3 & 1/3 \end{array}\right] \to \left[\begin{array}{cccccc} 1 & 0 & 0 & -1/3 & 2/3 & 1/3 \\ 0 & 1 & 0 & -5/3 & 1/3 & 2/3 \\ 0 & 0 & 1 & -1/3 & -1/3 & 1/3 \end{array}\right]$$

Inverses of triangular matrices:

$$U = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}, \quad L = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

A triangular matrix is invertible if and only if no entry on the main diagonal is zero.

The inverse of an upper (lower) triangular matrix is also an upper (lower) triangular matrix.

Inverse and Linear system:

If the coefficient matrix A of an n by n system $A\mathbf{x} = \mathbf{b}$ has an inverse, then $\mathbf{x} = A^{-1}\mathbf{b}$.

Properties: For a square matrix A of an n by n, the following statements are equivalent:

(i) A is invertible.

(ii) The homogeneous system of linear equations $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(iii) The system of linear equations $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in R^n.

(iv) Every $\mathbf{b}$ in R^n is a linear combination of the columns of A.

(v) A can be reduced to I.

(vi) There exists B such that $AB = I$.

(vii) There exists B such that $BA = I$.

Matrix Inverse and Matrix Transformation

Inverse transformation. If $T\mathbf{x} = M\mathbf{x}$ is a matrix transformation, and M is invertible, then $T^{-1}\mathbf{x} = M^{-1}\mathbf{x}$ is the inverse of T.

§4.2.2 Invertibility of Matrices

Let A be n by n matrix, the following statements are equivalent:

(a) A is invertible;
(b) The columns of A span $\mathbf{R}^n$;
(c) The rows of A are linearly independent;
(d) The columns of A are linearly independent;
(e) The rows of A span $\mathbf{R}^n$.

Examples. Test whether the following set of vectors is linearly independent:

$S = \{[1\ \ -2\ \ 3]^T, [5\ \ 6\ \ -1]^T, [3\ \ 2\ \ 1]^T\}$.

Solution: Calculate the determinant of the matrix formed from the three vectors.

§ 4.5 Orthogonality

1. Dot product in general

Suppose X = $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, Y = $\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

Dot product: $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + ... + x_n y_n$.

2. *Length* of X is defined as: ||X|| = $\sqrt{x_1{}^2 + ... + x_n{}^2}$.

Properties:

$X \cdot Y = Y \cdot X$;
$X \cdot (Y + Z) = X \cdot Y + X \cdot Z$;
$(aX) \cdot Y = a(X \cdot Y)$;
$X \cdot X = \| X \|^2$;
$\| X \| \geq 0$ and $\| X \| = 0$ iff $X = 0$;
$\| aX \| = | a | \, \| X \|$;
$| X \cdot Y | \leq \| X \| \, \| Y \|$; (Cauchy Inequality)
$\| X + Y \| \leq \| X \| + \| Y \|$; (Triangle Inequality)

Distance between two vectors:

$d(X, Y) = \| X - Y \|$.

Properties of distances:

$d(X, Y) \geq 0$ and $d(X, Y) = 0$ iff $X = Y$;
$d(X, Y) = d(Y, X)$;
$d(X, Y) \leq d(X, Z) + d(Z, X)$.

3. Orthogonality in general and Pythagoras' theorem

Definition. X are Y are *orthogonal* if $X \cdot Y = 0$, $X, Y \neq 0$. A set of vectors $S = \{X_1, X_2, \ldots, X_k\}$ is orthogonal if X_i and X_j are orthogonal for every pair $i \neq j$.

Example. The set $\{X_1, X_2, X_3, X_4\}$, where $X_1 = [1\ \ 1\ \ 1\ \ -1]^T$, $X_2 = [1\ \ 0\ \ 1\ \ 2]^T$, $X_3 = [-1\ \ 0\ \ 1\ \ 0]^T$, $X_4 = [-1\ \ 3\ \ -1\ \ 1]^T$ is orthogonal.

Properties: (i) An orthogonal set is linearly independent.

Indeed, if $c_1X_1 + c_2X_2 + \ldots + c_kX_k = 0$, then $X_i\,(c_1X_1 + c_2X_2 + \ldots + c_kX_k) = c_i \| X_i \|^2 = 0$. then $c_i = 0$.

(ii) If $S = \{X_1, X_2, \ldots, X_k\}$ is orthogonal, then

$\| X_1 + X_2 + \ldots + X_k \|^2 = \| X_1 \|^2 + \| X_2 \|^2 + \ldots + \| X_k \|^2$. (General Pythagoras's Theorem)

Definition. $S = \{X_1, X_2, \ldots, X_k\}$ is *orthonormal* if S is orthogonal and $\| X_i \| = 1$ for every i.

The standard basis $\{E_1, E_2, \ldots, E_n\}$ is orthonormal.

If S is an orthogonal set, $X_i \neq 0$, then we can *normalize* this set.

Example. Normalize the set $\{X_1, X_2, X_3, X_4\}$, where $X_1 = [1\ \ 1\ \ 1\ \ -1]^T$, $X_2 = [1\ \ 0\ \ 1\ \ 2]^T$, $X_3 = [-1\ \ 0\ \ 1\ \ 0]^T$, $X_4 = [-1\ \ 3\ \ -1\ \ 1]^T$ is orthogonal.

$\| X_1 \| = 2$, $\| X_2 \| = \sqrt{6}$, $\| X_3 \| = \sqrt{2}$, $\| X_4 \| = \sqrt{11}$. The set

$\{[1\ \ 1\ \ 1\ \ -1]^T/2,\quad [1\ \ 0\ \ 1\ \ 2]^T/\sqrt{6},\quad [-1\ \ 0\ \ 1\ \ 0]^T/\sqrt{2},\quad [-1\ \ 3\ \ -1\ \ 1]^T/\sqrt{11}\}$ is orthonormal.

4. Expansion Theorem

Let $B = \{X_1, X_2, \ldots, X_m\}$ be an orthogonal basis of a subspace U. Since B is a basis, every vector X in U is a linear combination of vectors in B:

$X = c_1X_1 + c_2X_2 + \ldots + c_mX_m$,

which is called Fourier expansion. Since $X \cdot X_i = c_i \| X_i \|^2$, and so $c_i = X \cdot X_i / \| X_i \|^2$, which are called Fourier coefficients.

Example. Expand $X = [-4\ \ 13\ \ 2\ \ 7]^T$ as a linear combination of the set $\{X_1, X_2, X_3, X_4\}$, where $X_1 = [1\ \ 1\ \ 1\ \ -1]^T$, $X_2 = [1\ \ 0\ \ 1\ \ 2]^T$, $X_3 = [-1\ \ 0\ \ 1\ \ 0]^T$, $X_4 = [-1\ \ 3\ \ -1\ \ 1]^T$.

$c_1 = X \cdot X_1 / \| X_1 \|^2 = 4/4 = 1$, $c_2 = X \cdot X_2 / \| X_2 \|^2 = 2$, $c_3 = 3$, $c_4 = 4$.

$X = X_1 + 2X_2 + 3X_3 + 4X_4$

Gram-Schmidt Algorithm

Let $\{X_1, X_2, \ldots, X_k\}$ be a basis of a subspace U. We construct an orthogonal set as follows:

$F_1 = X_1$;
$F_2 = X_2 - (X_2 \cdot F_1 / \| F_1 \|^2)F_1$;
…
$F_i = X_i - (X_i \cdot F_1 / \| F_1 \|^2)F_1 - (X_i \cdot F_2 / \| F_2 \|^2)F_2 - \ldots - (X_i \cdot F_{i-1} / \| F_{i-1} \|^2)F_{i-1}$,
…

$F_k = X_k - (X_k \cdot F_1 / \| F_1 \|^2)F_1 - (X_k \cdot F_2 / \| F_2 \|^2)F_2 - \ldots - (X_k \cdot F_{k-1} / \| F_{k-1} \|^2)F_{k-1}$.

Then $\{F_1, F_2, \ldots, F_k\}$ is orthogonal.

Let $j < i$. $F_j \cdot F_i = F_j \cdot ((X_i - (X_i \cdot F_1 / \| F_1 \|^2)F_1 - (X_i \cdot F_2 / \| F_2 \|^2)F_2 - \ldots - (X_i \cdot F_{i-1} / \| F_{i-1} \|^2)F_{i-1})$
$= F_j \cdot X_i - (X_i \cdot F_j / \| F_j \|^2)F_j \cdot F_j = 0$.

$\{F_1, F_2, \ldots, F_k\}$ is linearly independent, hence it is a basis of U.

Example. Use the Gram-Schmidt algorithm to convert the given basis

$$B=\{[1\ \ 0\ \ 1\ \ 2]^T, [0\ \ 1\ \ 2\ \ 0]^T,\ \ [0\ \ 0\ \ -1\ \ 1]^T\}$$

into orthogonal basis.

QR-Factorization

Theorem. If A is an m by n matrix with linearly independent columns, then

$$A=QR$$

where Q is an m by n matrix with orthonormal vectors, and R is an invertible upper triangular matrix.

§ 4.6 Projections and Approximation

Let U be a subspace of $\mathbf{R}^n$. The *orthogonal complement* of U is

$U^{\perp} = \{X \text{ in } \mathbf{R}^n; X\cdot Y = 0 \text{ for all } Y \text{ in } U\}$.

Properties: Let $U^{\perp}$ be a subspace of $\mathbf{R}^n$.

(i) Let U=span$\{F_1, F_2, \ldots, F_k\}$. X in $U^{\perp}$ iff $X\cdot F_i = 0$, $i = 1, 2, \ldots, k$.

(ii) $U^{\perp}$ is a subspace of $\mathbf{R}^n$.

(iii) $\dim(U) + \dim(U^{\perp}) = n$.

Finding $U^{\perp}$:

$U = \text{span}\,\{[1\ \ -1\ \ 2\ \ 0]^T, [1\ \ 0\ \ -2\ \ 3]^T\}$. $U^{\perp} = \{X; X\cdot[1\ \ -1\ \ 2\ \ 0]^T = 0, X\cdot[1\ \ 0\ \ -2\ \ 3]^T = 0\}$.

$x_1 - x_2 + 2x_3 = 0$, $x_1 - 2x_3 + 3x_4 = 0$.

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 1 & 0 & -2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -4 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & -4 & 3 \end{bmatrix}.$$

$x_3 = s$, $x_4 = t$, $x_1 = 2s - 3t$, $x_2 = 4s - 3t$.

$U^{\perp} = \text{span}\,\{[2\ \ 4\ \ 1\ \ 0]^T, [-3\ \ -3\ \ 0\ \ 1]^T\}$.

Orthogonal decomposition

Let U be a subspace of $\mathbf{R}^n$. Every vector X in $\mathbf{R}^n$ can be expressed as a sum of two vectors:

$X = X_1 + X_2$, where X_1, called the *orthogonal projection* of X on U, is in U, and X_2 is in $U^{\perp}$.

Indeed, let $U = \text{span}\,\{F_1, F_2, \ldots, F_k\}$, where $\{F_1, F_2, \ldots, F_k\}$ is an orthogonal basis of U. Then

$X_1 = \text{proj}_U(X) = (X\cdot F_1 / \| F_1 \|^2)F_1 + (X\cdot F_2 / \| F_2 \|^2)F_2 + \ldots + (X\cdot F_k / \| F_k \|^2)F_k$

$X_2 = X - X_1$ is in $U^{\perp}$.

Proof. $X_2 \cdot F_1 = X\cdot F_1 - X_1 \cdot F_1 = X\cdot F_1 - (X\cdot F_1 / \| F_1 \|^2)F_1 \cdot F_1 = 0$.

It can be shown that this decomposition is independent of the choice of the spanning set.

Example. Let $U = \text{span}\,\{[1\ \ 0\ \ 0]^T, [0\ \ 1\ \ 1]^T\}$. Decomposite the vector X= 1 0 $0]^T$ into two vectors, one is in U, one is in $U^{\perp}$.

Solution: Let $F_1= [1\ \ 0\ \ 0]^T$, $F_2= [0\ \ 1\ \ 1]^T$. Note that the two vectors in U are orthogonal. So

$X_1 = \text{proj}_U(X) = (X\cdot F_1 / \| F_1 \|^2)F_1 + (X\cdot F_2 / \| F_2 \|^2)F_2 = F_1 + 3F_2 = [1\ \ 3\ \ 3]^T$.

$X_2 = X - X_1 = [0\ \ -1\ \ 1]^T$.

§ 2.1 Cofactor Expansion

Determinant of a square matrix. Let A be $n\times n$ square matrix $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$.

Sub-matrix A_{ij} denotes the $(n-1)\times(n-1)$ matrix obtained from A by deleting the row i and the column j. C_{ij} (A)= $(-1)^{i+j}\det(A_{ij})$ is the (i, j)-cofactor.

Cofactor expansion along i-th row or j-th column:

$$\begin{aligned}\det(A) &= a_{i1} C_{i1}(A) + a_{i2} C_{i2}(A) + \ldots + a_{in} C_{in}(A) \\ &= a_{1j} C_{1j}(A) + a_{2j} C_{2j}(A) + \ldots + a_{nj} C_{nj}(A).\end{aligned}$$

Example. Down the first column:

$$\begin{vmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 1 & 4 & -2 \end{vmatrix} = 1\begin{vmatrix} 1 & -1 \\ 4 & -2 \end{vmatrix} - 2\begin{vmatrix} -2 & 1 \\ 4 & -2 \end{vmatrix} + 1\begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix} = 2 - 0 + 1 = 3.$$

Across the second row:

$$\begin{vmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 1 & 4 & -2 \end{vmatrix} = -2\begin{vmatrix} -2 & 1 \\ 4 & -2 \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ -2 & 4 \end{vmatrix} = 0 - 3 + 6 = 3.$$

Example. Along the second column:

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

$$= 2\times 3 + 2\left(1\begin{vmatrix} -1 & -1 \\ 1 & 4 \end{vmatrix} + (-2)\begin{vmatrix} 1 & -2 \\ -1 & -1 \end{vmatrix}\right) + \left(1\begin{vmatrix} -1 & -1 \\ 2 & 1 \end{vmatrix} + (-1)\begin{vmatrix} 1 & -2 \\ -1 & -1 \end{vmatrix}\right)$$

$$= 6 + 2(-3+6) + 1 + 3 = 16.$$

Special case: The determinant of a (upper or lower) triangular matrix is the product of diagonal entries. In particular, the determinant of an identity matrix is 1.

Elementary operations and determinants:

Let M be an $n\times n$ square matrix.

- If N is obtained from M by interchanging two rows or two columns, then det (N) = −det (M).

Corollary: If M has two identical row (columns), then det (M) = 0.

- If N is obtained from M by multiplying a row (or a column) by a constant k, then det (N) = k det (M).

Corollery: det (kM) = k^n det (M).

If N is obtained from M by adding a multiple of a row (column) to another row, then det (N) = det (M).

If $N = M^T$, then det (N) = det (M).

Exampsle.(i) If $\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix} = 6$, find $\begin{vmatrix} a+x & b+y & c+z \\ 3x & 3y & 3z \\ -p & -q & -r \end{vmatrix}$.

$$\begin{vmatrix} a+x & b+y & c+z \\ 3x & 3y & 3z \\ -p & -q & -r \end{vmatrix} = 3\begin{vmatrix} a+x & b+y & c+z \\ x & y & z \\ -p & -q & -r \end{vmatrix} = 3\begin{vmatrix} a & b & c \\ x & y & z \\ -p & -q & -r \end{vmatrix} = -3\begin{vmatrix} a & b & c \\ x & y & z \\ p & q & r \end{vmatrix} = 3\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix} = 18.$$

(ii) Use the properties to find determinant:

$$\begin{vmatrix} 1 & 0 & -2 & 1 \\ -1 & 2 & -1 & 0 \\ 2 & -2 & 1 & -1 \\ 1 & 1 & 4 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 & -3 & 1 \\ 0 & -2 & 5 & -3 \\ 0 & 1 & 6 & -3 \end{vmatrix} (R_3 \to R_3 + R_2) = \begin{vmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 & -3 & 1 \\ 0 & 0 & 2 & -2 \\ 0 & 1 & 6 & -3 \end{vmatrix} (C_3 \to C_3 + C_4)$$

$$= \begin{vmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 1 & 3 & -3 \end{vmatrix} \text{(Expand across first row, then third column)} = 2\begin{vmatrix} 2 & -2 \\ 1 & 3 \end{vmatrix} = 16.$$

(iii) Find $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$.

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = \begin{vmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 0 & z-x & z^2-x^2 \end{vmatrix} = (y-x)(z-x)\begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & 1 & z+x \end{vmatrix} = (y-x)(z-x)\begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & 0 & z-y \end{vmatrix}$$

$= (y-x)(z-x)(z-y).$

§ 2.2 Determinants and Inverses

Product Theorem: Let M and N be $n \times n$ square matrices. Then det (MN) = (det M)(det N).

If $N = M^{-1}$, then (det N)(det M) = det I = 1. Hence, det $M \neq 0$, and det (M^{-1}) = 1 / det (M).

M has an inverse if and only if det $M \neq 0$.

det M^T = det M.

If $M^{-1} = M^T$, M is *orthogonal*. If M is orthogonal, than (det M)(det M^T) = (det M)(det M^{-1}) = det I = 1 = (det M)^2. Hence, det $M = \pm 1$.

The *cofactor matrix* of M is $[c_{ij}(M)]$.

Adjugate (adjoint) of M: is defined as adj $M = [C_{ij}(M)]^T$.

Example.

$$adj\begin{bmatrix}1 & -2 & 1\\ 2 & 1 & -1\\ 1 & 4 & -2\end{bmatrix} = \begin{bmatrix} \begin{vmatrix}1 & -1\\ 4 & -2\end{vmatrix} & -\begin{vmatrix}2 & -1\\ 1 & -2\end{vmatrix} & \begin{vmatrix}2 & 1\\ 1 & 4\end{vmatrix} \\ -\begin{vmatrix}-2 & 1\\ 4 & -2\end{vmatrix} & \begin{vmatrix}1 & 1\\ 1 & -2\end{vmatrix} & -\begin{vmatrix}1 & -2\\ 1 & 4\end{vmatrix} \\ \begin{vmatrix}-2 & 1\\ 1 & -1\end{vmatrix} & -\begin{vmatrix}1 & 1\\ 2 & -1\end{vmatrix} & \begin{vmatrix}1 & -2\\ 2 & 1\end{vmatrix} \end{bmatrix}^T = \begin{bmatrix}2 & 3 & 7\\ 0 & -3 & -6\\ 1 & 3 & 5\end{bmatrix}^T = \begin{bmatrix}2 & 0 & 1\\ 3 & -3 & 3\\ 7 & -6 & 5\end{bmatrix}$$

If det $M \neq 0$, then $M^{-1} = \dfrac{1}{\det M}\operatorname{adj}M$.

Example. $$\begin{bmatrix}1 & -2 & 1\\ 2 & 1 & -1\\ 1 & 4 & -2\end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix}2 & 0 & 1\\ 3 & -3 & 3\\ 7 & -6 & 5\end{bmatrix}.$$

Example. Find (3,2)-entry of M^{-1}, where $M=\begin{bmatrix}1 & 2 & 3 & 4\\ 0 & 2 & 3 & 3\\ 0 & 3 & 1 & 2\\ 0 & 0 & 0 & 1\end{bmatrix}$.

Solution: (3,2)-entry is $C_{23} / \det M = (-1)^{2+3} \det\begin{bmatrix}1 & 2 & 4\\ 0 & 3 & 2\\ 0 & 0 & 1\end{bmatrix} /(-7) = 3/7.$

Cramer's Rule

Consider a system $A\mathbf{x} = \mathbf{b}$, where A is an n by n matrix. If $\det A \neq 0$, then $x_i = \dfrac{\det A_i(\mathbf{b})}{\det A}$, where $A_i(\mathbf{b})$ is the matrix obtained from A by replacing the i-th column by $\mathbf{b}$.

Example. Solve the system by Cramer's rule.

$2x - 2y + z = 1,$
$x + y - 2z = -3,$
$3x - y + 2z = 2.$

Solution: (i) Find the determinant of the coefficient matrix:

$$\det\begin{bmatrix} 2 & -2 & 1 \\ 1 & 1 & -2 \\ 3 & -1 & 2 \end{bmatrix} (C_2 \to C_2 + C_1) = \det\begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & -2 \\ 3 & 2 & 2 \end{bmatrix} (R_3 \to R_3 - R_2) = \det\begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & -2 \\ 2 & 0 & 4 \end{bmatrix}$$
$$= 2\begin{vmatrix} 2 & 1 \\ 2 & 4 \end{vmatrix} = 12.$$

(ii) Find $A_1(\mathbf{b})$, $A_2(\mathbf{b})$, $A_3(\mathbf{b})$:

$$\det A_1(\mathbf{b}) = \det\begin{bmatrix} 1 & -2 & 1 \\ -3 & 1 & -2 \\ 2 & -1 & 2 \end{bmatrix} (C_3 \to C_3 - C_1) = \det\begin{bmatrix} 1 & -2 & 0 \\ -3 & 1 & 1 \\ 2 & -1 & 0 \end{bmatrix} = -\begin{vmatrix} 1 & -2 \\ 2 & -1 \end{vmatrix} = -3.$$

$$\det A_2(\mathbf{b}) = \det\begin{bmatrix} 2 & 1 & 1 \\ 1 & -3 & -2 \\ 3 & 2 & 2 \end{bmatrix} (C_3 \to C_3 - C_2) = \det\begin{bmatrix} 2 & 1 & 0 \\ 1 & -3 & 1 \\ 3 & 2 & 0 \end{bmatrix} = -\begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = -1.$$

$$\det A_3(\mathbf{b}) = \det\begin{bmatrix} 2 & -2 & 1 \\ 1 & 1 & -3 \\ 3 & -1 & 2 \end{bmatrix} (C_2 \to C_2 + C_1) = \det\begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & -3 \\ 3 & 2 & 2 \end{bmatrix} (R_3 \to R_3 - R_2)$$
$$= \det\begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & -3 \\ 2 & 0 & 5 \end{bmatrix} = 2\begin{vmatrix} 2 & 1 \\ 2 & 5 \end{vmatrix} = 16.$$

Therefore, $x = -3 / 12 = -1 / 4$, $y = -1 / 12$, and $z = 16 / 12 = 4 / 3$.

§ 2.3 Diagonalization and Eigenvalues

(1) Motivation*:* Let M be an n by n matrix. Sometimes we need to calculate M^k, k-th power of M.

(2) Eigenvalues and Eigenvectors

Let $T\mathbf{v} = M\mathbf{v}$ be a linear transformation from R^n to R^n. We want $M\mathbf{v} = \lambda\mathbf{v}$ for some λ and $\mathbf{v} \neq \mathbf{0}$. Then $(I\lambda - M)\mathbf{v} = \mathbf{0}$, and $\det(I\lambda - M) = 0$. Polynomial $\det(I\lambda - M)$ is called the characteristic polynomial, denoted by $c_M(\lambda)$. The equation

$$\det(I\lambda - M) = 0$$

is called the *characteristic equation* of M. λ is an *eigenvalue* of M and $\mathbf{v}$ is a corresponding *eigenvector, called* λ*–eigenvector*. All eigenvectors form the *eigenspace*.

Examples. Finding Eigenvalues and Eigenvectors

(i) $M = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$. $\lambda I - M = \begin{bmatrix} \lambda-1 & -2 \\ -4 & \lambda-3 \end{bmatrix}$, $\det(\lambda I - M) = (\lambda - 1)(\lambda - 3) - 8 = \lambda^2 - 4\lambda - 5 = 0$.

$\lambda = 5$, $\lambda = -1$.

When $\lambda = 5$, $5I - M = \begin{bmatrix} 4 & -2 \\ -4 & 2 \end{bmatrix}$. An eigenvector is $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, which is a basic eigenvector.

When $\lambda = -1$, $-I - M = \begin{bmatrix} -2 & -2 \\ -4 & -4 \end{bmatrix}$. An eigenvector is $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(ii) $M = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. $\lambda I - M = \begin{bmatrix} \lambda-1 & 1 & -2 \\ 2 & \lambda-1 & 0 \\ -1 & 0 & \lambda-1 \end{bmatrix}$.

$\det(\lambda I - M) = (\lambda - 1)^3 - 4(\lambda - 1) = (\lambda - 1)(\lambda^2 - 2\lambda - 3) = 0$, $\lambda = 1, -1, 3$.

When $\lambda = 1$, $I - M = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$.

When $\lambda = -1$, $-I - M = \begin{bmatrix} -2 & 1 & -2 \\ 2 & -2 & 0 \\ -1 & 0 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ -2 & 1 & -2 \\ -1 & 0 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

When $\lambda = 3$, $-I - M = \begin{bmatrix} 2 & 1 & -2 \\ 2 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & -2 \\ -1 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

(iii) $M = \begin{bmatrix} 7 & 1 & -2 \\ -3 & 3 & 6 \\ 2 & 2 & 2 \end{bmatrix}$. $\lambda I - M = \begin{bmatrix} \lambda-7 & -1 & 2 \\ 3 & \lambda-3 & -6 \\ -2 & -2 & \lambda-2 \end{bmatrix}$.

det $(\lambda I - M) = (\lambda-7)\begin{vmatrix} \lambda-3 & -6 \\ -2 & \lambda-2 \end{vmatrix} + \begin{vmatrix} 3 & -6 \\ -2 & \lambda-2 \end{vmatrix} + 2\begin{vmatrix} 3 & \lambda-3 \\ -2 & -2 \end{vmatrix}$

$= (\lambda - 7)(\lambda - 3)(\lambda - 2) - 12(\lambda - 7) + 3(\lambda - 2) - 12 - 12 + 4(\lambda - 3) = \lambda^3 - 12\lambda^2 + 36\lambda.$

$\lambda = 0, 6, 6.$

When $\lambda = 0$, $-M = \begin{bmatrix} -7 & -1 & 2 \\ 3 & -3 & -6 \\ -2 & -2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 3 & -3 & -6 \\ -7 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & -6 & -9 \\ 0 & 6 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3/2 \\ 0 & 0 & 0 \end{bmatrix}$.

When $\lambda = 6$, $6I - M = \begin{bmatrix} -1 & -1 & 2 \\ 3 & 3 & -6 \\ -2 & -2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -2 \\ 3 & 3 & -6 \\ -1 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

(iv) $M = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$. $\lambda I - M = \begin{bmatrix} \lambda-2 & -4 & -3 \\ 4 & \lambda+6 & 3 \\ -3 & -3 & \lambda-1 \end{bmatrix}$.

det $(\lambda I - M) = (\lambda-2)\begin{vmatrix} \lambda+6 & 3 \\ -3 & \lambda-1 \end{vmatrix} + 4\begin{vmatrix} 4 & 3 \\ -3 & \lambda-1 \end{vmatrix} - 3\begin{vmatrix} 4 & \lambda+6 \\ -3 & -3 \end{vmatrix}$

$= (\lambda - 2)(\lambda + 6)(\lambda - 1) + 9(\lambda - 2) + 16(\lambda - 1) + 36 - 36 - 9(\lambda + 6) = \lambda^3 + 3\lambda^2 - 4 = 0.$

$\lambda = 1, -2, -2.$

When $\lambda = 1$, $I - M = \begin{bmatrix} -1 & -4 & -3 \\ 4 & 7 & 3 \\ -3 & -3 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 4 & 3 \\ 0 & -9 & -9 \\ 0 & 9 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

When $\lambda = -2$, $-2I - M = \begin{bmatrix} -4 & -4 & -3 \\ 4 & 4 & 3 \\ -3 & -3 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 4 & 4 & 3 \\ -4 & -4 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

Geometric interpretation of eigenvectors in R^2:

Let **v** be a non-zero vector in R^2, and let $L_\mathbf{v}$ be the (unique) line that goes through the origin and contains **v**. $L_\mathbf{v} = \{t\mathbf{v};\ t$ is a real number$\}$.

Let A be a 2×2 matrix. $L_\mathbf{v}$ is said to be A-*invariant* if **x** is in $L_\mathbf{v}$ implies that $A\mathbf{x}$ is $L_\mathbf{v}$. In other words, $\mathbf{x} = t_1\mathbf{v}$ implies $A\mathbf{x} = t_2\mathbf{v}$. When $t_1 \neq 0$, $A(t_1\mathbf{v}) = t_2\mathbf{v}$, $A\mathbf{v} = (t_2 / t_1)\mathbf{v}$. Vector **v** is an eigenvector of A.

On the other hand, if vector **v** is an eigenvector of A, the $L_\mathbf{v}$ is A-invariant.

Reflection about line $y = mx$ has this line as invariant line. Every vector in this line is an eigenvector of its matrix. Rotation through an angle that is not a multiple of π has no invariant line. Hence its matrix has no eigenvectors.

(3) Diagonalization

Motivation*:* Let A be an n by n matrix. Sometimes we need to calculate A^k, k-th power of A.

Diagonal matrix: $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ = diag ($\lambda_1, \lambda_2, \ldots, \lambda_n$). All entries off the main diagonal are zero.

The sum and product of diagonal matrices are still diagonal matrices.

Definition. Let A be an $n \times n$ matrix. If there is an invertible matrix P such that $D = P^{-1}AP$ is a diagonal matrix, then A is said to be diagonalizable, and P is a diagonalizing matrix of A.

If A is diagonalizable, then AP=PD. Write $P = [C_1\ \ C_2 \ldots C_n] = (a_{ij})$, an $n \times n$ matrix, and $D = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, then

$$PD = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} = \begin{bmatrix} a_{11}\lambda_1 & a_{12}\lambda_2 & \ldots & a_{1n}\lambda_n \\ a_{21}\lambda_1 & a_{22}\lambda_2 & \ldots & a_{2n}\lambda_n \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1}\lambda_1 & a_{n2}\lambda_2 & \ldots & a_{nn}\lambda_n \end{bmatrix}$$

$= (\lambda_1 C_1 \;\; \lambda_2 C_2 \;\; \ldots \;\; \lambda_n C_n)$.

Then $AC_i = \lambda_i C_i$. In other words, λ_i is an eigenvalue of A and C_i is an eigenvector of A.

Properties. Let A be an $n \times n$ matrix.

(1) A is diagonalizable if and only if we can find n eigenvectors $C_1, C_2, \ldots, C_n$ such that $P = [C_1 \;\; C_2 \ldots C_n]$ is invertible.
(2) If A has n distinct eigenvalues, then A is diagonalizable.
(3) If $(\lambda - \lambda_i)^k$ is a factor of the characteristic equation, λ_i is said to have *multiplicity* k. If, every eigenvalue with multiplicity k has k basic eigenvectors, then P is invertible.

Diagonalization algorithm: To diagonalize an $n \times n$ matrix A:

Step 1: Find the distinct eigenvalues of A.
Step 2: For each eigenvalue λ, compute basic eigenvectors from basic solutions of the homogeneous system $(\lambda I - A)X = 0$.
Step 3: A is diagonalizable if and only if there are n basic eigenvectors in all.
Step 4: If A is diagonalizable, then the matrix P with these n basic eigenvectors as columns is a diagonalizing matrix for A.

Examples.

(i) Diagonalize $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.

Solution: $\lambda I - A = \begin{bmatrix} \lambda - 1 & -2 \\ -4 & \lambda - 3 \end{bmatrix}$, $\det(\lambda I - A) = (\lambda - 1)(\lambda - 3) - 8 = \lambda^2 - 4\lambda - 5 = 0$.

$\lambda = 5,\ \lambda = -1$.

When $\lambda = 5$, $5I - A = \begin{bmatrix} 4 & -2 \\ -4 & 2 \end{bmatrix}$. An eigenvector is $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, which is a basic eigenvector.

When $\lambda = -1$, $-I - A = \begin{bmatrix} -2 & -2 \\ -4 & -4 \end{bmatrix}$. An eigenvector is $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

$P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$, $P^{-1} = -\frac{1}{3}\begin{bmatrix} -1 & -1 \\ -2 & 1 \end{bmatrix}$. $P^{-1}AP = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$.

(ii) $A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

$\det(\lambda I - A) = (\lambda - 1)^3 - 4(\lambda - 1) = (\lambda - 1)(\lambda^2 - 2\lambda - 3) = 0$, $\lambda = 1, -1, 3$.

When $\lambda = 1$, $I - A = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$.

When $\lambda = -1$, $-I - A = \begin{bmatrix} -2 & 1 & -2 \\ 2 & -2 & 0 \\ -1 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 0 \\ -2 & 1 & -2 \\ -1 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

When $\lambda = 3$, $-I - A = \begin{bmatrix} 2 & 1 & -2 \\ 2 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & -2 \\ -1 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

$P = \begin{bmatrix} 0 & -2 & 2 \\ 2 & -2 & -2 \\ 1 & 1 & 1 \end{bmatrix}$. $P^{-1} = \frac{1}{8}\begin{bmatrix} 0 & 2 & 4 \\ 2 & 1 & -2 \\ 6 & -3 & 6 \end{bmatrix}$. $P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

(iii) $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Solution. $\det(\lambda I - A) = \lambda(\lambda - 1)^2 = 0$, $\lambda = 0, 1, 1$.

When $\lambda = 0$, $0I - A = \begin{bmatrix} 0 & -1 & -2 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. When $\lambda = 1$, $I - A = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

P= $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Remark. This A is not invertible.

(iv) A= $\begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$ is not diagonalizable. $\lambda = 1, -2, -2$. When $\lambda = -2$, it has only one basic eigenvector, e.g., $[-1\ 1\ 0]^T$.

Remark. This A is invertible.

Finding Power of a Matrix

$P^{-1}AP = D$. $(P^{-1}AP)^k = P^{-1}A^kP = D^k$. $A^k = PD^kP^{-1}$.

Example. $A = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}$. Eigenvalues: $\lambda = 1, 0.7$, eigenvector of $\lambda = 1$ is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, eigenvector of $\lambda = 0.7$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Let $P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$. $P^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$. $P^{-1}AP = D = \begin{bmatrix} 1 & 0 \\ 0 & 0.7 \end{bmatrix}$.

$A^k = PD^kP^{-1} = \begin{bmatrix} -1+2(0.7^k) & 1-0.7^k \\ -2+2(0.7^k) & 2-0.7^k \end{bmatrix}$.

When $k \to \infty$, $M^k \to \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}$.

(4) Similar matrices

A and *B* are similar if there exists an invertible P such that $B = P^{-1}AP$. *We write A~B.*

§ 4.9 and 5.3 Linear Transformations

Transformations $\mathbf{R}^n \rightarrow \mathbf{R}^m$

For every vector **v** in R^n, a *transformation T* assigns a unique vector **w** in R^m, to be the *image* of **v**, denoted by **w** = T(**v**), while **v** is called the *pre-image* of **w**. We denote this by T: $R^n \rightarrow R^m$.

Examples

(i) For every vector **v** in R^n, T(**v**) = **v**. This is the *identity transformation*.

(ii) For every vector **v** in R^n, T(**v**) = **0**. The *zero transformation*.

Matrix transformation: Let A be a $m \times n$ matrix. For every vector **v** in R^n, T(**v**) = A**v** is a matrix transformation *induced by A*. The matrix A is called the standard matrix of T.

Examples of $R^2 \rightarrow R^2$:

Reflection about y-axis: $[a\ \ b]^T \rightarrow [-a\ \ b]^T := \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$.

Reflection about x-axis: $[a\ \ b]^T \rightarrow [a\ \ -b]^T := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$.

Reflection about line $y = x$: $[a\ \ b]^T \text{ -> } [b\ \ a]^T := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$.

Reflection about line $y = -x$: $[a\ \ b]^T \text{ -> } [-b\ \ -a]^T = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$.

Rotation counterclockwise about the origin through $\pi / 2$:

$$R_{\pi/2}\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -b \\ a \end{bmatrix}, \text{ or } R_{\pi/2}: (a, b)^T \text{ -> } (-b, a)^T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}.$$

Rotation counterclockwise about the origin through π: $(a, b)^T \text{ -> } (-a, -b)^T = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$.

Rotation counterclockwise about the origin through $3\pi / 2$: $(a, b)^T \text{ -> } (b, -a)^T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$.

Linear transformation: A transformation $T : R^n \to R^m$ is a *linear transformation* if

(i) $T(c\mathbf{v}) = cT(\mathbf{v})$ for every number c and every vector $\mathbf{v} \in R^n$, and

(ii) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all vectors $\mathbf{u}, \mathbf{v} \in R^n$.

Examples

(i) $n = m = 2$, $T([a \ \ b]^T) = [b \ \ a]^T$.

(ii) $n = 2$, $m = 3$, $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-2y \\ y \end{bmatrix}$.

(iii) Every matrix transformation is linear.

(iv) Every linear transformation from R^n to R^m is a matrix transformation.

Indeed, let T be a linear transformation. Let

$$I_n = [E_1 \ \ E_2 \ \dots \ E_n],$$

where $\{ E_1, \dots, E_n\}$ is called the standard basis of R^n.
Denote $TE_i = \mathbf{c}_i$, $i = 1, 2, \dots, n$, and let

$$M = [\mathbf{TE}_1 \ \mathbf{TE}_2 \ \dots \ \mathbf{TE}_n].$$

Then for any $\mathbf{x} = [x_1 \ \ x_2 \ \dots \ x_n]^T$, since T is linear, $T\mathbf{x} = \mathbf{c}_1 x_1 + \mathbf{c}_2 x_2 + \dots + \mathbf{c}_n x_n = M\mathbf{x}$.

(v) If U is a subspace, then **projection operator** $T(X) = \text{proj}_U(X): R^n \to R^n$ is a linear transformation.

In fact, let $\{F_1, F_2, \dots, F_k\}$ be an orthogonal basis of U. Then

$$T(X) = \text{proj}_U(X) = (X\cdot F_1 / \| F_1 \|^2)F_1 + (X\cdot F_2 / \| F_2 \|^2)F_2 + \dots + (X\cdot F_k / \| F_k \|^2)F_k$$

Properties

(i) $T(\mathbf{0}) = \mathbf{0}$.

(ii) $T(-\mathbf{v}) = -T(\mathbf{v})$.

(iii) $T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_k\mathbf{v}_k) = a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \dots + a_kT(\mathbf{v}_k)$.

(iv) Let $\{F_1, F_2, \dots, F_n\}$ be any basis of R^n, and let $\{Z_1, Z_2, \dots, Z_n\}$ be any n vectors in R^m. Then there is unique linear transformation T from R^n to R^m such that $T(F_i) = Z_i$.

Proof. For any X in R^n , X = $a_1F_1 + a_2F_2 + \ldots + a_n F_n$. Define T(X):

$T(X) = a_1Z_1 + a_2Z_2 + \ldots + a_n Z_n$.

Example. $T([a\ \ b]^T) = [a+1\ \ b]^T$ is not a linear transformation.

Example. Find the matrix M that induces the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-2y \\ y \end{bmatrix}$.

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, M = \begin{bmatrix} 1 & 1 \\ 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

Composition of matrix transformations:

Let $T_1(\mathbf{v}) = M_1\mathbf{v}$ be a matrix transformation $R^n \rightarrow R^m$, $T_2(\mathbf{v}) = M_2\mathbf{v}$ be a matrix transformation $R^m \rightarrow R^k$. Then $T(\mathbf{v}) = T_2(T_1(\mathbf{v})) = M_2M_1\mathbf{v}$ is a matrix transformation R^n -> R^k, called the *composition* of T_2 and T_1, can be written as $T_2 \circ T_1$.

Example. T_1 is the reflection about the y-axis. $T_1(\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = (a, -b)$. T_2 is the rotation counterclockwise about the origin through $\pi/2$: $T_2(\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = (-b, a)$. Then

$$T_1 \circ T_2(\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = (-b, -a).$$

$$T_2 \circ T_1 (\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = (b, a).$$

Inverse of matrix transformations:

If $T\mathbf{x} = M\mathbf{x}$ is a matrix transformation, and M is invertible, then $T^{-1}\mathbf{x} = M^{-1}\mathbf{x}$ is the inverse of T.

Example. Let T be the reflection about the y-axis. $T(\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = (a+2b, -b)$. Then

$$T^{-1} (\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}^{-1}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = (a\text{-}2b, -b).$$

Linear Transformation on Vector Space

Let V and W be vector spaces. Then a *linear transformation* from V to W is a function with domain V and range a subset of W satisfying

1) T(**u** + **v**) = L(**u**) + L(**v**) 2) T(c**u**) = cL(**u**)

for any vectors **u** and **v** in V and scalar c.

Example. Let V be the vector space of (infinitely) differentiable functions and define D to be the function from V to V given by

D(f(t)) = f '(t)

Then D is a linear transformation since

D(f(t) + g(t)) = (f(t) + g(t))' = f '(t) + g'(t) = D(f(t)) + D(g(t))

and

D(cf(t)) = (cf(t))' = c f '(t) = cD(f(t))

Example. Define T from $M_{2,2}$ to P_3 as follows:

$$T(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = ax^3 + bx^2 + cx + d\ .$$

Then T is a linear transformation

Example. Let V = P_2 and let W be the real numbers. Then T from V to W defined by

$L(at^2 + bt + c) = ab+c$

is not a linear transformation.

Properties: Assume that T is a linear transformation from a vector space V to a vector space W, and **u**, **v**, $\mathbf{v_1}$, $\mathbf{v_2}$, ... ,$\mathbf{v_n}$ are vectors in V. Then

1. L(0) = 0

2. L(**u** - **v**) = L(**u**) - L(**v**)

3. $L(c_1\mathbf{v_1} + c_2\mathbf{v_2} + ... + c_n\mathbf{v_n}) = c_1L(\mathbf{v_1}) + c_2L(\mathbf{v_2}) + ... + c_nL(\mathbf{v_n})$

4. Let S = $\{\mathbf{v_1}, \mathbf{v_2}, ... ,\mathbf{v_n}\}$ be a basis for V. Then T is completely determined by the image of the basis S.

Example. Let T be the linear transformation from P_1 to $M_{2,2}$ such that

$$T(1+t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ T(1-t) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Find L(3 + t).

Hint: Write 3+t = 2(1+t)+(1-t)

Kernel and Image Spaces

If $f : V \rightarrow W$ is linear, we define the kernel and the image or range of f by

$$\ker(f) = \{ x \in V : f(x) = 0 \}$$

$$\operatorname{im}(f) = \{ w \in W : w = f(x), x \in V \}$$

ker(f) is a subspace of V and im(f) is a subspace of W.

Rank-nullity theorem:

$$\dim(\ker(f)) + \dim(\operatorname{im}(f)) = \dim(V).$$

The number dim(im(f)) is also called the *rank of f* and written as rank(f); the number dim(ker(f)) is called the *nullity of f* and written as null(f).

If V and W are finite-dimensional, bases have been chosen and f is represented by the matrix A, then the rank and nullity of f are equal to the rank and nullity of the matrix A, respectively.