Groups and Linear Algebra

Mahipal Jadeja

Jaideep Mulherkar

Rishikant Rajdeep

E-mail address: jaideep_mulherkar@daiict.ac.in

2000 *Mathematics Subject Classification.* Primary
Key words and phrases. Linear Algebra, Group theory

Abstract.

Contents

List of Figures

CHAPTER 1

Groups

1.1. Symmetries

Consider the rotational symmetries of a regular tetrahedron. One can rotate about the axis L passing through the vertex 1 and the centre of the face determined by vertices $2, 3, 4$ (see figure 1) by angle $2\pi/3$ by a further angle $4\pi/3$. Rotating by another $2\pi/3$ totalling a rotation of 2π brings all the vertices back to their original positions and we call this the identity symmetry. There are four axis of the type L and hence there are 4×2 symmetries of this type. Another symmetry is by a rotation of angle π about axis M which passes through the mid point of side determined by vertices $1, 4$ and $2, 3$. Applying this symmetry twice we get to the identity. There are 3 symmetries of type M. Thus, along with the identity there are 12 rotational symmetries of a regular tetrahedron. Let r denote the rotation

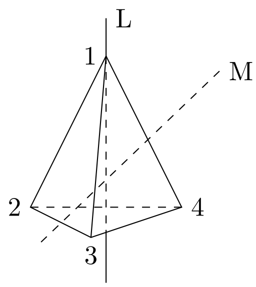

FIGURE 1. Axis of symmetries of a regular tetrahedron

(anti-clockwise) by an angle of $2\pi/3$ about the axis L and s be the rotation (anti-clockwise) by an angle of π about the axis M. Then $s^2 = 1$ and $r^3 = 1$. The rotational symmetries r and s permute the vertices as follows :

$$r : 1 \to 1, 2 \to 3, 3 \to 4, 4 \to 1$$

and

$$s\ :\ 1 \to 4, 2 \to 3, 3 \to 2, 4 \to 1$$

Let us see what happens when we apply r followed by s (figure 2) :

In general, we observe that

i Combining two rotations u and v gives another rotation w.

ii We also observe that the rotations u, v and w satisfy $u{*}(v{*}w) = (u{*}v){*}w$. This property is known as associativity.

iii for each rotation u there exists a rotation u^{-1} such that $u{*}u^{-1} = u^{-1}{*}u = e$.

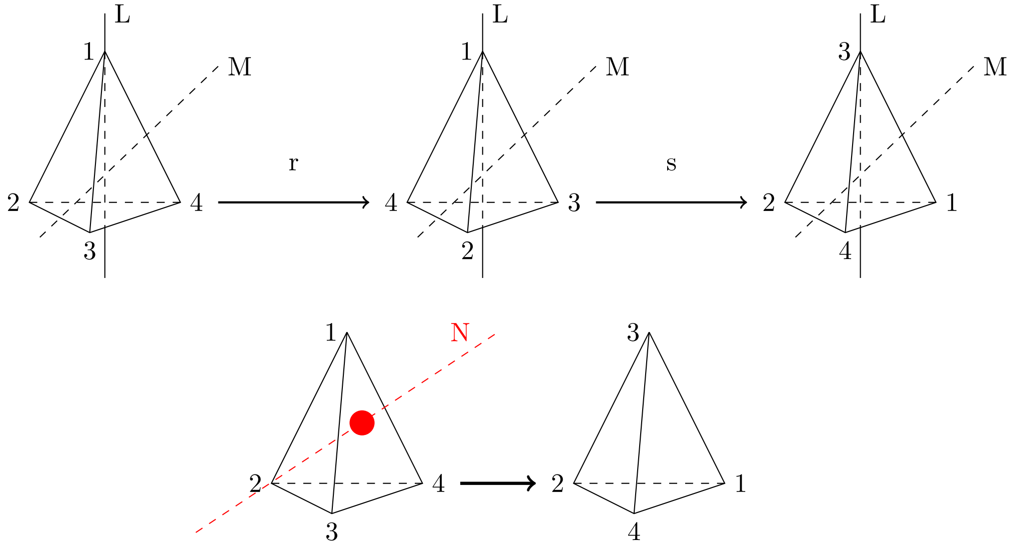

FIGURE 2. Applying the rotation r followed by rotation s results in the rotation about axis N as shown

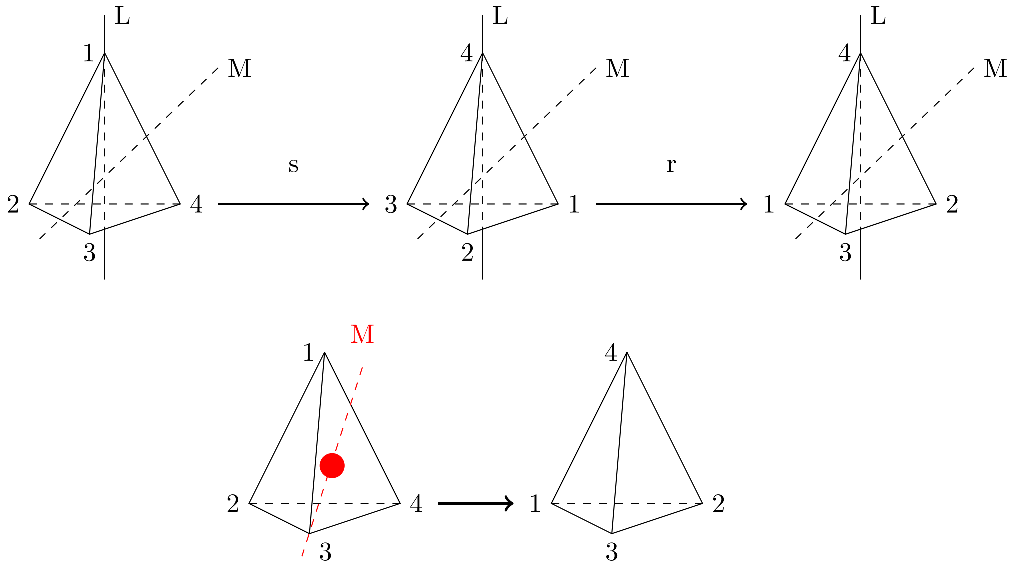

FIGURE 3. Rotation s followed r results in rotation about axis M as shown

If one applies rotation s followed by r then we get a rotation (clockwise) about the axis N as shown in figure 3 : Thus, rs is not equal to sr in general. The algebraic structure formed by the rotations of the regular tetrahedron is called a group. All the properties marked by $*$ in the above discussion will form the axioms of an algebraic structure called group.

1.2. Groups

DEFINITION 1.1. A non-empty set G along with a closed binary operation $*: G \times G \to G$ is called a group if

(1) $\forall u, v, w \in G$ the associativity property holds.
i.e. $\forall u, v, w \in G, u * (v * w) = (u * v) * w$.
(2) $\exists$ an element e called the identity such that, $u*e = e*u = u$ for all $u \in G$.
(3) For all $u \in G$, $\exists$ an element u^{-1} (called inverse of u) such that $u * u^{-1} = u^{-1} * u = e$.

We note that in a group the identity element and the inverse element are unique.

THEOREM 1.2. *In a group G the identity and inverse are unique.*

PROOF. If e and e' are both the identity elements of the group G then $e * e' = e = e'$.
Now if $x \in G$ and if y and z are inverses of x then $y = e*y = (z*x)*y = z*(x*y) = z * e = z$. □

1.2.1. Examples of groups.

(1) $(\mathbb{Z}, +), (\mathbb{R}, +), (\mathbb{C}, +)$: Integers, Real numbers, Complex numbers under addition.
(2) $(\mathbb{Z}^n, +), (\mathbb{R}^n, +), (\mathbb{C}^n, +)$: are groups under component wise addition.
(3) $M_n(\mathbb{R}), M_n(\mathbb{C})$: $n \times n$ matrices with entries in the real numbers (complex numbers) under matrix addition.
(4) $GL_n(\mathbb{R}), GL_n(\mathbb{C})$: The set of invertible matrices with entries in real (complex) numbers under matrix multiplication.
(5) S_n: Let $A = \{1, 2, ..., n\}$ and consider the set S_n of one-one and onto mappings from the set A onto A. The set S_n has $n!$ elements and each element of S_n is a permutation (one-one and onto mapping) on n-letters. S_n is a group (check!) under the operation function composition.
(6) The group of rotational symmetries of a regular tetrahedron as was studied in the previous chapter.
(7) Let $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$. Let us define an operation $\oplus_n$ (addition modulo n) on $\mathbb{Z}_n$

$$x \oplus_n y = x + y \qquad \text{if} \quad x + y < n$$
$$x \oplus_n y = x + y - n \qquad \text{if} \quad x + y \geq n$$

Verify that this is a group! What is the inverse of x?
(8) Is $(\mathbb{R}, \times)$, the set of real numbers a group under multiplication?
No! $x \times 0 = 0$. So 0 does not have an inverse! But if we exclude 0 then everything looks ok! So, $\mathbb{R} - \{0\}$ is group under multiplication.
(9) $\{z \in \mathbb{C} : z^n = 1\}$, the n^{th} roots of unity. Check that this is a group under multiplication.
(10) Consider ${\mathbb{Z}_n}^* = Z_n - \{0\} = \{1, 2, \ldots, n-1\}$ and operation

$$x \times_n y \quad = \quad x \cdot y \mod n$$

Is this a group?
Consider, $\mathbb{Z}_6^* = \{1, 2, 3, 4, 5\}$ under multiplication modulo 6. We observe that $2 \times_6 3 = 2 \cdot 3 \mod 6 = 0$, which does not belong to $\mathbb{Z}_6^*$. So $(\mathbb{Z}_6^*, \times_6)$ is not a group since it does not have closure property.

(11) What about ${\mathbb{Z}_7}^* = \{1, 2, 3, 4, 5, 6\}$. You can verify that this indeed is a group. Do you see any pattern? We will see later that if p is a prime then indeed ${\mathbb{Z}_p}^*$ is a group. In the previous example, we saw that by removing the element 0 we were able to form a group $\mathbb{R} - \{0\}$ under multiplication. Can something be done about ${\mathbb{Z}_n}^*$?

(12) Finally consider the group of symmetries of a regular polygon D_n. Shown in the figure are the various axis of symmetry for a triangle. There are six symmetries of a triangle and they form a group under composition.

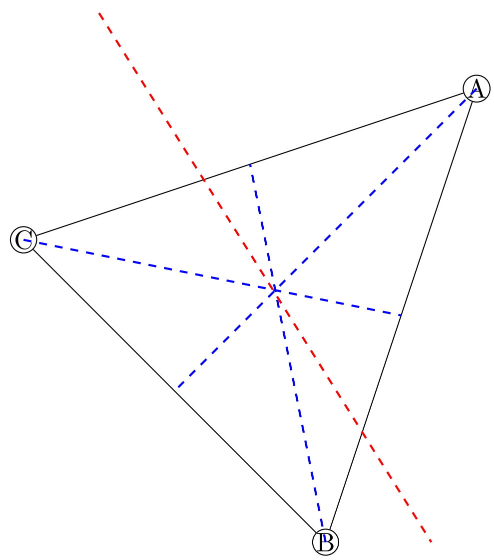

FIGURE 4. Symmetry for a triangle

1.3. Dihedreal group

The group of symmetries of a regular n-sided polygon is called dihedral group, denoted by D_n (some books denote it by D_{2n}). Consider first the equilateral triangle as shown in the figure.

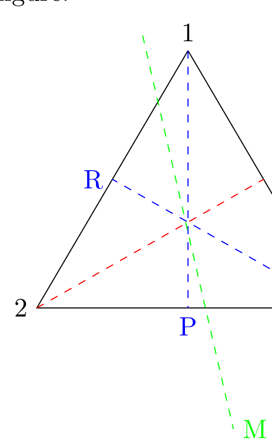

FIGURE 5. Symmetries of equilateral triangle

The axes of symmetries thes line M (through the paper) ,N,R and P. Let r denote the rotation about the axis M by $2\pi/3$. This takes the vertex 1 to 2, 2 to 3 and 3 back to 1. Let s denote the reflection about the axis P. s interchanges the vertices 2 and 3. We have the relations $r^3 = e$ and $s^2 = e$. Now rs (s followed by r) is the reflection about the axis R.

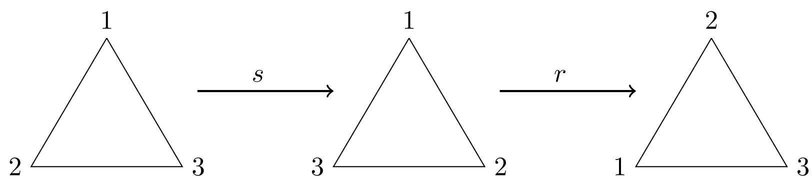

FIGURE 6. Symmetries of a regular triangle

Similarly r^2s is the reflection about the axis N. In this way we get all the symmetries of the equilateral triangle which is the set $\{e, r, r^2, s, rs, r^2s\}$. What about the element sr ? It turns out that $sr = r^2s$. This can be seen geometrically as shown in the figure. Then using this fact we can show that $sr^2 = rs$. Indeed, $sr^3 = (sr)r = (r^2s)r = r^2(sr) = r^2r^2s = r^4s = rs$, since $r^3 = e$.

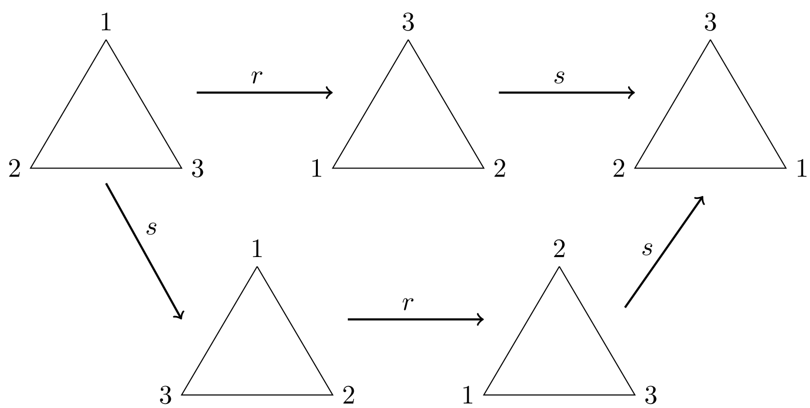

FIGURE 7. $sr = r^2s$

The corresponding group table (multiplication table) of the above dihedral group is shown below :

	e	r	r^2	s	rs	r^2s
e	e	r	r^2	s	rs	r^2s
r	r	r^2	e	rs	r^2s	s
r^2	r^2	e	r	r^2s	s	rs
s	s	r^2s	rs	e	r^2	r
rs	rs	s	r^2s	r	e	r^2
r^2s	r^2s	rs	s	r^2	r	e

For a general n-regular polygon we can generate the set : let r be a rotation by an angle $\frac{2\pi}{n}$ by the axis of symmetry that is perpendicular to the plane in which the regular n-gon lies and s be the reflection about a line that lies in the plane (

it does not matter which one), then again we get $r^n = e$ and $s^2 = e$. There are $2n$ symmetries given by $\{e, r, r^2, \ldots, r^{n-1}, s, rs, r^2s, \ldots, r^{n-1}s\}$. We also have that $sr = r^{n-1}s = r^{-1}s$ (check geometrically !). Using this fact we get

$$sr^j = r^{n-j}s \text{ (check!)}.$$

Each element of the group D_n has elements of the form r^a or r^as for $0 \leq a \leq n-1$ and we have

$$r^ar^b = r^k, \quad k = a +_n b \tag{1.1}$$

$$r^a(r^bs) = r^ks, \quad k = a +_n b \tag{1.2}$$

$$(r^as)r^b = r^ls, \quad l = a +_n (n-b) \tag{1.3}$$

$$(r^as)(r^bs) = r^l, \quad l = a +_n (n-b) \tag{1.4}$$

We say that r and s generate the group D_n.

Finally, the order of a group is the number of elements in the group. If a group has infinite elements then the group has infinite order. We denote the order of a group by $|G|$. If x is an element of G and if there is a positive integer such that $x^m = e$ then we say that x has finite order. The smallest positive integer m such that $x^m = e$ is called the order of x.

Examples :

(1) The order of D_n is $2n$. In D_3, r and r^2 have order 3, whereas s, rs and r^2s have order 2.
(2) Order of $\mathbb{Z}_6$ is six. The elements 1 and 5 have order 6, 2 and 4 have order 3 and 3 has order 2.
(3) $(\mathbb{R}, +)$ has infinite order and every element except 0 has infinite order.
(4) $C = \{z \in \mathbb{C} : |z| = 1\}$ is the unit circle. This is a group of infite order. All element of this group are of the form $e^{i\theta}$. The elements whose θ is a rational multiple of 2π have finite order.

1.4. Subgroups, cyclic groups, generators

Consider the following subset of the group D_6, $\{e, r^2, r^4, s, r^2s, r^4s\}$

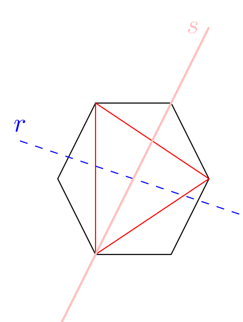

FIGURE 8. D_3 as subgroup of D_6

Notice that if we multiply any of these elements with each other we get another element within the set. The set is also closed under taking inverses. In other word, it is a subset of D_6 which is a group by itself. In fact, this group is the group of symmetries of the triangle, D_3, within the hexagon as shown in the figure 8.

DEFINITION 1.3. Let G be a group and $H \subset G$, then H is a subgroup of G (denoted by $H < G$) if

(i) $e \in H$
(ii) If $x, y \in H$ then $x * y \in H$
(iii If $x \in H$ then $x^{-1} \in H$

EXAMPLE 1.4. Examples of subgroups :

(1) $\mathbb{Z} < \mathbb{Q} < \mathbb{R}$ under addition.
(2) $\{e, r, r^2, \ldots, r^{n-1}\}$ is a subgroup of D_n.
(3) The set of diagonal matrices with non-zero entries is a subgroup of $GL_n(\mathbb{R})$.
(4) In $(\mathbb{Z}_6, +_6)$ the set $\{0, 2, 4\}$ is a subgroup.
(5) $\{\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : ac \neq 0\}$ is a subgroup of $GL_2(\mathbb{R})$.
(6) Let G be a group and let $x \in G$ then the set $\langle x \rangle = \{x^m : m \in \mathbb{Z}\}$ is a subgroup of G. The subgroup $\langle x \rangle$ is called the subgroup generated by x. We have $x^0 = e$, x^{-m} is the inverse of x^m and $x^k * x^p = x^{k+p}$. Recall that the order of an element $x \in G$ is the smallest positive integer m such that $x^m = e$. If $\langle x \rangle$ has infinite order then $\langle x \rangle$ consists of elements $\{\ldots, x^{-2}, x^{-1}, e, x, x^2, \ldots\}$. If $\langle x \rangle$ has finite order n then $\langle x \rangle$ has elements $e, x, x^2, \ldots, x^{n-1}$.

DEFINITION 1.5. A group G is called a cyclic group if there exists an element $a \in G$ such that $\langle a \rangle = G$. (The element a will be called the generator of the group G.)

EXAMPLE 1.6. (1) $\mathbb{Z}$ is an infinite cyclic group. Its generators are 1 and -1.
(2) $\mathbb{Z}_6$ is generated by 1 and 5, i.e., $\langle 1 \rangle = \langle 5 \rangle = \mathbb{Z}_6$. The subgroup generated by 2 is $\{0, 2, 4\}$.
(3) In D_3 we have,
- $\langle e \rangle = \{e\}$
- $\langle r \rangle = \langle r^2 \rangle = \{e, r, r^2\}$.
- $\langle s \rangle = \{e, s\}$.
- $\langle rs \rangle = \{e, rs\}$
- $\langle r^2 s \rangle = \{e, r^2 s\}$

We see that D_n is not cyclic but each of its elements can be written in terms of r and s, hence r and s together generate D_n. If X is a subset of a group G then a word in the elements of X is of the form $x_1^{m_1} x_2^{m_2} \ldots x_k^{m_k}$ where each $x_i \in X$ and $m_i \in \mathbb{Z}$. The collection of all words is a subgroup of G. (Check !). This group is called the subgroup generated by X. If this is the entire group G then the set X is called the generators of G. The set $\{r, s\}$ is the set of generators of D_6, and so is the set $\{s, rs\}$ (since $rs * s = r$, so any word using r and s can be converted to a word using rs and s.

THEOREM 1.7. *Let H be a non-empty subset of a group G then H is a subgroup of G iff xy^{-1} belongs to H whenever $x, y \in H$.*

PROOF. ($\Rightarrow$) Let $x, y \in H$ then since H is a subgroup $y^{-1} \in H$ and so $xy^{-1} \in H$.

($\Leftarrow$) Since $H \neq \phi$ so $\exists\, x \in H$. Then (i) $xx^{-1} = e \in H$, (ii) $ex^{-1} = x^{-1} \in H$, (iii) if $x, y \in H$ then $x, y^{-1} \in H$ and hence $x(y^{-1})^{-1} = xy \in H$. $\square$

This theorem gives an easy way to check if a subset of a group is a subgroup. For example to see if $\{\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : ac \neq 0\}$ is a subgroup of $GL_2(\mathbb{R})$ we check that

$$\begin{pmatrix} x & y \\ 0 & z \end{pmatrix} \begin{pmatrix} u & v \\ 0 & w \end{pmatrix}^{-1} = \begin{pmatrix} x & y \\ 0 & z \end{pmatrix} \begin{pmatrix} u & v \\ 0 & w \end{pmatrix}^{-1}$$

THEOREM 1.8. *Every subgroup of cyclic group is cyclic.*

PROOF. Let G be a cyclic group with generator x. Let $H < G$ then given $h \in H$ we have $x^j = h$ for some $j \in \mathbb{Z}$. Let m be the smallest positive integer such that $x^m \in H$.
claim : $\langle x^m \rangle = H$.
Let x^k be any element of H, then $k = qm + r$ with $0 \leq r < m$ (by Division algorithm theorem) and we get,

$$x^k = x^{qm+r} = x^{qm}x^r = (x^m)^q x^r$$

. Since, x^k and $(x^m)^q$ belongs to H therefore $x^r = x^k(x^m)^{-q} \in H$ but $r < m$ and by assumption m is the smallest positive integer with $x^m \in H$. Thus $r = 0$ and $x^k = (x^m)^q$ for some integer q. Therefore $H = \langle x^m \rangle$ and H is cyclic. $\square$

LEMMA 1.9. *Let $x \in G$ such that $|x| = k$, then if $x^m = e$ for some $m \in \mathbb{Z}^+$, then $k|m$.*

PROOF. $\square$

THEOREM 1.10. *Let $x \in G$ and $|x| = n$ then*

$$|x^a| = \frac{n}{gcd(n, a)}$$

PROOF. Let $y = x^a$. Let $d = gcd(n, a)$, then $\exists b, c \in \mathbb{Z}^+$ $n = bd$ and $a = cd$ with $gcd(b, c) = 1$. We need to show that $|y| = b$. Firstly,

$$y^b = (x^a)^b = (x^{cd})^b = (x^{bd})^c = (x^n)^c = e^c = e$$

Let $|y| = p$ then applying lemma 1.9 to $< y >$ we note that $p|b$, so

$$y^p = (x^a)^p = e$$

Again applying lemma 1.9 to $< x >$ we get $n|ap$ that is $bd|cdp$ so, $b|cp$. Since $gcd(b, c) = 1$ we get $b|p$. Since b and p divide each other it must be that $p = |y| = b$. $\square$

THEOREM 1.11. *Let G be a cyclic group of order n, then for each positive integer a such that $a|n$ there is a unique subgroup of G order a.*

<u>Examples:</u>

In $\mathbb{Z}_{36}$ the order of 1 is 36. Hence the order of 4 is $\dfrac{36}{gcd(36, 4)} = 9$.

In $\mathbb{Z}^*_{20}$ the order of $3 = 4$. Hence the order of $3^3 = 9$ is $\dfrac{4}{gcd(4, 3)} = 4$.

1.5. Permutation groups

A permutation of a set X is a bijection (one-one and onto mapping) from X onto itself. One can easily check that the set f permutations S_x is a group under the operation function composition. If α, β are permutations of X then we define the element $\alpha\beta(x)$=$\alpha((\beta(x)))$, for all $x \in X$. If we set X is the first n positive integers then S_x is written as S_n and is called the Symmetric group. For example the elements of the group S_3 are

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

If $\alpha = \begin{bmatrix}1 & 2 & 3\\2 & 1 & 3\end{bmatrix}$, and $\beta = \begin{bmatrix}1 & 2 & 3\\1 & 3 & 2\end{bmatrix}$ then $\alpha\beta$ (applying β first) is equal to $\begin{bmatrix}1 & 2 & 3\\2 & 3 & 1\end{bmatrix}$ whereas $\beta\alpha$ is given by $\begin{bmatrix}1 & 2 & 3\\3 & 1 & 2\end{bmatrix}$. So $\alpha\beta \neq \beta\alpha$ in general and S_n is a non-commutative group for all $n \geq 3$. One can write any permutation in cycle notation as discussed before. For example in $\begin{bmatrix}1 & 2 & 3 & 4 & 5 & 6\\5 & 4 & 3 & 6 & 1 & 2\end{bmatrix} \in S_6$ can be written as (15)(246) where a cycle is a permutation of the form $(a_1, a_2, a_3 \cdots, a_k)$ where $a_1 \to a_2$, $a_2 \to a_3$, $a_3 \to a_4 \cdots, a_k \to a_1$ while leaving all other elements fixed. $(a_1, a_2, a_3 \cdots, a_k)$ is called a k-cycle The procedure to write a permutation in cycle notation leads to a product of disjoint cycles and we have the following theorem.

THEOREM 1.12. *Every permutation in S_n can be written as a product of disjoint cycles.*

One can see that disjoint cycles commute. $(135)(24) = (24)(135)$. A 2-cycle is $(a_1, a2)$ is called a transposition.

THEOREM 1.13. *The transpositions generate S_n.*

PROOF. An arbitrary k-cycle $\{a_1, a_2, a_3 \ldots, a_k\}$ can be written as $\{a_1, a_2, a_3 \ldots, a_n\} = \{a_1, a_k\}\{a_1, a_{k-1}\} \ldots \{a_1, a_3\}\{a_1, a_2\}$ (check!)
Since every permutation of S_n is a product of disjoint cycles we can write any permutation as a product of transpositions. The decompositions of a permutation as a product of transpositions may not be unique. For example, $(15)(246) = (15)(26)(24) = (15)(46)(26)$. $\square$

THEOREM 1.14. i. *The transpositions $(12), (13), \ldots, (1n)$ generate S_n.*
ii. *The transpositions $(12), (23), \ldots, ((n-1)n)$ generate S_n.*

PROOF. i.$(ab) = (1a)(1b)(1a)$ then use previous theorem.
ii.$(1k) = ((k-1)k)((k-2)(k-1)) \ldots (34)(23)(12)(23) \ldots ((k-1)k)$ then use part i. $\square$

We already saw that given an element of S_n it can be decomposed as a product of transpositions in many different ways. However the number of transpositions that occur will always be either even or odd. Define a polynomial $P = \prod_{x_i < x_j} (x_i - x_j) = (x_1 - x_2)(x_1 - x_3) \ldots (x_{n-1} - x_n)$.If α is a permutation then $\alpha P = \prod_{x_i < x_j} (x_{\alpha(i)} - x_{\alpha(j)})$. Clearly, αP is either P or $-P$. If $\alpha P = P$ then we say that the sign of the permutation α sgn(α) is +1, otherwise if $\alpha P = -P$ then sgn(α) = −1. One note that if α, β are two permutations then sgn(α, β) =sgn(α)sgn(β).

Since a permutation can be written as a product of transpositions, the sign of a permutation is the product of the sign's of the transpositions (the sign of a transposition is -1). Hence a permutation can be written as either an even number or odd number of transpositions. The sign of a permutation is $+1$ if the permutation can be written as a product of even number of transpositions and these will be called Even Permutations. The sign of a permutation is -1 if the permutation can be written as a product of odd number of transpositions and these will be called Odd Permutations.

THEOREM 1.15. *The even permutations in S_n form a subgroup of order $\dfrac{n!}{2}$.*

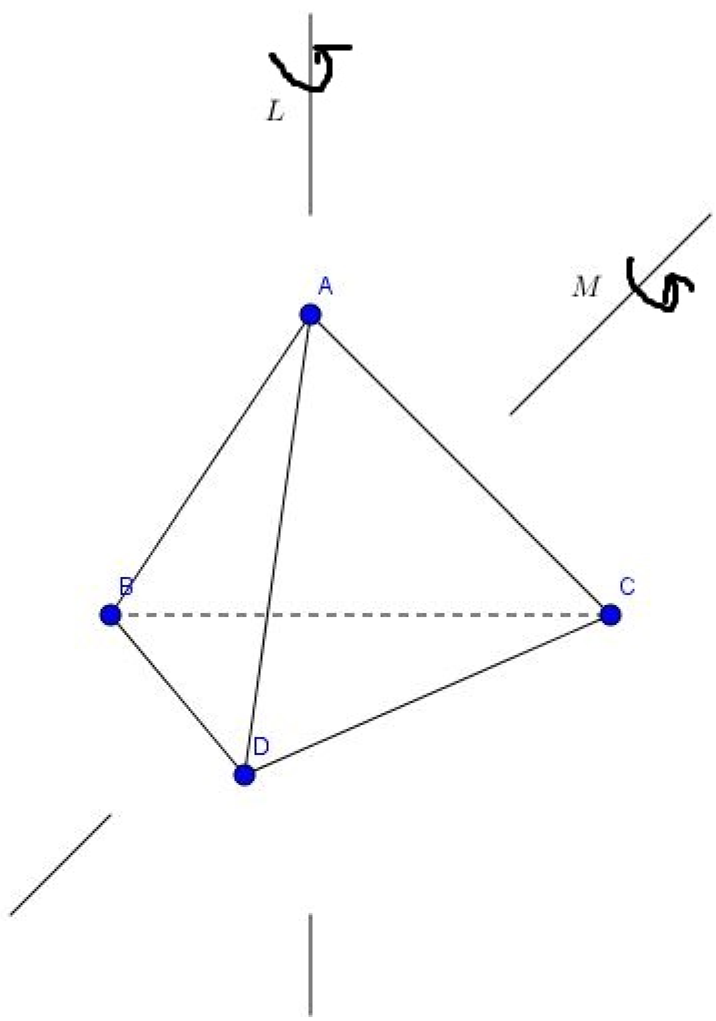

FIGURE 9. Rotational symmetries of the tetrahedron

PROOF. (1) e is an even permutation. $e = (12)(12)$.

(2) If α and β are even permutations then $\alpha\beta$ is an even permutation. (sum of two even numbers is even).

(3) If $\alpha = (a_1a_2)(a_3a_4)\dots(a_{n-1}a_n)$ is an even permutation then $\alpha^{-1} = (a_{n-1}a_n)(a_{n-2}a_{n-3})\dots(a_1a_2)$ is also even. There are exactly $\frac{n!}{2}$ even permutations of S_n as A_n since the mapping Φ : Even Permutations $\rightarrow$ Odd Permutations given by $\Phi(\alpha) = (12)\alpha$ is bijective.

$\square$

For example, subgroup A_4 of S_4 has the following elements. $\{e, (12)(34), (13)(24), (14)(23), (123), (132), (124), (142), (134), (143), (234), (243)\}$. Notice similarities with the group of rotational symmetries of the tetrahedron.

1.6. Group Isomorphism

Consider the symmetries of the chessboard as shown in the figure. There are four symmetries $\{e, r, q_1, q_2\}$. Here r is the rotation by angle π about the axis going through the centre and perpendicular to the plane in which the chessboard lies. q_1 and q_2 are reflections about the two diagonals. The group multiplicatin table is given below.

	e	r	q_1	q_2
e	e	r	q_1	q_2
r	r	e	q_2	q_1
q_1	q_1	q_2	e	r
q_2	q_2	q_1	r	e

Here consider the set $\{1, 3, 5, 7\}$ with operation $a * b = ab$ mod 8. This set forms a group and its group multiplication table is

	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

If we define a mapping $\Phi : \{e, r, q_1, q_2\} \rightarrow \{1, 3, 5, 7\}$. $\Phi(e) = 1$,$\Phi(r) = 3$, $\Phi(q_1) = 5$, $\Phi(q_2) = 7$. Then we see that the bijective map Φ obeys $\Phi(a * b) = \Phi(a) * \Phi(b)$. The two groups are structurally identical.

DEFINITION 1.16. Two groups G and K are said to be isomorphic if there exists a bijective map $\Phi : G \rightarrow K$ such that $\Phi(a * b) = \Phi(a) * \Phi(b)$ for all $a, b \in G$.

Examples of Isomorphism:
Example 1:
$(R, +)$ is ismorphic to $(R^+, \times)$.
The map $\Phi(x) = e^x$ is a bijection from R to R_+. Morever, $\Phi(x + y) = e^{x+y} = e^x.e^y = \Phi(x)\Phi(y)$. Here Φ is an isomorphism.
Example 2:
The group of rotational symmetries of a tetrahedron is isomorphic to A_4.
Example 3:
Every infinite cyclic group is isomorphic to Z.
Let G be an infinite cyclic group and let x be the generator of G then the mapping $\Phi(x_m) = m$ is an isomorphism (check!).
Example 4:

If G is a finite cyclic group of order m then it is isomorphic to Z_n.
$\Phi(x_k) = k \bmod m$ is the isomorphism (check!).
Example 5:
$\{1, -1, i, -i\}$ under multiplication is a group. Notice that it is a cyclic group. i and $-i$ are generators. It is a group of order 4. Here by example 4, it must be isomorphic to Z_4.

$$1 \to 0, -1 \to 2, i \to 1 \text{ and } -1 \to 3.$$

Example 6**:**
Q is not isomorphic to Q^+.
(Since Φ is an isomorphism), which implies that $\Phi(\frac{x}{2}) = \sqrt{2}$ which is a contradiction. Therefore, there is no isomorphic mapping from Q to Q^+. If $G \to K$ is an isomorphism then the following are true

(1) $\Phi(e) = e$.
(2) $\Phi(x^{-1}) = \Phi(x)^{(-1)}$.
(3) $|x| = |\Phi(x)|$.
(4) G is cyclic $\iff$ K is cyclic.
(5) G is abelian $\iff$ K is abelian.
(6) If $H \leq K$ then $\Phi(H) = \{\Phi(h) : h \in H\}$ is a subgroup of K.

Proofs:

(1)

$$\begin{aligned} \Phi(e_G * e_G) &= \Phi(e_G) * \Phi(e_G) \\ \Phi(e_G) &= \Phi(e_G) * \Phi(e_G) \\ \Phi(e_G) &= e_K \quad \text{Multiplying both sides by } \Phi(e_G)^{-1} \end{aligned}$$

(2)

$$\begin{aligned} \Phi(x * x^{-1}) &= \Phi(x) * \Phi(x^{-1}) \\ \Phi(e) &= \Phi(x) * \Phi(x^{-1}) \\ \Phi(x)^{-1} &= \Phi(x^{-1}) \quad \text{Multiplying both sides by } \Phi(x)^{-1} \end{aligned}$$

(3) Let $|x| = m$, then m is the smallest positive integer such that $x^m = e$. Since Φ is an isomorphism $\Phi(x^m) = \Phi(x)^m$, then m is also the smallest positive integer such that $\Phi(x)^m = e$. Indeed if $k < m$ such that $\Phi(x)^k = e$, then $\Phi(x)^k = \Phi(x^k) = e$, which implies that $x^k = e$, which contradicts the fact that $|x| = m$. Hence, $|\Phi(x)| = m$.
(4) Let $< x >= G$, then we claim that $< \Phi(x) >= K$. Let $k \in K$, since Φ is onto $\exists g \in G$ such that $\Phi(g) = k$. Since G is cyclic $g = x^m$ for some m and

$$k = \Phi(g) = \Phi(x^m) = \Phi(x)^m$$

Hence $< \Phi(x) >= K$, so K is cyclic. For the other direction if $< y >= K$ then a similar argument shows that $\Phi(y)^{-1}$ generates G and hence G is cyclic.

(5) Let $k_1, k_2 \in K$ then $\exists g_1, g_2 \in G$ such that $\Phi(g_1) = k_1$ and $\Phi(g_2) = k_2$. Therefore,

$$\begin{aligned} k_1 * k_2 &= \Phi(g_1) * \Phi(g_2) \\ &= \Phi(g_1 * g_2) \\ &= \Phi(g_2 * g_1) \quad \text{Since } G \text{ is Abelian} \\ &= \Phi(g_2) * \Phi(g_1) = k_2 * k_1 \end{aligned}$$

(6) We will use the subgroup criterea ($H \subset G$ is a subgroup if $x, y \in H \implies xy^{-1} \in H$). We have

$$\Phi(h_1) * \Phi(h_2)^{-1} = \Phi(h_1 * h_2^{-1}) = \Phi(h), \quad h = h_1 * h_2^{-1} \in H$$

Question Consider the three groups A_4, D_6 and Z_{12}. They are all groups of order 12. Are they isomorphic?
Ans.: Z_{12} is cyclic so it is not isomorphic to either A_4 or D_6 neither of which are cyclic. Moreever, D_6 has an element of order $6(r)$ but A_4 has no element of order 6. Therefore D_6 and A_4 are not isomorphic.

1.7. Products

Let G and K be groups. Consider the set $G \times K := \{(g, k) \ : \ g \in G \text{ and } k \in K\}$. Consider the operation on $G \times K$ defined as, if (g, k) and (g', k') are two elements of $G \times K$ then $(g, k)(g', k') = (gg', kk')$. *Claim* : $G \times K$ is a group under this operation.

(i) (e_G, e_k) is the identity since $(e_G, e_K)(g, k) = (e_G g, e_K k) = (g, k)$. Similarly, $(g, k)(e_G, e_K) = (g, k)$.
(ii) Associativity follows from associativity of the groups G and K, viz., $((g, k)(g', k'))(g'', k'') = (g, k)((g', k')(g'', k''))$
(iii) For each $(g, k) \in G \times K$, $(g^{-1}, k^{-1})(g, k) = (g, k)(g^{-1}, k^{-1}) = (e, e)$.

Note that we used (e, e) instead of (e_G, e_K).

Note that the subsets of $G \times K$ given by $\{(g, e) : g \in G\}$ and $\{(e, k) : k \in K\}$ are subgroups isomorphic to G and K, respectively. For example, $\mathbb{Z}_2 \times \mathbb{Z}_2$ is given by $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$. The corresponding group multiplication table is

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,1)	(1,0)	(1,1)
(0,1)	(0,1)	(0,0)	(1,1)	(1,0)
(1,0)	(1,0)	(1,1)	(0,0)	(0,1)
(1,1)	(1,1)	(1,0)	(0,1)	(0,0)

Notice that this is not a cyclic group. And this group is isomorphic to the group of symmetries of the chessboard and the group $\{1, 3, 5, 7\}$ under multiplication modula 8. Now look at the group table of the group $\mathbb{Z}_2 \times \mathbb{Z}_3$,

	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(1,2)
(0,0)	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(1,2)
(0,1)	(0,1)	(0,2)	(0,0)	(1,1)	(1,2)	(1,0)
(0,2)	(0,2)	(0,0)	(0,1)	(1,2)	(1,0)	(1,1)
(1,0)	(1,0)	(1,1)	(1,2)	(0,0)	(0,1)	(0,2)
(1,1)	(1,1)	(1,2)	(1,0)	(0,1)	(0,2)	(0,0)
(1,2)	(1,2)	(1,0)	(1,1)	(0,2)	(0,0)	(0,1)

In this group $(1,1) + (1,1) = (0,2), (0,2) + (1,1) = (1,0), (1,0) + (1,1) = (0,1), (0,1) + (1,1) = (1,2), (1,2) + (1,1) = (0,0)$. Hence $(1,1)$ is the generator of $\mathbb{Z}_2 \times \mathbb{Z}_3$ and this group is cyclic. This is a group of order 6 and as discussed in the previous lecture this group must be isomorphic to $\mathbb{Z}_6$. Hence, $\mathbb{Z}_2 \times \mathbb{Z}_3$ but $\mathbb{Z}_2 \times \mathbb{Z}_2 \ncong \mathbb{Z}_4$. We have the following theorem.

THEOREM 1.17. *$\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic iff $gcd(m,n) = 1$.*

PROOF. ($\Leftarrow$) Let $gcd(m,n) = 1$. We claim that $(1,1)$ generates $\mathbb{Z}_m \times \mathbb{Z}_n$. Let $(p,q) \in \mathbb{Z}_m \times \mathbb{Z}_n$
($\Rightarrow$) Let $\mathbb{Z}_m \times \mathbb{Z}_n$ be cyclic. Then, $\exists (x,y) \in \mathbb{Z}_m \times \mathbb{Z}_n$ of order mn and $(x,y)^{mn} = (0,0)$. But, $(x,y)^{lcm(m,n)} = (x^{lcm(m,n)}, y^{lcm(m,n)}) = (x^{k_1 m}, y^{k_2 n}) = (0,0)$ for some k_1, k_2. Since, mn was the smallest such power so $mn \leq lcm(m,n) \Rightarrow lcm(m,n) = mn \Rightarrow gcd(m,n) = 1$. □

1.8. Lagrange's theorem

THEOREM 1.18. ***Lagrange's theorem :*** *Let G be a finite group and H be a subgroup of G. Then the theorem states that $|H|$ divides $|G|$.*

PROOF. If $|H| = |G|$ then the result is trivial. If H is a proper subgroup of G then let $g_1 \in G$ such that $g_1 \notin H$ and consider the set $g_1 H = \{g_1 h : h \in H\}$. We have two claims

(i) $g_1 H \cap H = \phi$ and
(ii) $|g_1 H| = |H|$.

To prove the first part assume that $h \in g_1 H \cap H$, then $h = g_1 h_1$ for some $h_1 \in H$ which implies that $g_1 = h h_1^{-1}$. But this is a contradiction to the hypothesis that $g_1 \notin H$. Thus $g_1 H \cap H$.
To show the second part we define a mapping $\Phi\ H \longrightarrow\ g_1 H$ given by $\Phi(h) := g_1 h$. We claim that the map is bijective (check !) and hence the result follows.
Next if $g_1 H \cup H$ is proper subset of G then let $g_2 \in G$ such that $g_2 \notin g_1 H \cup H$, then again we have the claims

(iii) $g_2 H \cap H = \phi$, $g_2 H \cap g_1 H = \phi$ and
(iv) $|g_2 H| = |H|$.

We only need to prove that $g_2 H \cap g_1 H = \phi$, rest can be proved by the same arguments discussed in (i) and (ii). Let $x \in g_2 H \cap g_1 H$ then x is of the form $x = g_2 h = g_1 h'$ for some $h, h' \in H$. Therefore $g_2 = g_1 h' h^{-1}$ which implies that $g_2 \in g_1 H$ which is a contradiction. So, $g_2 H \cap g_1 H = \phi$.
Continuing in this manner till we exhaust all the elements of G (this happens because G is finite) we get a partition of G as shown

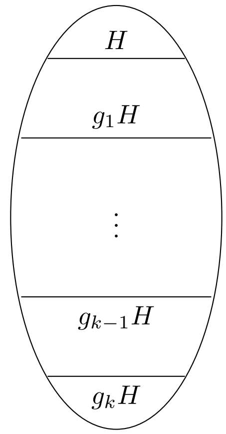

FIGURE 10. Partition of G

Then counting the elements of G we get

$$|G| = |H| + |g_1H| + \cdots + |g_kH|$$

$$|G| = (k+1)|H| \Rightarrow |H| \text{ divides } |G|$$

$\square$

Applications of Lagrange's theorem :

COROLLARY 1.19. Let G be a group and let $x \in G$ then $|\langle x \rangle|$ divides $|G|$.

COROLLARY 1.20. Let G be a group of prime order then G is cyclic.

PROOF. Let $x \in G$ such that $x \neq e$ and consider $\langle x \rangle$. By Corollary 1 $|\langle x \rangle|$ divides $|G|$. Since $|G|$ is prime so either $|\langle x \rangle| = 1$ or $|\langle x \rangle| = |G|$. Since $x \neq e$ so $|\langle x \rangle| = |G|$. Therefor G is cyclic generated by x. $\square$

REMARK 1.21. Note that if G is cyclic of prime order then each non-identity element generates G. However, in case of $|G| = 1$ identity is the generator.

COROLLARY 1.22. Let G be a group and x be any element of G then $x^{|G|} = e$.

PROOF. Let m be the order of x. From corollary 1 m divides G, so $|G| = km$ for some $k \in \mathbb{Z}$. Thus, $x^{km} = (x^m)^k = e^k = e$. $\square$

Consider the set $\mathbb{Z}_n^*$ consisting of elements that are less than n and relatively prime to n. For example $\mathbb{Z}_8^* = \{1,3,5,7\}$ and $\mathbb{Z}_{20}^* = \{1,3,7,9,11,13,17,19\}$. This forms a group under multiplication modulo n (check !). The order of this group is $\Phi(n)$, the Euler phi function.

COROLLARY 1.23. **Euler's theorem :** If $gcd(x,n) = 1$ then $x^{\Phi(n)} \equiv 1 \; mod \; n$.

PROOF. Let $x \in \mathbb{Z}_n^*$ then from corollary 3 we get $x^{\Phi(n)} = 1$. $\square$

COROLLARY 1.24. **Fermat's little theorem :** If p is a prime and x is not a multiple of p then $x^{p-1} \equiv 1 \; mod \; p$.

PROOF. Apply Euler's theorem with $n = p$. (Note that $\Phi(p) = p - 1$). $\square$

1.9. Equivalence relations and Partitions

Let M be a set and let $R \subseteq M \times M$. We say that x y (x is related to y) if $(x, y) \in R$. R is called an equivalence relation if a) $x \sim x$, b) if $x \sim y$ then $y \sim x$ and c) $x \sim y$ and $y \sim z \Rightarrow x \sim z$.

For example consider $M = \mathbb{Z}$ and the relation $x \sim y$ iff $x - y \equiv 0 \bmod 3$. One can easily check that this is an equivalence relation. Indeed $x \sim x$ since $x-x \equiv 0 \bmod 3$. If $x \sim y$ then $x-y \equiv 0 \bmod 3$ which implies that $y-x \equiv 0 \bmod 3$. Also if $x - y \equiv 0 \bmod 3$ and $y - z \equiv 0 \bmod 3$ then $(x - y) + (y - z) \equiv 0 \bmod 3$ or $x - z \equiv 0 \bmod 3$. Therefore if $x \sim y$ and $y \sim z$ then $x \sim z$. Therefore this is an equivalence relation. Let R be an equivalence relation on set M and $x \in M$. We call $R(x)$ to be all those elements that are related to x. It is called the equivalence class of x.

In the equivalence relation on $\mathbb{Z}$ just described $R(0) = \{\dots, -3, 0, 3, 6, \dots\}$, $R(1) = \{\dots, -2, 1, 4, 7, \dots\}$ and $R(2) = \{\dots, -1, 2, 5, 8, \dots\}$. The set $\mathbb{Z}$ is partitioned into three sets $R(0), R(1)$ and $R(2)$.

THEOREM 1.25. *If R is an equivalence relation on a non-empty set M then the distinct equivalence classes form a partition of M.*

PROOF. Let $R(x)$ and $R(y)$ be two distinct equivalence classes. If $R(x) \cap R(y) \neq \phi$ then $\exists z \in R(x) \cap R(y)$. Since $z \in R(x)$ so $x \sim z$ and since $z \in R(z)$ so $z \sim y$. Thus by transitive property of R $x \sim y$ which is a contradiction since it was assumed that $R(x)$ and $R(y)$ are two distinct equivalence classes. Next we need to show that $\cup_{x \in M} R(x) = M$. Clearly, $\cup_{x \in M} R(x) \subseteq M$. And the other inclusion is due to the fact that if $m \in M$ then $m \in R(m)$, i.e. every element of M is in its own equivalence class. $\square$

EXAMPLE 1.26. Consider the set $\mathbb{Z}$ with relation $x \sim y$ iff $x - y \equiv 0 \bmod n$. This is an equivalence relation (check !). The equivalence class of $m \in \mathbb{Z}$ is $R(m) := \{x \in \mathbb{Z} \,:\, x - m \equiv 0 \bmod n\}$.

EXAMPLE 1.27. Let G be a group and let H be a subgroup of G. We define the following relation on G as $x \sim y$ iff $y^{-1}x \in H$. This is an equivalence relation. Indeed, $x \sim x$ since $x^{-1}x \in H$. If $x \sim y$ then $y^{-1}x \in H \Rightarrow x^{-1}y = (y^{-1}x)^{-1} \in H$ so $y \sim x$. If $x \sim y$ and $y \sim z$ then $y^{-1}x$ and $z^{-1}y$ belongs to H. Hence, $z^{-1}x = (z^{-1}y)(y^{-1}x) \in H$ and so $x \sim z$. If $g \in G$ then the equivalence class $R(g) := \{x \in G : g^{-1}x \in H\}$., i.e. the equivalence class of g is all those elements $x \in G$ such that $g^{-1}x = h$ for some $h \in H$, or $x = gh$. This is precisely the set $gH = \{gh : h \in H\}$ as mentioned in Lagrange's theorem. The set gH will be called the left cosets of H. Note that $g_1H = g_2H$ iff $g_2^{-1}g_1 \in H$. We can define another equivalence relation $x \sim y$ iff $xy^{-1} \in H$. Check that This is an equivalence relation on G. The equivalence class of an element $g \in G$ will be $R(g) := \{x \in G : xg^{-1} \in G\}$ which leads to $R(g) = Hg$ where $Hg = \{hg : h \in H\}$ are the right cosets of H. Note that Hg in general will not be equal to gH.

EXAMPLE 1.28. Let G be a group and consider the relation $x \sim y$ iff $gxg^{-1} = y$ for some $g \in G$. This is an equivalence relation on G. a) $x \sim x$ since $exe^{-1} = x$. b) If $x \sim y$ then $\exists g \in G$ such that $gxg^{-1} = y$ then for $g_1 = g^1$ we get $g_1yg_1^{-1} = x$, therefore $y \sim x$. c) If $x \sim y$ and $y \sim z$ then $\exists g_1, g_2 \in G$ such that $y = g_1xg_1^{-1}$

and $z = g_2 y g_2^{-1}$. And we get $z = g_2(g_1 x g_1^{-1})g_2^{-1} = g_2 g_1 x g_1^{-1} g_2^{-1} = gxg^{-1}$, where $g = g_2 g_1$. Therefore, $x \sim z$. The equivalence class of an element $x \in G$ is $R(x) := \{gxg^{-1} : g \in G\}$. These are called the conjugacy classes of G.

1.10. Conjugacy classes

Consider the equivalence relation given in the previous lecture. Let G be a group and $x, y \in G$ then $x \sim y$ iff $\exists\ g \in G$ such that $gxg^{-1} = y$. This relation was shown to be an equivalence relation. The equivalence class of an element x is $R(x) = \{gxg^{-1} \ : \ g \in G\}$. The equivalence classes under this relation will be called conjugacy classes. For example in D_6, $\{e\}$ is obviously in its own an equivalence class since $geg^{-1} = e \ \forall\ g \in G$. Let us now compute the equivalence class of r. Since r commutes with elements of the type r^a, $r^a r r^{-a} = r$, and so we do not get any new element in the conjugacy class of r by conjugating with elements of type r^a. Now $srs^{-1} = r^{-1} = r^5$ and $r^a sr(r^a s)^{-1} = r^5$. Therefore, the conjugacy class of r is $R(r) = \{r, r^5\}$. Similarly lets compute the equivalence class of r^2. Since we do not get any new element by conjugating with elements of the type r^a we check $sr^2s^{-1} = r^4$ and as before $r^a sr^2(r^a s)^{-1} = r^4$. Thus we get $R(r^2) = \{r^2, r^4\}$. In the case of r^3 since $r^{-3} = r^3$ we get only a single element i.e. $R(r^3) = \{r^3\}$. For computing the equivalence class of elements of the type $r^a s$ we observe that $r^k r^a s r^{-k} = r^{2k+a}s$, $(k = 0, 1, 2, \ldots, 5)$ and $r^k sr^a s(r^k s)^{-1} = r^{2k-a}s$, $(k = 0, 1, \ldots, 5)$. thus the conjugacy class of rs is $R(rs) = \{rs, r^3s, r^5s\}$ and that of r^2s is $R(r^2s) = \{r^2s, r^4s, s\}$. Hence the equivalence classes of D_6 are $R(e) = \{e\}$, $R(r) = \{r, r^5\}$, $R(r^2) = \{r^2, r^4\}$, $R(r^3) = \{r^3\}$, $R(rs) = \{rs, r^3s, r^5s\}$, $R(r^2s) = \{r^2s, r^4s, s\}$.

In general we can write down the conjugacy classes of D_n in the following way. Notice that if n is even then there are two conjucgacy classes with singleton elements, namely $\{e\}$ and $\{r^{n/2}\}$. If n is odd then there is only trivial conjugacy class that is a singleton i.e. $\{e\}$. What are these one element conjugacy classes? They are all those elements of G that satisfy $\{x \in G \ : \ gxg^{-1} = x \ \forall\ x \in G\}$. But $gxg^{-1} = x \implies gx = xg$. Therefore the single element conjugacy classes are precisely those elements of G that commute with every other elements of G.

What about the conjugacy classes of S_n? Lets start with S_3. One can verify that the conjugacy classes of S_3 are $\{e\}, \{(1,2), (2,3), (1,3)\}$ and $\{(1,2,3), (1,3,2)\}$. The following theorem helps to classify all the conjugacy classes of S_n.

THEOREM 1.29. *Two elements of S_n are in the same conjugacy class iff they have the same cycle structure.*

PROOF. ($\Leftarrow$) Let θ and ϕ be elements of S_n that have the same cycle structure. We need to show that $\exists\ g \in S_n$ such that $g\theta g^{-1} = \phi$. To make things clear lets take an example in S_9. Put $\theta = (67)(2539)(14)$ and $\phi = (12)(38)(5467)$. Check that θ and ϕ have the same cycle structure. To find g write θ and ϕ in the decreasing order of cycle length, and write the elements that remain fixed as cycles of length 1.

$$\theta = (2539)(67)(14)(8)$$

$$\phi = (5467)(12)(38)(9)$$

Take g to be $(136)(254897)$ and then claim $g\theta g^{-1} = \phi$. We need to check if $g\theta g^{-1}(x) = \phi(x)\ \forall\ x$.

incomplete

($\Rightarrow$) We need to show that if two elements of S_n are conjugate then they have the same cycle structure. We will show that if $\theta \in S_n$ then $g\theta g^{-1}$ has the same cycle structure for all $g \in G$. Let $\theta = \theta_1\theta_2 \dots \theta_t$ be the representation of θ as a product of disjoint cycles. Then,

$$g\theta g^{-1} = g_1^\theta g^{-1} g\theta_2 g^{-1} \dots g\theta_t g^{-1}.$$

It is enough to show that $g\theta_i g^{-1}$ has the same cycle structure as θ_i. Let $\theta_i = (a_1 a_2 \dots a_k)$, then

$$\begin{aligned} g\theta_i g^{-1}(g(a_1)) &= g(a_2) \\ g\theta_i g^{-1}(g(a_2)) &= g(a_3) \\ &\vdots \\ g\theta_i g^{-k}(g(a_k)) &= g(a_1) \end{aligned}$$

Therefore $g\theta_i g^{-1} = (g(a_1)\ g(a_2)\ \dots\ g(a_k))$ which is the same cycle structure as θ_i. The conjugacy classes of S_4 are

$\{e\}$
$\{(12), (13), (14), (23), (24), (34)\}$
$\{(123), (132), (124), (142), (134), (143), (234), (243)\}$
$\{(1234), (1243), (1324), (1342), (1423), (1432)\}$
$\{(12)(34), (13)(24), (14)(23)\}$

$\square$

1.11. Quotient groups

We learnt about two equivalence relations on a group G, which lead to the left and right cosets.

$H \leq G$, $x \sim y$ iff $y^{-1}x \in H$ leads to the left cosets $R(g) = gH$.
$H \leq G$, $x \sim y$ iff $xy^{-1} \in H$ leads to the right cosets $R(g) = Hg$.

As we saw before although the left cosets and the right cosets form a partition of G they may bot be the same. For example let $G = D_6$ and $H =< s >$, then $r < s >= \{r, rs\}$ and $< s > r = \{r, r^5 s\}$.

image

If $gH = Hg$ for all $g \in G$ then if $h \in H$ then $\exists h' \in H$ such that $gh = h'g$ which implies that $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$. By a similar argument if $ghg^{-1} \in H$ for all $g \in G$ then $gH = Hg$. But gHg^{-1} $\forall h \in H$ and $\forall g \in G$ means that H is a union of conjugacy classes.

Definition 1.30. Normal subgroup. H is a normal subgroup, denoted as $H \lhd G$, if H is a union of conjugacy classes.

We have the following proposition.

Proposition 1.31. (1) $H \lhd G$.
(2) $gH = Hg$ $\forall g \in G$.

(3) $ghg^{-1} \in H \; \forall h \in H, \forall g \in G$

EXAMPLE 1.32. If G is abelian group and H is any subgroup of G then $H \lhd G$.

EXAMPLE 1.33. A subgroup H of G within index 2 (index of a subgroup is the number of right or left cosets) is normal. Since there are only two cosets and one of them is H we have $gH = Hg \; \forall g \notin H$.

EXAMPLE 1.34. $Z(G) = \{g \in G : gx = xg \forall x \in G\}$ is called the center of the group G. It is a subgroup (check!). It is also the union of single element conjugacy classes the elements of which commute with all the elements of the group, hence it is a normal subgroup.

EXAMPLE 1.35. $A_n \lhd S_n$, since it is subgroup of index 2. $< r > \lhd D_n$ for the same reason.

EXAMPLE 1.36. If n is even then $Z(D_n) = \{e, r^{n/2}\}$ hence $< r^{n/2} > \lhd D_n$. If n is odd then $Z(D_n) = \{e\}$.

EXAMPLE 1.37. In S_n, $\langle(123)\rangle = \{e, (123), (132)\}$. So, $\langle(123)\rangle \lhd S_3$ since it is a subgroup of index 2, but $\langle(12)\rangle = \{e, (12)\}$ is not normal in S_3 since it is not a union of conjugacy classes. Recall that $\{(12), (23), (13)\}$ are in the same conjugacy class.

If $H \trianglelefteq G$ then we can form another group called G/H (G mod H). The group elements are the cosets $(gHorHg)$ and the group operation $(g_1H)(g_2H) = g_1g_2H$. Note that this is a valid operation only if $gH = Hg$.

claim: The cosets under the given operation form a group. Firstly $g_1H.g_2H$ is a well defined operation since each element of g_1Hg_2H is of the form g_1hg_2h' but since $H \lhd \; G \; hg_2 = g_2h"$ for some $h" \in H$. Therefore, $g_1hg_2h' = g_1g_2h"h' \in g_1g_2H$. Moreover if $g_1H = g_1'H$ and $g_2H = g_2'H$ then $g_1'Hg_2'H = g_1g_2H$ so the operation is well defined on the cosets. Also the identity element of G/H is H sicne $eHg_1H = eg_1H = g_1H$. Associativity follows from the associativity in G. Finally $g_1H \in G/H \; g_1^{-1}H.g_1H = g_1^{-1}g_1H = H$. So inverse also exists. Hence G/H is a group.

EXAMPLE 1.38. Let $G = \mathbb{Z}$ and $H = n\mathbb{Z} = \{\dots, -3n, -2n, -n, 0, n, 2n, 3n, \dots\}$ since $\mathbb{Z}$ is abelian $n\mathbb{Z}$ is a normal subgroup. $\mathbb{Z}/n\mathbb{Z}$ consists of the equivalence classes $n\mathbb{Z}, 1 + n\mathbb{Z}, 2 + n\mathbb{Z}, \dots, (n-1) + n\mathbb{Z}$. For example if $n = 3$ then there are three partitions of Z, i.e.

$$0 + 3\mathbb{Z} = \{\dots, -6, -3, 0, 3, 5, \dots\}$$
$$1 + 3\mathbb{Z} = \{\dots, -5, -2, 1, 4, 7, \dots\}$$
$$2 + 3\mathbb{Z} = \{\dots, -4, -1, 2, 5, 8, \dots\}$$

The group multiplication table is

	$3\mathbb{Z}$	$1 + 3\mathbb{Z}$	$2 + 3\mathbb{Z}$
$3\mathbb{Z}$	$3\mathbb{Z}$	$1 + 3\mathbb{Z}$	$2 + 3\mathbb{Z}$
$1 + 3\mathbb{Z}$	$1 + 3\mathbb{Z}$	$2 + 3\mathbb{Z}$	$3\mathbb{Z}$
$2 + 3\mathbb{Z}$	$2 + 3\mathbb{Z}$	$3\mathbb{Z}$	$1 + 3\mathbb{Z}$

One can see that $\mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}_3$. In general $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$ under the isomorphism mapping $k + n\mathbb{Z} \longrightarrow k$.

EXAMPLE 1.39. Consider $G = D_6$ and $H = \langle r^3 \rangle$ then H is a normal subgroup since it is equal to $Z(D_6)$. $D_6/\langle r^3 \rangle$ consists of the following elements.

$\langle r^3 \rangle = \{e, r^3\}$
$s\langle r^3 \rangle = \langle r^3 \rangle s = \{s, r^3 s\}$
$rs\langle r^3 \rangle = \langle r^3 \rangle rs = \{rs, r^4 s\}$
$r^2 s\langle r^3 \rangle = \langle r^3 \rangle r^2 s = \{r^2 s, r^5 s\}$
$r\langle r^3 \rangle = \langle r^3 \rangle r = \{r, r^4\}$
$r^2\langle r^3 \rangle = \langle r^3 \rangle r^2 = \{r^2, r^3\}$

Now $D_6/\langle r^3 \rangle \cong D_3$ and the isomorphism is given by

$$\begin{aligned} \langle r^3 \rangle &\mapsto e, & r^s\langle r^3 \rangle &\mapsto r^2 s \\ s\langle r^3 \rangle &\mapsto s, & r\langle r^3 \rangle &\mapsto r \\ rs\langle r^3 \rangle &\mapsto rs, & r^2\langle r^3 \rangle &\mapsto r^2 \end{aligned}$$

Let G and H be groups and let $\Phi : G \to H$ be an homomorphism

DEFINITION 1.40. $\mathrm{Ker}(\Phi) = \{g \in G | \Phi(g) = e\}$ is the Kernel of the homomorphism

DEFINITION 1.41. $\mathrm{Im}(\Phi) = \{\Phi(g) | g \in G\}$ is the Image of the homomorphism

THEOREM 1.42. *Let $\Phi : G \to H$ be an homomorphism then*

$$\begin{aligned} Im(\Phi) &\leq H \\ Ker(\Phi) &\trianglelefteq G \end{aligned}$$

PROOF. Let $h_1, h_2 \in \mathrm{Im}(\Phi)$ then $\exists g_1, g_2 \in G$ such that $\Phi(g_1) = h_1$ and $\Phi(g_2) = h_2$. Then,

$$\begin{aligned} h_1 h_2^{-1} &= \Phi(g_1)\Phi(g_2)^{-1} \\ &= \Phi(g_1 g_2^{-1}) = \Phi(g) \quad g = g_1 g_2^{-1} \end{aligned}$$

So $\mathrm{Im}(\Phi)$ is a subgroup by the subgroup criterion.

Let $g_1, g_2 \in \mathrm{Ker}(\Phi)$ then

$$\Phi(g_1 g_2^{-1}) = \Phi(g_1)\Phi(g_2)^{-1} = ee^{-1} = e$$

Hence, $\mathrm{Ker}(\Phi) < G$. Moreover if $k \in \mathrm{Ker}(\Phi)$, and $g \in G$ then

$$\begin{aligned} \Phi(gkg^{-1}) &= \Phi(g)\Phi(k)\Phi(g)^{-1} \\ &= \Phi(g)e\Phi(g)^{-1} = e \end{aligned}$$

Hence $\mathrm{Ker}(\Phi) \trianglelefteq G$. □

THEOREM 1.43 (First Isomorphism Theorem). *Let $\Phi : G \to H$ be a homomorphism of groups G and H, then*

$$G/Ker(\Phi) \equiv Im(\Phi)$$

PROOF. Consider the mapping $\Psi : G/\mathrm{Ker}(\Phi) \to \mathrm{Im}(\Phi)$, $g\mathrm{Ker}(\Phi) \mapsto \Phi(g)$. We claim that this is an isomorphism. We will first show that the map Ψ is well defined. Let $g \sim g'$ then $g^{-1}g' \in \mathrm{Ker}(\Phi)$, so $\Phi(g^{-1}g') = e$. Therefore,

$$\Phi(g^{-1}g') = \Phi(g^{-1})\Phi(g')) = e \implies \Phi(g) = \Phi(g')$$

Now,

$$\begin{aligned}
\Psi(g_1\mathrm{Ker}(\Phi) * g_2\mathrm{Ker}(\Phi)) &= \Psi(g_1 g_2\mathrm{Ker}(\Phi)) \\
&= \Phi(g_1 g_2) \\
&= \Phi(g_1)\Phi(g_2) \\
&= \Psi(g_1\mathrm{Ker}(\Phi))\Psi(g_2\mathrm{Ker}(\Phi))
\end{aligned}$$

which shows that the map Ψ is a homomorphism. Let $h \in \mathrm{Im}(\Phi)$, then $\exists g \in G$ such that $\Phi(g) = h$, then $\Psi(g\mathrm{Ker}(\Phi)) = \Phi(g) = h$. Therefore, Ψ is onto. Moreover, if

$$\begin{aligned}
\Psi(g\mathrm{Ker}(\Phi)) &= \Psi(g'\mathrm{Ker}(\Phi)) \quad \text{then} \\
\Phi(g) &= \Phi(g') \\
\Phi(g^{-1})\Phi(g') &= \Phi(g^{-1}g') = e \implies
\end{aligned}$$

$g^{-1}g' \in \mathrm{Ker}(\Phi)$ or $g\mathrm{Ker}(\Phi) = g'\mathrm{Ker}(\Phi)$, so Ψ is 1-1. Hence Ψ is an isomorphism. □

Exercise 1

1. Find all the rotational symmetries of the cube.

2. If G is a group then show the following
   i. The identity element of G is unique
   ii. For $x \in G$ then x has a unique inverse.
   iii. For $a, b \in G$ there is a unique x such that $a * x = b$.

3. Determine whether the binary operation $*$ gives a group structure
   i. Let $*$ be defined on $\mathbb{Z}$ by $a * b = ab$.
   ii. Let $*$ be defined on $\mathbb{R}^+$ by $a * b = \sqrt{ab}$.
   iii. Let $*$ be defined on $\mathbb{R} - \{0\}$ by $a * b = \frac{a}{b}$.

4. Let $G = \{a + \sqrt{2}b \in \mathbb{R} | a, b \in \mathbb{Q}\}$
   i. Prove that G is a group under addition.
   ii. Prove that the non-zero elements of G are a group under multiplication.

5. Let S be the set of all real numbers except -1. Define an operation $*$ on S by
   $$a * b = a + b + ab$$
   i. Show that $< S, * >$ is a group.
   ii. Find the solution to the following equation in S.
   $$2 * x * 3 = 7$$

6. Show that a group of three elements is commutative.

7. If x and y are elements of a group show that $(x * y)^{-1} = y^{-1} * x^{-1}$.

8. Prove that if $x^2 = 1$ for all $x \in G$ then G is a commutative (abelian) group.

9. Show that the following subsets of D_4 are actually subgroups.
   i. $\{1, r^2, s, r^2s\}$.
   ii. $\{1, r^2, rs, r^3s\}$

10. Determine if the following set of matrices are subgroups of $GL_n(\mathbb{R})$.
    i. The diagonal $n \times n$ matrices with no zeros on diagonal.
    ii. The $n \times n$ matrices with determinant -1.
    iii. The set of $n \times n$ matrices such that $A^T A = \mathrm{I}$.

11. Find the order of
    i. 2,6,10 in the additive group $\mathbb{Z}_{36}$.
    ii. 2 in the multiplicative group $\mathbb{Z}_{13}^*$.

12. What are the generators of $\mathbb{Z}_5$? What about $\mathbb{Z}_9$ and $\mathbb{Z}_{12}$? Do you notice a pattern?

13. Show that D_n is generated by by two elements rs and r^2s.

14. Let x and g be elements of a group G. Show that x and gxg^{-1} have the same order. Now show that xy and yx have the same order for any two elements x, y in G.

15. Consider the group of invertible 2×2 matrices with entries in real numbers under matrix multiplication $GL_2(\mathbb{R})$. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ belong to $GL_2(\mathbb{R})$. Compute the order of A, B, AB, BA

16. Let σ be the permutation $1 \mapsto 3$, $2 \mapsto 4$,$3 \mapsto 5$,$4 \mapsto 2$,$5 \mapsto 1$
    and let τ be the permutation $1 \mapsto 5$, $2 \mapsto 3$,$3 \mapsto 2$,$4 \mapsto 4$,$5 \mapsto 1$
    Find the cycle decompositions of σ^2, $\sigma\tau$ and $\tau^2\sigma$

17. Show that if σ is the m-cycle $(a_1 a_2 ... a_m)$ then $|\sigma| = m$.

18. Compute the order of the element $(13)(246)$ in S_6.

19. Show that if $n \geq m$ then the number of m-cycles in S_n is given by

$$\frac{n(n-1)(n-2)\cdots(n-m+1)}{m}$$

20. Let $\mathbb{Z}_n^*$ be set of integers that are less than n and relatively prime to n. For example $\mathbb{Z}_8^* = \{1, 3, 5, 7\}$. In class we saw that $\mathbb{Z}_n^*$ is a group under multiplication modulo n. $|\mathbb{Z}_n^*| = \phi(n)$ and its not always cyclic like $\mathbb{Z}_n$.

Show that $\mathbb{Z}_9^*$ is isomorphic to $\mathbb{Z}_6$.

21. Show that $\mathbb{Z}_{20}^*$ is not isomorphic to $\mathbb{Z}_8$.

22. An isomorphism of a group onto itself is called an automorphism. Let G be a group and let g be an element of G. Show that the mapping $x \mapsto gxg^{-1}$ is an automorphism of G.

23. Let G be a group. Show that the mapping $x \mapsto x^{-1}$ is an automorphism G iff G is abelian.

24. Show that if G and H are groups and if $G \times H$ is cyclic then G and H are both cyclic.

25. Show that $\mathbb{Z} \times \mathbb{Z}$ is not isomorphic to $\mathbb{Z}$

26. How many different isomorphisms are there from S_3 to D_3 ?

27. **Fact**: $\mathbb{Z}_m^* \times \mathbb{Z}_n^*$ is isomporphic to $\mathbb{Z}_{mn}^*$ if m,n are relatively prime.
Use this fact to show that $\mathbb{Z}_{20}^*$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_4$

28. Show that if G is a group of order 4 that is not cyclic then it is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

29. Which are the subsets of $\mathbb{R} \times \mathbb{R}$ are equivalence relations
i) $\{(x, y)|x - y \text{ is and even integer}\}$
ii) $\{(x, y)|x - y \text{ is rational}\}$
i) $\{(x, y)|x + y \text{ is rational}\}$
i) $\{(x, y)|x - y \geq 0\}$

30. Let G be a group and $|G|= pq$ where p and q are primes. Show that any proper subgroup of G is cyclic.

31. The remainder when 3^{64} is divided by 20 is?

32. If H and K are subgroups of a group G such that their orders are relatively prime, show that H and K only have the identity element in common.

33. Find all the cosets of
i) The subgroup $4\mathbb{Z}$ of $\mathbb{Z}$.
ii)The subgroup $< 2 >$ of $\mathbb{Z}_{12}$.

34. Find all the normal subgroups of D_4

35. Let G a group then $Z(G) = \{x \in G | xg = gx \, \forall g \in G\}$ is the set all elements of G that commute with every other element of G. It is called the center of the group G. Show that
i) $Z(G)$ is a subgroup of G.
ii)Show that $Z(G)$ is the union of all single element conjugacy classes.
iii)$Z(G)$ is a normal subgroup of G.

36. Show that the $Z(D_n) = \{e\}$ when n is odd and $Z(D_n) = \{e, r^{\frac{n}{2}}\}$ when n is even

37. Find all the conjugacy classes of D_4 and D_5 and in general D_n.

38. Find all the conjugacy classes of S_6.

CHAPTER 2

Vector Spaces and Linear Transformations

2.1. Vector spaces and subspaces

DEFINITION 2.1. Let V be a set and $\mathbb{F}$ be a field of scalar. Let $+ : V \times V \to V$ and $\cdot : \mathbb{F} \times V \to V$ be two operations (addition and scalar multiplication) then V is called a vector space over $\mathbb{F}$ if

(1) $(V, +)$ is a commutative group.
(2) The following properties hold
  i. $1v = v \, \forall v \in V$
  ii. $\alpha(\beta v) = (\alpha\beta)v \, \forall \alpha, \beta \in \mathbb{F}, v \in V$
  iii.$\alpha(u + v) = \alpha u + \alpha v \, \forall u, v \in V, \alpha \in \mathbb{F}$
  iv. $(\alpha + \beta)u = \alpha u + \beta u \, \forall u \in V, \alpha, \beta \in \mathbb{F}$

EXAMPLE 2.2. $\mathbb{R}^n(\mathbb{C}^n)$ are vector spaces over $\mathbb{R}(\mathbb{C})$. An element of $\mathbb{R}^n(\mathbb{C}^n)$ is a n-tuple $(a_1, a_2, ..., a_n)$ with each $a_i \in \mathbb{R}(\mathbb{C})$. It is easy to verify that all the properties of a vector space are satisfied.

EXAMPLE 2.3. The space of $n \times n$ matrices with entries in $\mathbb{R}$ called $M_n(\mathbb{R})$ and the space of $n \times n$ matrices with entries in $\mathbb{C}$ called $M_n(\mathbb{C})$ are vector spaces under element-wise matrix addition.

EXAMPLE 2.4. The space of functions $\{f : \mathbb{R} \to \mathbb{R}\}$ is a vector space under the addition operation $(f + g)(x) = f(x) + g(x)$ and scalar multiplication operation $(\alpha f)(x) = \alpha f(x)$.

EXAMPLE 2.5. / The space of polynomials with coefficients in the rational numbers $\mathbb{Q}[x]$ is also a vector space over the scalar field $\mathbb{Q}$.

DEFINITION 2.6. A subset $U \subseteq V$ of a vector space is a subspace if it is closed under addition and scalar multiplication.

EXAMPLE 2.7. The subset $\{(\alpha_1, \alpha_2, ..., \alpha_n) | \alpha_i \in \mathbb{R}, \alpha_1 = 0\}$ is a subspace of $\mathbb{R}^n$

EXAMPLE 2.8. The subset of vectors in $\{(x_1, x_2, x_3) | x_1 + 2x_2 + x_3 = 0\}$ is a subspace of $\mathbb{R}^3$

EXAMPLE 2.9. The subset of $M_n(\mathbb{R})$ of symmetric matrices, that is, $\{A \in M_n(\mathbb{R}) | A = A^T\}$ is a subspace.

EXAMPLE 2.10. The subset of $M_n(\mathbb{C})$ given by the Hermitian matrices, that is, matrices that satisfy $\{A \in M_n(\mathbb{C}) | A = A^\dagger\}$, where $(A^\dagger)_{ij} = \bar{A}_{ij}$ denotes the conjugate transpose is not a subspace of $M_n(\mathbb{C})$. Indeed if A is Hermitian then iA is not Hermitian because $(iA)^\dagger = -iA$. Hence this subset is not closed under scalar multiplication. Note that the space of Hermitian matrices is a subspace if we consider the space of complex matrices with scalar field $\mathbb{R}$.

EXAMPLE 2.11. Let $\mathbb{Q}^{(n)}[x]$ be the space of polynomials over $\mathbb{Q}$ with degree less than or equal to n. This is a subspace of $\mathbb{Q}[x]$. Also, let $W = \{p(x) \in \mathbb{Q}[x] | p(a) = 0\}$ for some $a \in \mathbb{R}$. Then $W \subseteq \mathbb{Q}[x]$.

The following theorem states that the intersection of two subspaces is also a subspace.

THEOREM 2.12. *Let V be a vector space over a field $\mathbb{F}$ and let $U_1 \subseteq V$, $U_2 \subseteq V$ then $U_1 \cap U_2 \subseteq V$*

PROOF. Let $u, v \in U_1 \cap U_2$ then $u + v \in U_1$ and $u + v \in U_2$ since U_1, U_2 are subspaces. Also, if $w \in U_1 \cap U_2$ then for $\alpha w \in U_1$ and $\alpha w \in U_2$ for any $\alpha \in \mathbb{F}$. Hence $w \in U_1 \cap U_2$. Since $U_1 \cap U_2$ is closed under addition and scalar multiplication it is a subspace of V. □

The sum of two subspaces $U_1 + U_2 := \{u_1 + u_2 | u_1 \in U_1,\, u_2 \in U_2\}$ is also a subspace.

THEOREM 2.13. *Let V be a vector space over $\mathbb{F}$. If $U_1, U_2 \subseteq V$ then $U_1 + U_2 \subseteq V$*

PROOF. Let $u, v \in U_1 + U_2$, then $u = u_1 + u_2$ for some $u_1 \in U_1$ and $u_2 \in U_2$ and $v = u_1' + u_2'$ for some $u_1' \in U_1$ and $u_2' \in U_2$.

$u + v = (u_1 + u_2) + (u_1' + u_2') = (u_1 + u_1') + (u_2 + u_2')$ (Addition is commutative)

Since U_1, U_2 are subspaces $u_1 + u_1' \in U_1$ and $u_2 + u_2' \in U_2$. Therefore $u + v \in U_1 + U_2$. Also, if $u \in U_1 + U_2$ then $u = u_1 + u_2$ for some $u_1 \in U_1$ and $u_2 \in U_2$ then for any $\alpha in \mathbb{F}$

$$\alpha u = \alpha(u_1 + u_2) = \alpha u_1 + \alpha u_2$$

But, $\alpha u_1 \in U_1$ and $\alpha u_2 \in U_2$ because U_1, U_2 are subspaces. Therefore $u \in U_1 + U_2$. Since $U_1 + U_2$ is closed under addition and scalar multiplication it is a subspace of V. □

DEFINITION 2.14. (Direct Sum) Let V be a vector space and U_1, U_2 be subspaces of V, then $U_1 + U_2$ is called a direct sum of U_1 and U_2 if every vector in $U_1 + U_2$ can be written uniquely as a vector in U_1 and a vector in U_2.

We shall denote the direct sum of U_1 and U_2 as $U_1 \oplus U_2$. The next example clarifies the difference between an ordinary sum and a direct sum of subspaces.

EXAMPLE 2.15. $U_1 = \{(x, y, 0) | x, y \in \mathbb{R}\}$ and $U_2 = \{(0, w, z) | w, z \in \mathbb{R}\}$ are subspaces of $\mathbb{R}^3$then $U_1 + U_2 = \mathbb{R}^3$ since any vector $(a_1, a_2, a_3) \in \mathbb{R}^3$ can be written as

$$(a_1, a_2, a_3) = \overbrace{(a_1, a_2, 0)}^{U_1} + \overbrace{(0, 0, a_3)}^{U_2}$$

But, $\mathbb{R}^3$ is not a direct sum of U_1 and U_2 since

$$(a_1, a_2, a_3) = \overbrace{(a_1, a_2 - x, 0)}^{U_1} + \overbrace{(0, x, a_3)}^{U_2}$$

If $x \neq 0$ then this is another way to write the vector (a_1, a_2, a_3) as a sum of vectors in U_1 and U_2.

EXAMPLE 2.16. If $U_1 = \{(x, y, 0) | x, y \in \mathbb{R}\}$ and $U_2 = \{0, 0, z) | z \in \mathbb{R}\}$ then $\mathbb{R}^3 = U_1 \oplus U_2$. It is left to the reader to check that this is indeed a direct sum.

EXAMPLE 2.17. Consider the space of all functions $F = \{f : \mathbb{R} \to \mathbb{R}\}$. As previously observed, this is a vector space. Let $F^o = \{f \in F | f(x) = f(-x)\}$ be all the even functions and $F^o = \{f \in F | f(x) = f(-x)\}$ be all the odd functions. One can show that F^o, F^e are subspaces of F. In fact $F = F^e \oplus F^o$.

The next theorem gives an easy way to check if a vector space is a direct sum of subspaces.

THEOREM 2.18. *If U_1 and U_2 are subspaces of a vector space V then $V = U_1 \oplus U_2$ iff i)$V = U_1 + U_2$ ii) $U_1 \cap U_2 = 0$.*

PROOF. $\Rightarrow$ (i) is true by definition. Suppose there exist $u \in U_1 \cap U_2$ s.t. $u \neq 0$. Now, if $v \in U_1 \oplus U_2$, then $v = u_1 + u_2$ for some $u_1 \in U_1$ and $u_2 \in U_2$. But $v = v_1 + v_2$ where $v_1 = (u_1 + u) \in U_1$ and $v_2 = (u_2 - u) \in U_2$. This contradicts the fact that there should be unique way to represent v.
$\Leftarrow$ Suppose conditions i) and ii) hold and assume that $v \in V$ can be written as

$$v = u_1 + u_2 = u_1' + u_2'$$

then

$$\overbrace{u_1 - u_1'}^{U_1} = \overbrace{u_2 - u_2'}^{U_2} = 0$$

Last equality is because condition ii) states $U_1 \cap U_2 = 0$. Hence $u_1 = u_1', u_2 = u_2'$. Hence the representation of v is unique. Therefore $V = U_1 \oplus U_2$. ☐

2.2. Span,linear independence, basis

DEFINITION 2.19. (Span) Let V be a vector space over $\mathbb{F}$, then the span of vectors $(v_1, v_2, ..., v_n) \in V$ is

$$\text{span}(v_1, v_2, ..., v_n) = \{\sum_{i=1}^{n} a_i v_i | a_i \in \mathbb{F}\}$$

LEMMA 2.20. *span$(v_1, v_2, ..., v_n)$ is a subspace of V.*

PROOF. Let $u, v \in \text{span}(v_1, v_2, ..., v_n)$ then there exist $\{a_i, b_i \in \mathbb{F}\}$ such that

$$u = \sum_{i=1}^{n} a_i v_i, \quad v = \sum_{i=1}^{n} b_i v_i$$

then

$$u + v = \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{n} b_i v_i = \sum_{i=1}^{n} (a_i + b_i) v_i \in F$$

If $\alpha \in \mathbb{F}$ then,

$$\alpha u = \alpha(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} \alpha a_i v_i \in \mathbb{F}$$

Therefore, span$(v_1, v_2, ..., v_n)$ is a subspace. ☐

DEFINITION 2.21. (Linear independence) A set of vectors $\{v_1, v_2, ..., v_n\} \in V$ is called linearly independent if the only solution to $a_1 v1 + a_2 v2 + \cdots + a_n v_n = 0$ is $a_i = 0 \; \forall i$.

Otherwise the set is linearly dependent, that is, at least some $a_i's$ that are non zero such that the equation is satisfied.

EXAMPLE 2.22. The vectors $e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)$ span $\mathbb{R}^3$ since any vector $v = (a,b,c)$ in $\mathbb{R}^3$ can be written as a linear combination $ae_1+be_2+ce_3$. The set $\{e_1, e_2, e_3\}$ is also linearly independent since if $a_1e_1 + a_2e_2 + a_3e_3 = 0$ then this implies $a_1 = a_2 = a_3 = 0$.

EXAMPLE 2.23. In the vector space of polynomials with rational coefficient with degree less than or equal to n, $\mathbb{Q}^{(n)}[x]$ the set $\{1, x, x^2, ...x^n\}$ is a linearly independent set. $\text{span}(1, x, x^2..., x^n) = \mathbb{Q}^{(n)}[x]$.

EXAMPLE 2.24. In $\mathbb{R}^3$ the vectors $v_1 = (1,1,1), v_2 = (0,1,-1), v_3 = (1,2,0)$ are not linearly independent, Since if

$$a_1(1,1,1) + a_2((0,1,-1) + a_3(1,2,0) = (0,0,0)$$

Thus,

$$a_1 + a_3 = 0, a_1 + a_2 + 2a_3 = 0, a_1 - a_2 = 0$$

which leads to $a_1 = a_2 = -a_3$. Hence $a_1 = a_2 = 1$ and $a_3 = -1$ is a non-zero solution to $a_1v_1 + a_2v_2 + a_3v_3 = 0$.

EXAMPLE 2.25. In $M_2(\mathbb{R})$ the matrices $v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, v_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ is an linearly independent set. Moreover, $span(v_1, v_2, v_3, v_4) = M_2(\mathbb{R})$.

THEOREM 2.26. *If $v_1, v_2, ...v_n$ is a linearly independent set of vectors then any vector in $span(v_1, v_2, ..., v_n)$ can be written uniquely as a linear combination of $v_1, v_2, ..., v_n$.*

PROOF. Let $v \in \text{span}(v_1, v_2, ..., v_n)$ then

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n \tag{2.1}$$

for some scalars $a_1, a_1, ..., a_n$. Suppose

$$v = b_1v_1 + b_2v_2 + \cdots + b_nv_n \tag{2.2}$$

Subtracting equation (2.1) from (2.2) we get

$$0 = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n$$

Since $(v_1, v_2, ..., v_n)$ is linearly independent it follows that $a_i = b_i$ for each i. □

Next we will work our way towards showing three important theorems i) In a vector space any two basis have the same number of elements ii) Any spanning set of a vector space can be reduced to a basis (Basis Reduction)iii) Any linearly independent set (not necessarily spanning set) can be extended to a basis (Basis Extension). Towards, these we will make use some lemmas.

LEMMA 2.27. *Let $(v_1, v_2, ..., v_n)$ be a linearly dependent set such that $v_1 \neq 0$ then there exists an index $j \in \{2, 3, ..., n\}$ such that*
i)$v_j \in span(v_1, v_2, ..., v_{j-1})$
ii) $span(v_1, v_2, ..., \hat{v}_j, v_{j+1}, ..., v_n) = span(v_1, v_2, ...v_n)$

PROOF. Since $(v_1, v_2, ..., v_n)$ is linearly dependent $\exists a_1, a_2, ..., a_m$ not all zero such that $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Choose j to be the largest index such that $a_j \neq 0$, then

$$a_1v_1 + a_2v_2 + \cdots a_jv_j = 0 \implies \tag{2.3}$$

$$v_j = -\frac{a_1}{a_j}v_1 - \frac{a_2}{a_j}v_2 - \cdots - \frac{a_{j-1}}{a_j}v_{j-1} \tag{2.4}$$

Let $u \in \text{span}(v_1, v_2, ..., v_n)$ then

$$u = b_1v_1 + b_2v_2 + \cdots + b_nv_n$$

Substituting v_j from equation (2.3) we see that one can write u as a linear combination of $(v_1, v_2, ..., \hat{v}_j, v_{j+1}, ..., v_n)$. Hence

$$\text{span}(v_1, v_2, ..., \hat{v}_j, v_{j+1}, ..., v_n) = \text{span}(v_1, v_2, ...v_n)$$

$\square$

In example (2.24) $v_1 = (1, 1, , 1), v_2 = (0, 1, -1)$ and $v_3 = (1, 2, 0)$ in $\mathbb{R}^3$ is a linearly dependent set since $v_1 + v_2 - v_3 = 0$. $v_3 \in \text{span}(v_1, v_2$ and $\text{span}(v_1, v_2, v_3) = \text{span}(v_1, v_2)$.

LEMMA 2.28. *Let V be a vector space and let $v_1, v_2, ..., v_m$ be a linearly independent set that spans V and $w_1, w_2, ..., w_n$ be any other set that spans V then $m \leq n$*

PROOF. The set $(v_1, w_1, w_2, ..., w_n)$ is linearly dependent with $\text{span}(v_1, w_1, w_2, ...w_n) = V$. Applying lemma (2.27) we see that there exists and index i_i such that

$$\text{span}(v_1, ..., \hat{w}_{i_1}, ...w_n) = \text{span}(v_1, w_1, ..., w_n) = V$$

Call the set $\text{span}(v_1, ..., \hat{w}_{i_1}, ...w_n)$ of n elements as S_1. Iteratively construct sets $S_1, S_2, ..., S_k$ by adding $v_1, v_2, ..., v_k$ respectively and removing $w_{i_1}, w_{i_2}, ..., w_{i_k}$. If all $w_i's$ are not removed by step $k = m$ then we have $S_k = (v_1, v_2, ..., v_m, w_{m+1}, ...w_n)$ and obviously in this case we have $m \leq n$. On the other hand if all $w_i's$ are removed by step k then we will have $\text{span}(v_1, v_2, ..., v_k) = V$ with $k \leq m$. If $k < m$ then we arrive at a contradiction that $\text{span}(v_1, v_2, ..., v_k) = V$ but $v_{k+1} \in V$ does not belong to $\text{span}(v_1, v_2, ..., v_k)$ (If it did that would contradict the linear independence of $(v_1, v_2, ..., v_n)$). Hence $n \geq k \geq m$. Hence $n \leq m$. $\square$

DEFINITION 2.29. (Basis) A set $v_1, v_2, ...v_n$ is called a basis of a vector space V if
i)$\text{span}(v_1, v_2, ..., v_n) = V$.
ii)$v_1, v_2, ..., v_n$ is linearly independent.

THEOREM 2.30. *If $v_1, v_2, ..., v_n$ is a basis of a vector space V and $w_1, w_2, ..., w_m$ is another basis of V then $m = n$.*

PROOF. We apply theorem (2.28) to show that $n \leq m$ and in the other direction to show $m \leq n$. Hence $n = m$. $\square$

DEFINITION 2.31. (Dimension) If V is a vector space then any two basis have the same number of elements and this number is called the dimension of the vector space.

Note: The dimension of a vector space can be infinite. For example $\mathbb{Q}[x]$ has basis $\{1, x, x^2, ...\}$ is an infinite dimensional vector space.

LEMMA 2.32. *If $\{u_1, u_2, \dots, u_k\}$ is a linearly independent set then if $u_{k+1} \notin$ span$(\{u_1, u_2, \dots, u_k\})$ then $\{u_1, u_2, \dots, u_k, u_{k+1}\}$ is linearly independent.*

PROOF. If $a_1u_1 + a_2u_2 + \cdots a_{k+1}u_{k+1} = 0$ then $a_{k+1} = 0$, otherwise we can divide by a_{k+1} to get $u_{k+1} = -\frac{a_1}{a_{k+1}}u_1 - \frac{a_2}{a_{k+1}}u_2 - \cdots - \frac{a_k}{a_{k+1}}u_k$. This contradicts the fact that $u_{k+1} \notin \text{span}(u_1, u_2, ..., u_k)$. Hence $a_{k+1} = 0$ and due to the linear independence of $(u_1, u_2, ..., u_k)$ we also get $a_1 = a_2 = \cdots = a_k = 0$. $\square$

The following theorem states that if we have a spanning set then we can always extract a linearly independent set out of it.

THEOREM 2.33. *(Basis reduction) If $V = span(v_1, v_2, ..., v_n)$ be a vector space, then either $(v_1, v_2, ..., v_n)$ is a basis of V of some v_i can be removed to obtain a basis of V.*

PROOF. If $(v_1, v_2, ..., v_n)$ is a linearly independent set then we are done. If not, then we follow the following procedure. Initially set $j = 1$
Step 1: If $v_j = 0$ then remove v_j
Step 2: If $v_{j+1} \in \text{span}(v_1, v_2, ..., v_j)$ then remove v_{j+1}; If $j \neq n$ then $j = j + 1$; Go to Step 1 else output the list of remaining vectors.
The final list spans V since a vector was discarded only if it was in the span of the previous vectors. Also, since no vector is in the span of the previous vectors by lemma 2.32 we get a set of linearly independent vectors. $\square$

THEOREM 2.34. *Every linearly independent set can be extended to a basis of V.*

PROOF. Let $(v_1, v_2, ..., v_m)$ be a linearly independent set. Let $(w_1, w_2, ..., w_n)$ be a basis of V. We do the following procedure. Let $\mathcal{S} = (v_1, v_2, ..., v_m)$
Step 1: If $w_1 \in \text{span}(v_1, v_2, ..., v_m)$ then $\mathcal{S} = (v_1, v_2, ..., v_m)$ otherwise $\mathcal{S} = (v_1, v_2, ..., v_m, w_1)$.
Step k: If $w_k \in \text{span}(\mathcal{S})$ then leave $\mathcal{S}$ unchanged, otherwise adjoin w_k to $\mathcal{S}$.
i) By lemma 2.32 after each step the list $\mathcal{S}$ is linearly independent.
ii)After n-steps $w_k \in \text{span}(\mathcal{S})$ for all $k = 1, 2, ..., n$. Since $(w_1, w_2, ..., w_n)$ was as spanning set therefore $\mathcal{S}$ spans V.
By arguments i) and ii) we see that $\mathcal{S}$ is a basis of V. $\square$

EXAMPLE 2.35. Given that the in $\mathbb{R}^3$ the set of vectors

$$\mathcal{S} = \{(1, -1, 0), (2, -2, 0), (-1, 0, 1), (0, -1, 1), (0, 1, 0)$$

forms a spanning set we can use Basis Reduction procedure to obtain a basis:

$$\mathcal{B} = \{(1, -1, 0), (-1, 0, 1), (0, 1, 0)$$

EXAMPLE 2.36. Given the linearly independent set

$$\mathcal{S} = \{v_1 = (1, 1, 0, 0), v_2 = (1, 0, 1, 0)\}$$

in $\mathbb{R}^4$ we can use the Basis extension procedure to extend this set to a basis. We know that $e_1 = (1, 0, 0, 0), e_2 = (0, 1, 0, 0), e_3 = (0, 0, 1, 0), e_4 = (0, 0, 0, 1)$ span $\mathbb{R}^4$. Using the extension procedure we get $(v_1, v_2, e_1, e_4$ as the extended basis of $\mathbb{R}^4$.

EXERCISE 2.37. Show that if V is a vector space and $\dim(V) = n$, then any set of n vectors that span V are linearly independent

EXERCISE 2.38. Show that if V is a vector space and $\dim(V) = n$, then any set of n vectors that are linearly independent also span V.

THEOREM 2.39. *Let U and W be subspaces of a vector space V then*

$$dim(U+W) = dim(U) + dim(W) - dim(U \cap W)$$

PROOF. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of of $U \cap W$. By the basis extension theorem we can extend this to $\{v_1, v_2, \ldots, v_n, u_1, u_2, \ldots, u_k\}$ to obtain a basis of U. Similarly we can extend the basis of $U \cap W$ as $\{v_1, v_2, \ldots, v_n, w_1, w_2, \ldots, w_l\}$ to obtain a basis of W.
Claim:

$$\mathcal{S} = \{v_1, v_2, \ldots, v_n, u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_l\}$$

is a basis of $U + W$.
If we show the claim then a simple counting argument leads to the proof of the main theorem. It is clear that $\mathcal{S}$ spans $U + W$, since it contains a basis of U and W. To show linear independence, let

$$\begin{aligned} \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{k} b_i u_i + \sum_{i=1}^{l} c_i w_i &= 0 \quad \text{then,} \\ \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{k} b_i u_i + &= -\sum_{i=1}^{l} c_i w_i \end{aligned}$$

The L.H.S. of the last equality is an element of U while the R.H.S. is an element of of W hence this vector has to be an element of $U \cap W$ and therefore,

$$\begin{aligned} \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{k} b_i u_i &= \sum_{i=1}^{n} a'_i v_i \quad \text{So,} \\ \sum_{i=1}^{n} (a_i - a'_i) v_i + \sum_{i=1}^{k} b_i u_i &= 0 \end{aligned}$$

But since $\{v_1, v_2, \ldots, v_n, u_1, u_2, \ldots, u_k\}$ is a linearly independent set (It is a basis for U) therefor we get

$$a_i = a'_i \quad b_i = 0 \quad \text{for all i}$$

Since $b_i = 0$ for all i we get

$$\sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{l} c_i w_i = 0$$

But since $\{v_1, v_2, \ldots, v_n, w_1, w_2, \ldots, w_l\}$ is also linearly independent (It is a basis for W), we get that all a_i and c_i are 0. Therefore $\mathcal{S}$ is a basis of $U + W$ □

EXAMPLE 2.40. Consider the following subspaces of $M_2(\mathbb{R})$

$$U = \{\begin{pmatrix} x & y \\ z & 0 \end{pmatrix} | x, y, z \in \mathbb{R}\} \quad W = \{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} | x, y, z \in \mathbb{R}\}$$

$dim(U) = 3$, $dim(W) = 2$ and $U \cap W = \{\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} | x \in \mathbb{R}\}$, so $dim(U \cap W) = 1$. So by theorem 2.39 we see that $dim(U + W) = 3 + 2 - 1 = 4$. Hence $U + W = M_2(\mathbb{R})$.

2.3. Linear transformations

DEFINITION 2.41. Let V, W be vector spaces over a field $\mathbb{F}$, then a map $T : V \to W$ is a linear map from V to W if
i)$T(v_1 + v_2) = T(v_1) + T(v_2)\, \forall v_1, v_2 \in V$
ii)$T(\alpha v) = \alpha T(v)\, \forall v \in V$ and $\alpha \in \mathcal{F}$

EXAMPLE 2.42. Zero map $O : V \to W$ that maps every vector in V to the zero vector of W can easily verified to be a linear map

EXAMPLE 2.43. $\mathbb{1} : V \to V$, the identity map that maps every vector in V to itself is a linear map.

EXAMPLE 2.44. $D : \mathbb{Q}[x] \to \mathbb{Q}[x]$, be the differentiation map $D(p(x)) = p'(x)$. The differentiation map is linear since $D(p(x) + q(x)\frac{d}{dx}(p(x) + q(x)) = \frac{d(p(x))}{dx} + \frac{d(q(x))}{dx}$.

EXAMPLE 2.45. Let $C[0,1]$ be the vector space of real valued continuous functions on the interval $[0,1]$. $I : C[0,1] \to C[0,1]$, be the integration map $I(f(x)) = \int_0^x f(u)du$. The integration map is linear since

$$\begin{aligned} I(f(x) + g(x)) &= \int_0^x f(u) + g(u)du = \int_0^x f(u)du + int_0^x g(u)du. \\ I(\alpha f(x)) &= \int_0^x \alpha f(u)du = \alpha \int_0^x f(u)du = \alpha I(f(x)) \end{aligned}$$

EXAMPLE 2.46. In the vector space $\mathbb{R}^2$ the rotation of a vector by angle θ map that takes $(x, y) \in \mathbb{R}^2$ to $(x\cos(\theta) - y\sin(\theta), x\sin(\theta) + y\sin(\theta)$ is a linear map.

EXAMPLE 2.47. In the vector space $\mathbb{R}^3$ the projection map of a vector (x, y, z) on to the $x - y$ plane that takes it to $(x, y, 0) \in \mathbb{R}^3$ is a linear map.

EXAMPLE 2.48. The reflection of a vector by angle about the $x - y$ plane is a linear map that takes $(x, y, z) \in \mathbb{R}^3$ to $(x, y, -z) \in \mathbb{R}^3$

We will study more about rotations, projections and reflections in the next chapter. Given $T : V \to W$ a linear operator it induces two important subspaces.

DEFINITION 2.49. (Null space/Kernel) Let $T : V \to W$ be a linear map then $\mathrm{Null}(T) = \{v \in V : T(v) = 0\}$.

DEFINITION 2.50. (Range) $\mathrm{Range}(T) = \{w \in W : \exists v \in V\, s.t. Tv = w\}$

THEOREM 2.51. *$T : V \to W$ is a linear transformation then Null(T) and Range(T) are subspaces of V and W respectively.*

PROOF. Left as an exercise. □

The $\dim(\mathrm{Null}(T))$ is called the <u>nullity</u> of T and $\dim(\mathrm{Range}(T))$ is called the <u>rank</u> of T

THEOREM 2.52. *(Rank-Nullity) Let $T : V \to W$ be a linear transformation the*

$$Rank(T) + Nullity(T) = dim(V)$$

PROOF. Let $dim(V) = n$ and let $(v_1, v_2, \ldots, v_k)$ be a basis of Null(T). By the basis extension theorem we can extend this so that $(v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_{n-k})$ is a basis of V.
Claim: $(Tw_1, Tw_2, \ldots, Tw_{n-k})$ is a basis of Range(T).
Let $w \in$ Range(T), then there exists a $v \in V$ such that $Tv = w$. Expanding v in the basis $(v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_{n-k})$ we have,

$$\begin{aligned} v &= a_1v_1 + a_2v_2 + \cdots + a_kv_k + b_1w_1 + b_2w_2 + \cdots b_{n-k}w_{n-k} \\ & \text{Therefore,} \\ Tv &= T(a_1v_1 + a_2v_2 + \cdots + a_kv_k + b_1w_1 + b_2w_2 + \cdots b_{n-k}w_{n-k}) \\ &= a_1T(v_1) + a_2T(v_2) + \cdots + a_kT(v_k) + b_1T(w_1) + b_2T(w_2) + \cdots b_{n-k}T(w_{n-k}) \\ &= 0 + b_1T(w_1) + b_2T(w_2) + \cdots b_{n-k}T(w_{n-k}) = w \end{aligned}$$

Therefore $w \in$ span$(Tw_1, Tw_2, \ldots, Tw_k)$.

Moreover,

$$\begin{aligned} c_1T(w_1) + c_2T(w_2) + \cdots + c_{n-k}Tw_{n-k} &= 0, \text{ then} \\ T(c_1w_1 + c_2w_2 + \cdots + c_{n-k}w_{n-k}) &= 0 \implies \\ c_1w_1 + c_2w_2 + \cdots + c_{n-k}w_{n-k} &\in \text{Null}(T) \implies \\ c_1w_1 + c_2w_2 + \cdots + c_{n-k}w_{n-k} &= d_1v_1 + d_2v_2 + \cdots + c_kd_k \text{ for some } d_i \end{aligned}$$

Therefore, $c_1w_1 + c_2w_2 + \cdots + c_{n-k}w_{n-k} - d_1v_1 - d_2v_2 - \cdots - d_kv_k = 0$. Since $(v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_{n-k})$ is a linearly independent set, we get that $c_i's = 0$ and $d_j's = 0$ for $i = 0..n-k$ and $j = 0..k$. Hence $(Tw_1, Tw_2, \cdots Tw_{n-k})$ is a linearly independent set that spans Range(T). So Nullity$(T) = k$ and Range$(T) = n - k$. Hence the theorem. $\square$

2.4. Matrix representation and Inverse of a linear operator

Let $T : V \to W$ be a linear transformation from a vector space V to W. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ be a basis of V. It is easy to see that T is completely determined by its action on the basis $\mathcal{B}$. Indeed, if $v \in V$, then since

$$\begin{aligned} v &= \sum_{i=1}^{m} a_iv_i, \quad \text{therefore,} \\ T(v) &= \sum_{i=1}^{m} a_iT(v_i) \end{aligned}$$

Now if $T(v_k) = u_k \in W$, then we can expland u_k in terms of a basis $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ of W. Thus, we can write

$$T(v_j) = \sum_{i=1}^{n} a_{ij}w_i \quad \text{for } j \in (0..m)$$

$$[T]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{12} & \cdots & a_{nm} \end{pmatrix}$$

is the matrix representation of the linear transformation T in the basis $\mathcal{B}$ of V and $\mathcal{B}'$ of W

EXAMPLE 2.53. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear operator given by $T(x,y) = (ax+by, cx+dy)$ then to write the matrix reprentation of T in the standard basis $\mathcal{B} = \mathcal{B}' = (1,0),(0,1)$ we see that

$$\begin{aligned} T(1,0) &= (a,c) = a(1,0) + c(0,1) \\ T(0,1) &= (b,d) = b(1,0) + d(0,1) \end{aligned}$$

Thus,

$$[T]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

EXAMPLE 2.54. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear map given by $T(x,y) = (y, x+2y, x+y$. Let $\mathcal{B} = \{(1,0),(0,1)$, $\mathcal{B}'' = \{(1,0,0),(0,1,0),(0,0,1)\}$ and $\mathcal{B}' = \{(1,2),(0,1)\}$, then

$$[T]_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad [T]_{\mathcal{B}',\mathcal{B}''} = \begin{pmatrix} 2 & 1 \\ 5 & 2 \\ 3 & 1 \end{pmatrix}$$

Let $B : V \to U$ be a linear transformation pace V to a vector space U, and let $A : U \to W$ be a linear transformation such that $\mathrm{dom}(A) \subseteq \mathrm{range}(B)$. Let U, V, W be finite dimensional vector spaces and let $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$, $\mathcal{B}'' = \{u_1, u_2, \ldots, u_l\}$ and $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ be basis for V, U and W respectively. Is $AB : V \to W$ linear? What is the matrix representation of the the linear transformation AB in the basis $\mathcal{B}, \mathcal{B}'$? Clearly AB is linear since

$$\begin{aligned} AB(x+y) &= A(Bx + By) = AB(x) + AB(y) \\ AB(\alpha x) &= A(\alpha B(x)) = \alpha AB(x) \end{aligned}$$

follows from the linearity of A and B. Now,

$$\begin{aligned} B(v_j) &= \sum_{k=1}^{l} b_{kj} u_k \\ A(u_k) &= \sum_{i=1}^{n} a_{ik} w_i \end{aligned}$$

Therefore,

$$\begin{aligned} AB(v_j) &= \sum_{i=1}^{l} b_{kj} A(u_k) \\ AB(v_j) &= \sum_{i=1}^{l} b_{kj} \sum_{i=1}^{n} a_{ik} w_i \\ AB(v_j) &= \sum_{i=1}^{n} (\sum_{k=1}^{l} a_{ik} b_{kj}) w_k \end{aligned}$$

So the matrix representation of the linear transformation $AB : V \to W$ in the basis $\{\mathcal{B}, \mathcal{B}'\}$ is given by the matrix entries

$$(AB)_{ij} = \sum_{k=1}^{l} a_{ik} b_{kj}$$

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{il} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \dots & \dots & b_{1j} & \dots \\ \dots & \dots & b_{2j} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \dots & \dots & b_{lj} & \dots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \dots & \dots & \sum_{k=1}^{l} a_{ik}b_{kj} & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Definition 2.55 (Inverse). Let $A : V \to W$ be a linear operator, then if there exists a linear operator $B : W \to V$ such that $AB = \mathbb{1}_V$ and $BA = \mathbb{1}_W$, then B is called the inverse of A.

The inverse of A is denoted as A^{-1}.

Exercise 2

Show that the inverse of a linear mapping is unique

Theorem 2.56. *A linear map $A : V \to W$ is invertible iff it is one-one (injective) and onto (surjective)*

Proof. $\Longrightarrow$ Let $A : V \to W$ be invertible, then if $A(x) = A(y)$, applying A^{-1} to both sides we get $A^{-1}A(x) = A^{-1}A(y)$, so $x = y$ and hence A is one-one. Let $y \in V$, then since $A^{-1} : V \to V$ exists and is a linear operator on V, let $A^{-1}(y) = x$, then $AA^{-1}y = Ax$, that is $Ax = y$, so A is surjective.

$\Longleftarrow$ Let $A : V \to W$ be injective and surjective. Since A is surjective for each $w \in W$ there is a $v \in V$ such that $Av = w$. Moreover, since A is injective this v is uniquely determined. Define $B : W \to V$ such that $Bw = v$, the $ABw = Av = w$. So $AB = \mathbb{1}_W$. Since $Av = w$, $BAv = Bw = v$, so $BA = \mathbb{1}_V$. Hence B is the inverse of A. □

Theorem 2.57. *If $T : V \to W$ is invertible then T^{-1} is a linear mapping from W to V*

Theorem 2.58. *For $w_1, w_2 \in W$*

$$T(T^{-1}w_1 + T^{-1}w_2) = T(T^{-1}w_1) + T(T^{-1}w_1) = w_1 + w_2$$

Hence if $v = ((T^{-1}w_1 + T^{-1}w_2)) \in V$ then $Tv = w_1 + w_2$. Hence,

$$T^{-1}(w_1 + w_2) = v = (T^{-1}w_1 + T^{-1}w_2)$$

Similarly, for $w \in W$

$$T(\alpha T^{-1}(w)) = \alpha T(T^{-1}w) = \alpha w$$

Hence $T^{-1}(\alpha w) = \alpha T^{-1}(w)$.

Definition 2.59. Two vector spaces V and W are isomorphic if there exists an invertible linear mapping $T : V \to W$

Theorem 2.60. *Two finite dimensional vector spaces V and W over a field $\mathbb{F}$ are isomorphic iff $dim(V) = dim(W)$*

Theorem 2.61. *Let $A : V \to V$ be a linear operator on a finite dimensional vector space V, then the following are equivalent*
1.A linear map $A : V \to V$ is invertible
2.A is one-one (injective)
3.A is onto (surjective)

PROOF. 1 $\implies$ 2 If A is invertible then $A^{-1} : V \to V$ exists. If $Ax = Ay$ then $A^{-1}Ax = A^{-1}Ay$ so $x = y$ and A is injective.
2 $\implies$ 3 Let $A(x) = 0 = A(0)$, then since A is injective this implies that $x = 0$. So, injectivity implies that $\ker(A) = 0$. By the rank nullity theorem $\text{rank}(A) = dim(V)$. Since $\text{range}(A) \subseteq V$, so $\text{range}(A) = V$ and A is surjective.
3 $\implies$ 1 From the rank-nullity theorem it follows that $\text{nullity}(A) = 0$. Hence A is injective and surjective. Hence A is invertible. □

THEOREM 2.62. *Let $A : V \to V$ be a linear operator on a finite dimensional vector space V, then A is invertible $\iff$ The matrix representation of A in some basis $\mathcal{B}$, $[A]_{\mathcal{B}}$, has linearly independent columns*

PROOF. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ then $[A]_{\mathcal{B}} = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_n \end{bmatrix}$
$\implies$ If A is invertible then by theorem 2.61 A has full rank which is equal to $\dim(V) = n$. Therefore, the n columns of $[A]_{\mathcal{B}}$, $Av_1, Av_2, \ldots, Av_n$ span V. Therefore the columns are linearly independent
$\impliedby$ If the columns of $[A]_{\mathcal{B}}$ are linearly independent then
□

EXAMPLE 2.63. Let $A : \mathbb{R}^2 \to \mathbb{R}^2$ is such that $A(x, y) = (2x + y, x - y)$. Let

$$u = 2x + y \quad v = x - y$$

Solving for x and y we get

$$x = \frac{u + v}{3} \quad y = \frac{u - 2v}{3}$$

Hence the inverse of A is $A^{-1}(x, y) = (\dfrac{x + y}{3}, \dfrac{x - 2y}{3})$

The matrix representation of A and A^{-1} in the basis $\{(1, 0), (0, 1)\}$ is

$$A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \quad A^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{pmatrix}$$

2.5. Change of Basis

Consider a basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ of a vector space V. Let $x \in V$ then $\exists$ scalars α_i such that

$$x = \sum_i \alpha_i \alpha_i \tag{2.5}$$

The co-ordinate representation of the the vector x in the basis $\mathcal{B}$ is given as

$$[x]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

If $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ is another basis of V and let's suppose we are interested in the coordinate representation of the vector x in basis $\mathcal{B}'$.

$$w_j = \sum_{i=1}^{n} P_{ij} v_i$$

Now if

$$x = \sum_{j=1}^{n} \alpha'_j w_j$$

then, we get

$$\begin{aligned} x &= \sum_{j=1}^{n} \alpha'_j \sum_{i=1}^{n} P_{ij} v_i \\ x &= \sum_{i=1}^{n} (\sum_{j=1}^{n} P_{ij} \alpha'_j) v_i \end{aligned}$$

Comparing with equation (2.5) we get

$$\alpha_i = (\sum_{j=1}^{n} P_{ij} \alpha'_j) \quad \text{for each } i$$

This is just the matrix equation

$$[x]_{\mathcal{B}} = P[x]_{\mathcal{B}'}$$

Moreover the columns of the matrix P are $w_i = Pv_i$. Since w_i is a basis of V from theorem (2.62) we see that P is invertible. Hence we get

$$[x]_{\mathcal{B}'} = P^{-1}[x]_{\mathcal{B}}$$

EXAMPLE 2.64. Let $\mathcal{B} = \{e_1, e_2\}$ be the standard basis in $\mathbb{R}^2$. Let $\mathcal{B}' = \{e'_1, e'_2\}$ be another basis that is obtained from $\mathcal{B}$ by rotation anticlockwise by an angle θ. We see that

$$\begin{aligned} e'_1 &= \cos\theta e_1 + \sin\theta e_2 \\ e'_2 &= -\sin\theta e_1 + \cos\theta e_2 \end{aligned}$$

Thus

$$P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and

$$P^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

Thus if $[x]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ then $[x]_{\mathcal{B}'} = P^{-1}[x]_{\mathcal{B}} = \begin{pmatrix} \cos\theta\alpha_1 + \cos\theta\alpha_2 \\ -\sin\theta\alpha_1 + \cos\theta\alpha_2 \end{pmatrix}$

EXAMPLE 2.65. Let $\mathcal{B} = \{e_1, e_2, e_3\}$ be the standard basis of $\mathbb{R}^3$ and let basis $\mathcal{B}' = \{e'_1, e'_2, e'_3\}$ be another basis where $e_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 5 \\ -3 \\ 8 \end{pmatrix}$ then

$$P = \begin{pmatrix} -1 & 4 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$P^{-1} = \begin{pmatrix} -1 & 2 & \frac{11}{8} \\ 0 & \frac{1}{2} & \frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{pmatrix}$$

A vector x whose representation in in basis $\mathcal{B}$ is $[x]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 2 \\ 8 \end{pmatrix}$ has co-ordinate representation

$$[x]_{\mathcal{B}'} = P^{-1}[x]_{\mathcal{B}} = \begin{pmatrix} -1 & 2 & \frac{11}{8} \\ 0 & \frac{1}{2} & \frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 12 \\ \frac{11}{8} \\ 1 \end{pmatrix}$$

in basis $\mathcal{B}'$

Now consider a linear operator $T : V \to V$ which has some representation as a matrix in a basis $\mathcal{B}$ which is given by $[T]_{\mathcal{B}}$. Let $[x]_{\mathcal{B}}$ be a representation . We have

$$[x]_{\mathcal{B}} = P[x]_{\mathcal{B}'}$$

We also have

$$[Tx]_{\mathcal{B}} = [T]_{\mathcal{B}}[x]_{\mathcal{B}}$$

Therefore,

$$\begin{aligned} [Tx]_{\mathcal{B}} &= P[Tx]_{\mathcal{B}'} \\ [T]_{\mathcal{B}}[x]_{\mathcal{B}} &= P[T]_{\mathcal{B}'}[x]_{\mathcal{B}'} \\ [T]_{\mathcal{B}}[x]_{\mathcal{B}} &= P[T]_{\mathcal{B}'}P^{-1}[x]_{\mathcal{B}} \end{aligned}$$

So we get

$$[T]_{\mathcal{B}} = P[T]_{\mathcal{B}'}P^{-1}$$

And therefore,

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P$$

EXAMPLE 2.66. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear operator that projects onto the vector $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then the matrix representation of T in the basis $\mathcal{B} = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ can be written as follows. We have

$$\begin{aligned} Te_1 &= 1.e_1 + 0.e_2 \\ Te_1 &= 0.e_1 + 0.e_2 \end{aligned}$$

Hence $[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. If $\mathcal{B}' = \{e_1', e_2'\}$ where

$$e_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, e_2' = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

then

$$\begin{aligned} e_1' &= e_1 + e_2 \\ e_2' &= 2e_1 + e_2 \end{aligned}$$

Hence $P = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ and $P^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$ Hence

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix}$$

2.6. System of Linear Equations

Consider a system of m linear equations with n unknowns

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

This can be written as a matrix equation $Ax = b$ where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m \times n} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n \times 1} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}_{m \times 1}$$

We observes that we can do the following operations on a system of linear equations without modifying the solution set

(1) Interchange any two equations
(2) Multiply an equation with a scalar
(3) Adding a scalar multiple of one equation to another

To solve the system of linear equations we look at the augmented matrix $[A|b]$. Performing the above operations is equivalent to performing row operations on the augmented matrix. The goal is to reduce the augmented matrix to a tractable form called the reduced row echelon form (RREF). The RREF of a matrix is as follows:

$$\begin{pmatrix} 1 & * & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(1) All zero entry rows are stacked at the bottom.
(2) The first non-zero entry of any row is strictly to the right of the first non-zero entry of the previous row
(3) The first non-zero entry in every row is 1.
(4) All the entries above the first non-zero entry of a row are zero

If only the first two conditions are satisfied then the matrix is said to be in a row echelon form. The first non-zero entries of a row are called pivots.

Theorem 2.67. *Every matrix can be reduced to RREF that is unique using elementary row operations*

EXAMPLE 2.68. Solve the following system of linear equations:

$$\begin{aligned} x_1 + 2x_2 + 2x_3 + 3x_4 &= 4 \\ 2x_1 + 4x_2 + x_3 + 3x_4 &= 5 \\ 3x_1 + 6x_2 + x_3 + 4x_4 &= 7 \end{aligned}$$

We do row operations to reduce the augmented matrix to RREF form

$$\begin{pmatrix} 1 & 2 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 & 5 \\ 3 & 6 & 1 & 4 & 7 \end{pmatrix} \xrightarrow[R_3 \leftarrow R_3 - 3R_3]{R_2 \leftarrow R_2 - 2R_2} \begin{pmatrix} 1 & 2 & 2 & 3 & 4 \\ 0 & 0 & -3 & -3 & -3 \\ 0 & 0 & -5 & -5 & -5 \end{pmatrix} \xrightarrow[R_2 \leftarrow -\frac{1}{3}R_2]{R_3 \leftarrow -\frac{1}{5}R_3}$$

$$\begin{pmatrix} 1 & 2 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{R_3 \leftarrow R_3 - R_2} \begin{pmatrix} 1 & 2 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1 \leftarrow R_1 - 2R_2}$$

$$\begin{pmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{RREF}$$

Rewriting the equations

$$\begin{aligned} x_1 + 2x_2 + x_4 &= 2 \\ x_3 + x_4 &= 1 \end{aligned}$$

The variables x_1 and x_3 correspond to the pivot elements of the RREF of the augmented matrix and will be called pivot variables. x_2 and x_4 will be called the free variables. Writing the pivot variables in terms of the free variables we get

$$\begin{aligned} x_1 &= 2 - 2x_2 - x_4 \\ x_3 &= 1 - x_4 \end{aligned}$$

The space of solutions is given by

$$\begin{pmatrix} 2 - 2x_2 - x_4 \\ x_2 \\ 1 - x_4 \\ x_4 \end{pmatrix} \qquad x_2, x_3 \in \mathbb{R}$$

$$\begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \qquad x_2, x_4 \in \mathbb{R}$$

EXAMPLE 2.69. Solve the system of linear equations

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 2 \\ x_1 + 2x_2 + 3x_3 &= 1 \\ x_1 + x_3 &= -3 \\ 2x_1 + 4x_2 &= 8 \end{aligned}$$

The augmented matrix and its resutant RREF is

$$\begin{pmatrix} 2 & 4 & 6 & 2 \\ 1 & 2 & 3 & 1 \\ 1 & 0 & 1 & -3 \\ 2 & 4 & 0 & 8 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

In this case we have a unique solution given by $x = \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}$

EXAMPLE 2.70. Solve the following system of linear equations

$$\begin{aligned} x_1 + 2x_2 + x_3 + 3x_4 &= 3 \\ 2x_1 + 4x_2 + 4x_4 &= 4 \\ x_1 + 2x_2 + 3x_3 + 5x_4 &= 5 \\ 2x_1 + 4x_2 + 4x_4 &= 7 \end{aligned}$$

The augmented matrix and its resutant RREF is

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 2 & 4 & 0 & 4 & 7 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The second last row of the RREF of the augmented matrix gives the equation

$$0.x_1 + 0.x_2 + 0.x3 + 0.x4 = 1$$

So there is no solution to this system of linear equations. This system of linear equations will be called inconsistent. Notice that this situation happens when the number of pivot elements of $A <$ number of pivot elements of $[A|b]$.

We can summarize as follows:
A system of linear equations has a solution (Is Consistent) **if the number of pivot of A (in RREF) = number of pivot elements of the augmented matrix** $[A|b]$
A system of linear equations has no solution (Is Inconsistent) **if the number of pivot of A (in RREF) < number of pivot elements of the augmented matrix** $[A|b]$
Moreover a consistent system of linear equations has a **unique solution the number of pivot of A (in RREF) = number of pivot elements of the augmented matrix** $[A|b]$ **= number of unknowns** n.
The number of free variables corresponds to the dimension of the solution space.

EXAMPLE 2.71. Are the vectors $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$ in $\mathbb{R}^3$ linearly independent ?
Assume that

$$a_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + a_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} = 0$$

This is a system of linear equations $Ax = 0$ where $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$, $x = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$. The the set of vectors are linearly dependent if the null space of A has a non-zero vector. To that end we reduce A to RREF.

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

This gives

$$a_1 + a_3 \quad = \quad 0 \quad a_2 - a_3 = 0$$

This gives $a_1 = -a_3$ and $a_2 = a_3$. Therefore the null space of A is $\operatorname{Ker}(A) = \begin{pmatrix} -a_3 \\ a_3 \\ a_3 \end{pmatrix}$ with $a_3 \in \mathbb{R}$. Therefore the Kernel of A is one dimensional span of $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$. Since there is a non-zero vector in the Kernel of A, the three given vectors are linearly dependent.

The row operations to reduce a matrix to RREF correspond to multiplying the matrix (from the left) by elementary matrices. These matrices are of the form $(\mathbb{1} - uv^T)$, where u and v are $n \times 1$ column vectors. The elementary matrices are invertible and one can verify that

$$(\mathbb{1} - uv^T)^{-1} = \mathbb{1} - \frac{uv^T}{v^T u - 1}$$

The three row operations correspond to the following elementary matrices. The elementary matrix E_1 arises by interchanging rows i and j and is given by

$$E_1 = \mathbb{1} - uu^T \quad \text{where}\, u = e_i - e_j$$

' The elementary matrix E_2 arises by multiplying row i by α and is given by

$$E_2 = \mathbb{1} - (1 - \alpha) e_i e_i^T$$

The elementary matrix E_3 arises by adding a multiple (α) of row i to row j and is given by

$$E_3 = \mathbb{1} + \alpha e_i e_j^T$$

Reducing a matrix to its RREF preserves the column relationships between the reduced matrix and the original matrix. Indeed if R is the row echelon form of A then there exists an invertible matrix P such that $PA = R$ (P is a product of invertible elementary matrices and hence is invertible). Denote A_{*j} as the j^{th} column of A. If $A_{*k} = \sum_j \alpha_j A_{*j}$ then since $PA = R$, that is $PA_{*j} = R_{*j}$. Therefore

$$\begin{aligned} R_{*k} &= PA_{*k} = P \sum_j \alpha_j A_{*j} \\ &= \sum_j \alpha_j P A_{*j} = \sum_j \alpha_j R_{*j} \end{aligned}$$

So the same linear relationship exists amongst the columns of the RREF of A. Since $PA = R$, therefore $P^{-1}R = A$ and similarly we can show that if the reduced matrix columns have a certain linear relationship then the same holds for the original matrix A. Since the columns of A span the range of A we can use the RREF to find the basis of the range of A.

EXAMPLE 2.72. Find the basis for the range and kernel of a linear transformation from $A : \mathbb{R}^6 \to \mathbb{R}^4$ given by following matrix

$$A = \begin{pmatrix} 1 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 4 & 4 & 3 & 1 \\ 2 & 2 & 4 & 4 & 2 & 2 \\ 3 & 5 & 8 & 6 & 5 & 3 \end{pmatrix}$$

We first reduce A to RREF

$$A = \begin{pmatrix} 1 & 1 & 2 & 2 & 1 & 1 \\ 2 & 2 & 4 & 4 & 3 & 1 \\ 2 & 2 & 4 & 4 & 2 & 2 \\ 3 & 5 & 8 & 6 & 5 & 3 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The columns corresponding to the pivot variables are linearly independent and the remaining columns are linear combinations of the pivot columns. Hence the basis of the range of A are the columns of A corresponding to the pivot variables. Therefore the basis for range of A is

$$\{ \begin{pmatrix} 1 \\ 2 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \\ 5 \end{pmatrix} \}$$

To determine the kernel of A we neeed to solve the equation $Ax = 0$. Looking at the RREF, the number of free variables is the dimension of the kernel of A. Indeed the RREF leads to the following equivalent equations

$$\begin{aligned} x_1 + x_3 + 2x_4 + x_6 &= 0 \\ x_2 + x_3 + x_6 &= 0 \\ x_5 - x_6 &= 0 \end{aligned}$$

Writing the pivot variables in terms of the free variables we get that the space of solutions is

$$\begin{pmatrix} 1 - x_3 - 2x_4 \\ -x_3 - x6 \\ x_3 \\ x_4 \\ x_6 \\ x_6 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

So the basis for the kernel of A is

$$\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \}$$

EXAMPLE 2.73. Find the Inverse of $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & -1 \\ 2 & 1 & 3 \end{pmatrix}$

If a square matrix A is invertible then a series of of elementary row operations will result in the RREF of A to be the identity matrix. As we saw that row operations are equivalent to multiplying the matrix A from the left by elementary matrix. Thus if a matrix A is invertible then

$$E_{i_k} E_{i_{k-1}} \cdots E_{i_1} A = \mathbb{1}$$

Therfore $A^{-1} = E_{i_k} E_{i_{k-1}} \cdots E_{i_1}$. Therefore A^{-1} can be obtained by doing the series of row operations $E_{i_k} E_{i_{k-1}} \cdots E_{i_1}$ on the identity. Thus to find the inverse of a matrix we write the augmented matrix $[A|\mathbb{1}]$ and do the elementary operations on the augmented matrix till we get the form $[\mathbb{1}|A^{-1}]$. In our case writing the augmented matrix we get

$$\begin{pmatrix} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{3} & \frac{5}{3} & \frac{2}{3} \\ 0 & 1 & 0 & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & 0 & -1 & 0 \end{pmatrix}$$

Hence

$$A^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{5}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -1 & 0 \end{pmatrix}$$

Finally we prove the following theorem

THEOREM 2.74.

$$rank(A) = rank(A^T)$$

2.7. Least Squares

Consider the system of m linear equations with n unknowns given by $Ax = b$. Here A is a $m \times n$ matrix, x is a $n \times 1$ vector and b is a $m \times 1$ vector. If the vector b is not in the range of A then this system does not have a solution (is inconsistent). Given a system of linear equations the associated system of normal equations are formed by multiplying the left and right hand side by A^T, so we get the system of normal equations

$$A^T A x = A^T b \tag{2.6}$$

We will proceed to show that the system given by equation (2.6) is always consistent (even if the original system $Ax = b$ is not).We will also show that if $Ax = b$ is consistent then the solutions of this system is identical to the solutions of the associated normal system (2.6). We will also answer the the question about the solutions of the normal equations (2.6) when $Ax = b$ is not consistent, that is when b is not in the range of A. The solutions to the normal equations will be called the least square solutions corresponding to the system $Ax = b$.

To this end we first prove the following theorem

THEOREM 2.75. *Let $B : V \to U$ and $A : U \to W$ be linear transformations on vector spaces, then*

$$rank(AB) = rank(B) - dim(ker(A) \cap range(B))$$

PROOF. Let $\{x_1, x_2, \ldots, x_s\}$ be a basis of $\ker(A) \cap \mathrm{range}(B)$ and extend this basis by adding the vectors $\{z_1, z_2, \ldots, z_t\}$ so that $\{x_1, x_2, \ldots, x_s, z_1, z_2, \ldots, z_t\}$ is a basis for range(B). We claim that $\{Az_1, Az_2, \ldots, Az_t\}$ is a basis for range(B). We shall first prove that $\{Az_1, Az_2, \ldots, Az_t\}$ spans the range of AB. Indeed, let $u \in \mathrm{range}(AB)$, then there exists a vector $v \in V$ such that $ABv = u$. Now the vector $Bv \in \mathrm{range}(AB)$ so,

$$Bv = \sum_{i=1}^{s} a_i x_i + \sum_{i=1}^{t} b_i z_i$$

Therefore,

$$u = ABv = \sum_{i=1}^{s} a_i A x_i + \sum_{i=1}^{t} b_i A z_i$$

Since each x_i belongs to the $\ker(A)$ we get

$$u = \sum_{i=1}^{t} b_i A z_i$$

and so $\{Az_1, Az_2, \ldots, Az_t\}$ spans range of AB. Next we will show that $\{Az_1, Az_2, \ldots, Az_t\}$ is a linearly independent set. If

$$\begin{aligned} \sum_{i=1}^{t} \alpha_i A z_i &= 0 \\ \sum_{i=1}^{t} A(\alpha_i z_i) &= 0 \end{aligned}$$

So, $\sum_{i=1}^{t} \alpha_i z_i \in \ker(A)$ but since each $z_i \in \mathrm{range}(B)$ so $\alpha_i z_i \in \ker(A) \cap \mathrm{range}(B)$ □

THEOREM 2.76. *Let $B : V \to U$ and $A : U \to W$ be linear transformations on vector spaces, then*

$$\begin{aligned} \mathit{ker}(B) &\subseteq \mathit{ker}(AB) \\ \mathit{range}(AB) &\subseteq \mathit{range}(A) \end{aligned}$$

PROOF. Let $v \in \ker(B)$ then

$$ABv = A(Bv) = A(0) = 0$$

So $v \in \ker(AB)$

Let $u \in \mathrm{range}(AB)$ then there exists $v \in V$ such that $ABv = u$, so $Ax = u$ where $x = Bv$. Hence, $u \in \mathrm{range}(A)$ □

THEOREM 2.77.

$$\begin{aligned} \mathit{range}(A^T A) &= \mathit{range}(A^T) \\ \mathit{ker}(A^T A) &= \mathit{ker}(A) \end{aligned}$$

PROOF. First we claim that $\ker(A^T) \cap \mathrm{range}(A) = 0$. Indeed if $x \in \ker(A^T) \cap \mathrm{range}(A)$ then $A^T x = 0$ and there exists y such that $Ay = x$. So,

$$x^T x = y^T A^T x = y(0) = 0$$

Therefore $x = 0$. Now, due to theorem 2.75 we have

$$\text{rank}(A^T A) = \text{rank}(A) - \dim(\ker(A^T) \cap \text{range}(A))$$

So,

$$\text{rank}(A^T A) = \text{rank}(A) = \text{rank}(A^T)$$

But, due to theorem 2.76 $\text{range}(A^T A) \subseteq \text{range}(A^T)$. So we must have $\text{range}(A^T A) = \text{range}(A^T)$.

Now, by the rank-nullity theorem

$$\text{rank}(A^T A) + \text{nullity}(A^T A) = n$$

Since, $\text{rank}(A^T A) = \text{rank}(A)$

$$\text{rank}(A) + \text{nullity}(A^T A) = n$$
$$n - \text{nullity}(A) + \text{nullity}(A^T A) = n$$

So, $\text{nullity}(A^T A) = \text{nullity}(A)$. Moreover, due to theorem 2.76 $\ker(A) \subseteq \ker(A^T A)$. Hence we get $\ker(A) = \ker(A^T A)$ □

Coming back to the system $Ax = b$ we have the following theorem

Theorem 2.78. *Let $Ax = b$ be a system of equations where A is a $m \times n$ matrix x and b are $n \times 1$ and $m \times 1$ vectors respectively and let $A^T Ax = A^T b$ be the associated normal equations then*
i)$A^T Ax = A^T b$ is always consistent (even if $Ax = b$ is not)
ii)The solutions of $A^T Ax = A^T b$ are ones that minimize the distance from Ax to b that is they minimize the quantity $(Ax-b)^T(Ax-b)$ and are called the least squares solutions to $Ax = b$
iii)If $Ax = b$ is consistent then the set of solutions of $Ax = b$ and $A^T Ax = A^T Ab$ are the same.
iv)If rank$(A) = n$ then $A^T A$ is invertible, and in this case the least squares solution is unique and is given by $x = (A^T A)^{-1} A^T b$

Proof. i)This follows from theorem 2.77 since $\text{rank}(A^T A) = \text{rank}(A^T)$
ii) Let $x \in \mathbb{R}^n$ and let $f(x) = (Ax - b)^T(Ax - b)$. To determine the vector x that minimizes f we use minimization techniques from calculus to differentiate the function

$$f(x_1, x_2, \dots, x_n) = (Ax - b)^T(Ax - b) = x^T A^T Ax - 2x^T A^T b + b^T B$$

We use some simple facts on matrix differentiation. Matrix differentiation is defined in the following way

$$\left[\frac{\partial A}{\partial x}\right]_{ij} := \frac{\partial a_{ij}}{\partial x}$$

and we have the product rule

$$\frac{\partial AB}{\partial x} = \frac{\partial A}{\partial x}B + A\frac{\partial B}{\partial x}$$

So,

$$\frac{\partial f}{\partial x_i} = \frac{\partial x^T}{\partial x_i}A^T Ax + x^T A^T Ax\frac{\partial f}{\partial x_i} - 2\frac{\partial x^T}{\partial x_i}A^T b$$

Since $\frac{\partial x}{\partial x_i} = e_i$ (the i^{th} unit vector) we get

$$\frac{\partial f}{\partial x_i} = e_i^T A^T Ax + x^T A^T Axe_i - 2e_i^T A^T b = 2e_i^T A^T x - 2e_i^T A^T b$$

Using the fact that $e_i^T A^T = (A^T)_{i*}$, the i^{th} column of A^T setting $\frac{\partial f}{\partial x_i} = 0$ gives

$$(A^T)_{i*}Ax = (A^T)_{i*}b \quad \text{for } i = 1, 2, \ldots, n$$

This can be written as a single matrix equation

$$A^T Ax = A^T b$$

iii)If $Ax = b$ this system is consistent then the solution set $\mathcal{S} = p + \ker(\mathrm{A})$. Here p is a particular solution that satisfies $Ap = b$. Now, since by theorem 2.77 range($A^T A$) = range(A^T), the system $A^T Ax = A^T b$ is always consistent. Since, $A^T Ap = A^T b$ p is also a particular solution to the system $A^T Ax = A^T b$. Also by theorem 2.77 $\ker(A^T A) = \ker(A)$ the solution set $\mathcal{S}'$ of $A^T Ax = A^T b$ is

$$\mathcal{S}' = p + \ker(A^T A) = p + \ker(A) = \mathcal{S}$$

iv)If $\mathrm{rank}(A) = n$ then by theorem 2.77 $\mathrm{rank}(A) = \mathrm{rank}(A^T A) = n$, therefore the square matrix $A^T A$ is invertible as $\ker(A^T A) = 0$ and therefore the unique least square solution is given by $x = (A^T A)^{-1} A^T b$

$\square$

The least squares method is a very powerful method of estimation and prediction.

Example 2.79. An experiment is conducted to measure the loss in gms of a pint of ice-cream. The loss is measured by keeping the pint of ice-cream for a certain duration (in hrs) and at a certain temperature (in Farenheit). The data is given in the table

Time(weeks)	Temp (^oF)	Loss (gms)
1	-10	0.15
1	-5	0.18
1	0	0.20
2	-10	0.17
2	-5	0.19
2	0	0.22
3	-10	0.20
3	-5	0.23
3	0	0.25

What would be our prediction on the estimated loss of gms of ice-cream when the ice-cream is stored for 9 weeks at a temperature of -35^oF ?

Assuming a linear relationship between the loss in gms of ice-cream and the

$$y = \alpha_0 + \alpha_1 t_1 + \alpha_2 t_2$$

Here the variable y which represents the loss in gms of ice-cream is the dependent variable and the variables t_1 (the time in weeks) and t_2 the temperature in Fahrenheit are dependent variables. The parameters α_0 and α_1 will be determined by the least squares method.

Assuming a linear relation between the dependent variables and the independent variable the table gives a linear equation $Ax = b$ where

$$A = \begin{pmatrix} 1 & 1 & -10 \\ 1 & 1 & -5 \\ 1 & 1 & 0 \\ 1 & 2 & -10 \\ 1 & 2 & -5 \\ 1 & 2 & 0 \\ 1 & 3 & -10 \\ 1 & 3 & -5 \\ 1 & 3 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_1 \end{pmatrix} \quad b = \begin{pmatrix} 0.15 \\ 0.18 \\ 0.20 \\ 0.17 \\ 0.19 \\ 0.22 \\ 0.20 \\ 0.23 \\ 0.25 \end{pmatrix}$$

In practice, most of these systems like the present one can be shown to be inconsistent. The associated normal equations are $A^T Ax = A^T b$ which gives

$$A^T A = \begin{pmatrix} 9 & 18 & -45 \\ 18 & 42 & -90 \\ -45 & -90 & 375 \end{pmatrix} \quad x = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_1 \end{pmatrix} \quad A^T b = \begin{pmatrix} 1.79 \\ 3.73 \\ -8.2 \end{pmatrix}$$

In this case $\operatorname{rank}(A) = 3$ and hence the unique least squares solution is

$$x = (A^T A)^{-1} A^T b = \begin{pmatrix} 0.174 \\ 0.025 \\ 0.005 \end{pmatrix}$$

So $y = 0.174 + 0.025t_1 + 0.005t_2$. So, the loss in gms of ice-cream that is stored for 9 weeks at a temp of $-35^o F$ is

$$y = 0.174 + 0.025 \times 9 + 0.005 \times (-35) = 0.224 \,\text{gms}$$

2.8. Invariant Subspaces

2.9. Linear Functionals

Exercise 3

1. Let V be a set of real sequences $(a_1, a_2, ..., a_n, ...)$ such that $\sum_i a_i^2$ is finite. Prove that V is a vector space over $\mathbb{R}$.

2. Show that the space $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1 + 2x_2 + 2x_3 = 0\}$ forms a vector space.

3. Let W_1 and W_2 be two subspaces of of a vector space V .
(a) Prove that $W_1 \cap W_2$ is a subspace of V .
(b) Prove that $W_1 \cup W_2$ is a subspace of V if and only if $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

4. Let V be the vector space of all functions f from $\mathbb{R}$ into $\mathbb{R}$. Which of the following sets of functions are subspaces of V ?
(a) all f such that $f(x^2) = f(x)^2$
(b) all f such that $f(0) = f(1)$
(c) all f such that $f(3) = 1 + f(-5)$
(d) all f such that $f(-1) = 0$;

(e) all f which are continuous.

5. Let V be the vector space of all $n \times n$ matrices over $\mathbb{C}$. Which of the following matrices A in V are subspaces of V
(a) all invertible A
(b) all A such that $AB = BA$, where B is some fixed matrix
(c) all A such that $A^2 = A$

6. Let V be the vector space of all functions from $\mathbb{R}$ to $\mathbb{R}$. Show that the space of even functions and the set of odd functions are subspaces of V.

7. Find three vectors in $\mathbb{R}^3$ that are linearly dependent and such that any two of them are linearly independent.

8. Consider the complex vector space $V = \mathbb{C}^3$ and the list (v_1, v_2, v_3) of vectors in V , where $v_1 = (i, 0, 0)$, $v_2 = (i, 1, 0)$, $v_3 = (i, i, -1)$:
(a) Prove that $\text{span}(v_1, v_2, v_3) = V$.
(b) Prove or disprove: (v_1, v_2, v_3) is a basis for V .

9. Are the vectors $x_1 = (1, 1, 2, 4), x_2 = (2, -1, -5, 2), x_3 = (1, -1, -4, 0), x_4 = (2, 1, 1, 6)$ linearly independent in $\mathbb{R}^4$? If not find a basis for the subspace of $\mathbb{R}^4$ spanned by the four vectors.

10. Let V be the vector space of 2×2 matrices over the complex numbers. As seen this has dimension 4. Find a basis of V consisting of matrices A_1, A_2, A_3 and A_4 such that $A_j^2 = A_j$ for each j.

11. Determine the dimension of each of the following subspaces of $\mathbb{R}^4$.
(a) $\{(x_1, x_2, x_3, x_4) | x_4 = 0\}$.
(b) $\{(x_1, x_2, x_3, x_4) | x_4 = x_1 + x_2\}$.
(c) $\{(x_1, x_2, x_3, x_4) | x_4 = x_1 + x_2, x_3 = x_1 - x_2\}$.

12. Let V be the vector space of 2×2 matrices over the complex numbers. Let W_1 be the matrices of the form $\begin{pmatrix} x & -x \\ y & z \end{pmatrix}$ and let W_2 be the set of matrices of the form $\begin{pmatrix} a & b \\ -a & c \end{pmatrix}$.
i) Show that W_1 and W_2 are subspaces of V.
ii)Find the dimensions of $W_1, W_2, W_1 + W_2$ and $W_1 \cap W_2$.

13. Which of the following from $\mathbb{R}^2$ to $\mathbb{R}^2$ is a linear transformation
(a)$T(x_1, x_2) = (1 + x_1, x_2)$
(b)$T(x_1, x_2) = (x_2, x_1)$
(c)$T(x_1, x_2) = (x_1^2, x_2)$
(d)$T(x_1, x_2) = (\sin(x_1), x_2)$

(e)$T(x_1, x_2) = (x_1 - x_1, 0)$

14. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix}$ for all $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$
(a) Show that T is surjective.
(b) Find dim (null (T)).
(c) Find the matrix for T with respect to the canonical basis of $\mathbb{R}^2$.

15. Let T be a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ defined by
$T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1)$
(a) What is the matrix of T relative to the standard ordered basis in $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively.
(b) What is the matrix of T relative to the ordered basis in $B = \{u_1, u_2, u_3\}$ in $\mathbb{R}^3$ and $B' = \{v_1, v_2\}$ in $\mathbb{R}^3$ where
$u_1 = (1, 0, -1), u_2 = (1, 1, 1), u_3 = (1, 0, 0)$ and $v_1 = (0, 1), v_2 = (1, 0)$.

16. Let T be a linear operator on $\mathbb{R}^3$ the matrix representation in the standard ordered basis is $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{pmatrix}$. Find the basis for the range of T and the null space of T.

17. Let T be the linear operator on the space of 2×2 complex matrices such that $T(A) = A^t$. Find a matrix representation of T with respect the basis E_{ij} where E_{ij} are matrices that have the $(i, j)^{th}$ element 1 and 0 elsewhere.

18. Show that the vectors $v_1 = (1, 1, 0, 0), v_2 = (0, 0, 1, 1), v_3 = (1, 0, 0, 4), v_4 = (0, 0, 0, 2)$ form a basis of $\mathbb{R}^4$. Find the coordinates of the vector $v = (1, 3, -1, 2)$ in this basis.

19. Let T be a linear operator defined on $\mathbb{R}^3$ defined by
$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3)$
a) What is matrix of T in the stadard basis of $\mathbb{R}^3$
b)What is matrix of T relative to the basis $(\alpha_1, \alpha_2, \alpha_3)$ where $\alpha_1 = (1, 0, 1), \alpha_2 = (-1, 2, 1), \alpha_3 = (2, 1, 1)$
c) Find matrix for T^{-1} in both these bases.

20. Let T be a linear operator on $\mathbb{R}^3$, the matrix of T in the standard basis is given by

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

Find a basis for the range of T and the null space of T.

21. Check if the following system of equations has a solution. If yes find the solution(s)

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 2 \\ x_1 + 2x_2 + 3x_3 &= 1 \\ x_1 + x_3 &= -3 \\ 2x_1 + 4x_2 &= 8 \end{aligned}$$

22. Check if the following system of equations has a solution. If yes find the solution(s)

$$\begin{aligned} x_1 + 2x_2 + 2x_3 + 3x_4 &= 4 \\ 2x_1 + 4x_2 + x_3 + 3x_4 &= 5 \\ 3x_1 + 6x_2 + x_3 + 4x_4 &= 7 \end{aligned}$$

23. Show that if A is $m \times n$ and B is $n \times p$ then
i) $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$
ii)$\text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB)$
(Hint: Use a. $\text{rank}(AB) = \text{rank}(B) - \dim(\text{Ker}(A) \cap \text{Range}(B))$
b. $\text{rank}(A) = \text{rank}(A^T)$)

24. A small company has been in business for four years and has recorded annual sales (in tens of thousands of Rs) as follows
    Assuming a linear relationship between time and sales predict a best estimate of the sales of the company in year 5.

Year	Sales
1	23
2	27
3	30
4	34

25. Cancer researchers hypothesize that the number of malignant cells (y) in a particular tissue grows exponentially with time (t), that is $y = \alpha_0 e^{\alpha_1 t}$. Determine the least squares estimate of the parameters α_0 and α_1 from the observed data given below.

t(days)	y (cells)
1	16
2	27
3	45
4	74
5	122

26. A hypothesis is that change in the price of bread is a linear combination of wheat and change in price of the minimum wage, that is

$$B = \alpha W + \beta M$$

The following is change (In Rupees) in price of the bread, wheat and minimum wages for three consecutive years Estimate the change in price of bread in Year 4 if wheat prices and minimum wage each fall by Rs 1.

	Year 1	Year 2	Year 3
B	+1	+1	+1
W	+1	+2	+0
M	+1	+0	-1

CHAPTER 3

Inner product space

3.1. Inner Products and Norms

DEFINITION 3.1. Let V be a vector space over the field $\mathbb{C}$ then an inner product is a function $\langle\cdot,\cdot\rangle :\ V \times V \longrightarrow \mathbb{C}$ such that

i) (Linearity) $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle\ \forall\ u, v, w \in V$ and $\alpha, \beta \in \mathbb{F}$.
ii) (Positive definiteness) $\langle v, v\rangle \geq 0\ \forall\ v \in V$ and $\langle v, v\rangle = 0 \Leftrightarrow v = 0$.
iii) (Conjugate symmetry) $\langle u, v\rangle = \overline{\langle v, u\rangle}\ \forall\ u, v \in V$.

Note that the inner product is conjugate linear in the second argument, i.e., $\langle u, \alpha v + \beta w\rangle = \bar{\alpha}\langle u, v\rangle + \bar{\beta}\langle u, w\rangle$. Indeed,

$$\begin{aligned}
\langle u, \alpha v + \beta w\rangle &= \overline{\langle \alpha v + \beta w, u\rangle} \text{ (due to iii)}\\
&= \overline{\alpha\langle v, u\rangle + \beta\langle w, u\rangle} \text{ (due to i)}\\
&= \bar{\alpha}\overline{\langle v, u\rangle} + \bar{\beta}\overline{\langle w, u\rangle}\\
&= \bar{\alpha}\langle u, v\rangle + \bar{\beta}\langle u, w\rangle
\end{aligned}$$

If the field is the field of real numbers $\mathbb{R}$ then the real inner product is a function $\langle\cdot,\cdot\rangle\ :\ V \times V \longrightarrow \mathbb{R}$ with the properties $i)$,$ii)$ and the third property $iii)$ is symmetric, i.e., $\langle u, v\rangle = \langle v, u\rangle$.

EXAMPLE 3.2. Define an inner product on $\mathbb{C}^n$ as $\langle u, v\rangle = \sum_i u_i \bar{v}_i$ where u_i and $\bar{v}_i$ are the i^{th} components of u and v, respectively. On $\mathbb{R}^n$ one can define a real inner product as $\langle u, v\rangle = \sum_i u_i v_i$.

EXAMPLE 3.3. Let $\mathcal{C}(\mathbb{R})$ be the space of continuous complex valued functions on $\mathbb{R}$. One can define inner product as $\langle f, g\rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}dx$.

EXAMPLE 3.4. Let $M_n(\mathbb{C})$ and $M_n(\mathbb{R})$ be the space of $n \times n$ matrices with entries in the complex numbers and the real numbers, respectively. Define $\langle A, B\rangle = Trace(AB^\dagger)$ where $B^\dagger$ denotes the conjugate transpose of B. This forms an inner product on $M_n(\mathbb{C})$. Similarly, $\langle A, B\rangle = Trace(AB^t)$ is an inner product on $M_n(\mathbb{R})$.

EXERCISE 3.5. Check that all the above examples satisfy the three properties of an inner product.

DEFINITION 3.6. (Norm) Let V be an inner product on $\mathbb{R}$ or $\mathbb{C}$ then a function $\|\cdot\|\ :\ V \to \mathbb{R}$ is called a norm if $\forall\ u, v \in V$,

i) (Homogeneity) : $\|\alpha v\| = |\alpha|\|v\|$
ii) (Positive definiteness) : $\|v\| \geq 0$, $\|v\| = 0$ iff $v = 0$
iii) (Triangle inequality) : $\|u + v\| \leq \|u\| + \|v\|$

Given an inner product, $\|v\| = \sqrt{\langle v, v\rangle}$ defines a norm on V.

3.2. Orthogonality, Orthogonal Basis and Gram-Schmidt procedure

DEFINITION 3.7. (Orthogonality) Given an inner product space V, we say $u, v \in V$ is orthogonal (denoted by $u \perp v$) if $\langle u, v\rangle = 0$.

THEOREM 3.8. *(Pythagoras) If $u \perp v$ then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.*

PROOF.

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \langle v,u\rangle + \langle u,v\rangle + \langle v,v\rangle \\
&= \|u\|^2 + 0 + 0 + \|v\|^2 \\
&= \|u\|^2 + \|v\|^2
\end{aligned}$$

□

This in fact generalizes to $u_1, u_2, \ldots, u_n$ mutually orthogonal, i.e.,

$$\left\|\sum_i u_i\right\|^2 = \sum_i \|u_i\|^2 \; (\langle u_i, u_j\rangle = \delta_{ij})$$

DEFINITION 3.9. (Orthonormal basis) A basis $\{v_1, v_2, \ldots, v_n\}$ is called an orthonormal basis if $\langle v_i, v_j\rangle = \delta_{ij}$ and $\|v_i\| = 1 \; \forall \; i = 1, 2, \ldots, n$.

If u is a vector in an inner product space with ONB $\{v_1, v_2, \ldots, v_n\}$ then one can write

$$u = \sum_{i=1}^{n} \alpha_i v_i$$

Taking inner product with v_j for some $j \in 1, 2, \ldots n$

$$\langle u, v_j\rangle = \langle \sum_{i=1}^{n} \alpha_i v_i, v_j\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_j\rangle = \alpha_j$$

Hence a vector u in an inner product space with ONB $\{v_1, v_2, \ldots, v_n\}$ can be written as

$$u = \sum_{i=1}^{n} \langle u, v_i\rangle v_i$$

This is called the Fourier expansion of the vector u in the ONB $\{v_1, v_2, \ldots, v_n\}$

EXAMPLE 3.10.

EXAMPLE 3.11.

Given a basis $\{w_1, w_2, \ldots, w_n\}$ of a vector space V one can iteratively produce an ONB by using the following procedure.
Let

$$\begin{aligned}
v_1 &= \frac{w_1}{\|w_1\|} \\
v_k &= \frac{w_k - \sum_{i=1}^{k-1} \langle w_k, v_i\rangle v_i}{\|w_k - \sum_{i=1}^{k-1} \langle w_k, v_i\|} \quad \text{for } k = 2, 3, \ldots, n
\end{aligned}$$

The dividing by the norm in the denominator is to normalize the vector. One can check that $\langle v_i, v_j\rangle = 0$ for $i \neq j$.

EXERCISE 3.12. Show that $\langle v_i, v_j\rangle = 0$ for $i \neq j$.

DEFINITION 3.13. (Orthogonal complement) Let $U \subseteq V$ then $U^\perp = \{v \in V : \langle v, u\rangle = 0 \;\forall\; u \in U\}$ is called the orthogonal complement of U.

EXERCISE 3.14. Show that $U^\perp$ is a subspace of V.

THEOREM 3.15. *Let $U \subseteq V$ then $U \oplus U^\perp = V$.*

PROOF. Let $v \in V$ and $\{w_1, w_2, \ldots, w_k\}$ be a basis of U. Use the Gram-Schmidt procedure to form an ONB $\{v_1, v_2, \ldots, v_k\}$ for U. Then,

$$v = \underset{u_1}{\left(\sum_{i=1}^{k}\langle v, v_i\rangle v_i\right)} + \underset{u_2}{\left(v - \sum_{i=1}^{k}\langle v, v_i\rangle v_i\right)}$$

Now, $u_1 \in U$ since it is a linear combination of basis vectors of U. Moreover $u_2 \in U^\perp$ since one can check that $\langle u_2, v_j\rangle = 0 \;\forall\; j = 1, 2, \ldots, k$

$$\begin{aligned}
\left\langle v - \sum_{i=1}^{k}\langle v, v_i\rangle, v_j\right\rangle &= \langle v, v_j\rangle - \left\langle \sum_{i=1}^{k}\langle v, v_i\rangle v_i, v_j\right\rangle \\
&= \langle v, v_j\rangle - \sum_{i=1}^{k}\langle v, v_i\rangle\langle v_i, v_j\rangle \\
&= \langle v, v_j\rangle - \sum_{i=1}^{k}\langle v, v_i\rangle\delta_{ij} \\
&= \langle v, v_j\rangle - \langle v, v_j\rangle \\
&= 0
\end{aligned}$$

Hence any vector $v \in V$ can be written as $v = u_1 + u_2$ where $u_1 \in U$ and $u_2 \in U^\perp$. If $U \cap U^\perp \neq \{0\}$ then $\exists\; u \neq 0$ such that $u \in U$ and $u \in U^\perp$ then $\langle u, u\rangle = 0$, but this is a contradiction since $\langle u, u\rangle = 0 \Leftrightarrow u = 0$. Thus, $U \cap U^\perp = \{0\}$. Hence, $V = U \oplus U^\perp$.

□

3.3. Orthogonal Projections closest point theorem

Let $U \subset V$ then as we have shown any $v \in V$ can be written as $v = u_1 + u_2$ where $u_1 \in U$ and $u_2 \in U^\perp$. Consider an operator $P_U : V \longrightarrow V$ defined as $P_U(v) = u_1$.

P_U projects onto the subspace U. It assigns v to u_1. From teh previous discussion we know that any $v \in V$ has a unique decomposition $v = \sum_{i=1}^{k}\langle v, v_i\rangle v_i + \left(v - \sum_{i=1}^{k}\langle v, v_i\rangle v_i\right)$ where $\{v_i\}_{i=1}^{k}$ is any ONB of U. Now we observe that $Range(P_U) = U$ and $Null(P_U) = U^\perp$.

THEOREM 3.16. *Show that if P_U is an orthogonal projection then $Range(P_U) = U$ and $Null(P_U) = U^\perp$ hence $V = Range(P_U) \oplus Null(P_U)$.*

PROOF. □

THEOREM 3.17. *(CLosest point) Let $U \subseteq V$ and let $v \in V$ then*

$$\|v - P_U v\| \leq \|v - u\| \;\forall\; u \in U$$

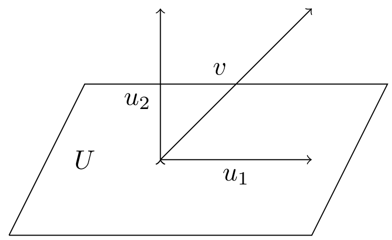

FIGURE 1

PROOF. We have

$$\begin{aligned} \|v - P_U v\|^2 &\leq \|v - P_U v\|^2 + \|P_U v - u\|^2 \\ &= \|v - P_U v + P_U v - u\|^2 \text{ by Pythagoras theorem} \\ &= \|v - u\|^2 \end{aligned}$$

□

Thus the projected vector onto the subspace U has the smallest distance from v among all $u \in U$.

THEOREM 3.18 (Cauchy-Schwarz). *If u, v are vectors in an inner product space V then*

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

PROOF. Consider the projection of the vector v onto the subspace U spanned by the unit vector $\frac{u}{\|u\|}$

$$P_U(v) = \frac{\langle v, u\rangle}{\|u\|^2}$$

Now, $P_U(v) \in U$ and $v - P_U(v) \in U^{\perp}$, so since

$$v = P_U(v) + v - P_U(v)$$

by Pythagoras theorem we have

$$\begin{aligned} \|v\|^2 &= \|P_U(v)\|^2 + \|v - P_U(v)\|^2 \\ \|v\|^2 &\geq \|P_U(v)\|^2 \\ \|v\|^2 &\geq \left\| \frac{\langle v, u\rangle}{\|u\|^2} \right\|^2 = \frac{|\langle v, u\rangle|^2}{\|u\|^2} \end{aligned}$$

□

Application of the Cauchy-Schwarz inequality for different vector spaces leads to the following inequalities

EXAMPLE 3.19. In $\mathbb{R}^n$ with $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$

$$\sum_i x_i y_i \leq \sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}$$

In $\mathcal{C}[0,1]$ where $\langle f,g\rangle = \int_0^1 f(t)g(t)dt$

$$\left|\int_0^1 f(t)g(t)dt\right| \leq \sqrt{\int_0^1 f(t)^2}\sqrt{\int_0^1 g(t)^2}$$

In $M_n(\mathbb{C})$ where $\langle A,B\rangle = \mathrm{Tr}(AB^\dagger)$

$$|\mathrm{Tr}(AB^\dagger)| \leq \sqrt{\mathrm{Tr}(AA^\dagger)}\sqrt{\mathrm{Tr}(BB^\dagger)}$$

3.4. Operators on Inner Product Spaces

Let A be the matrix representation of a linear operator in an ONB $\{v_1, v_2, \ldots, v_n\}$ then

$$Av_j = \sum_i a_{ij} v_i$$

Here a_{ij} are the matrix elements of A in this basis. Taking the inner product with vector v_k on both sides we get

$$\begin{aligned}\langle Av_j, v_k\rangle &= \langle \sum_i a_{ij} v_i, v_k\rangle \\ &= \sum_i a_{ij}\langle v_i, v_k\rangle \\ &= a_{kj}\end{aligned}$$

Definition 3.20 (Adjoint). Let $A: V \to V$ be a linear operator on an inner product space V then $\exists$ a unique linear operator $A^\dagger$ (called the adjoint of A) that satisfies $\langle Au, v\rangle = \langle u, A^\dagger v\rangle \; \forall u, v \in V$.

Exercise 3.21. Show that
i. $(A^\dagger)^\dagger = A$ b. $(A+B)^\dagger = A^\dagger + B^\dagger$ c. $(AB)^\dagger = B^\dagger A^\dagger$

Now,

$$\langle Av_j, v_k\rangle = \langle v_j, A^\dagger v_k\rangle = \overline{\langle A^\dagger v_k v_j\rangle} = \overline{a^*_{jk}}$$

Hence the matrix elements of $A^\dagger$ are the conjugate-transpose of the matrix elements of A.

The operators that will play an important role will be

Complex vector spaces

i) Hermitian operators (or sefl adjoint operators) : $A = A^\dagger$

ii) Unitary operators : $U^\dagger U = UU^\dagger = I$

Real vector spaces

iii) Symmetrix operators : $A^T = A$

iv) Orthogonal operators : $OO^T = O^T O = I$ (real rotation)

Let A be the matrix representation of a Hermitian operator in an ONB $\mathcal{B}$ then since $A = A^\dagger$ we get $a_{ij} = \overline{a_{ji}}$ so the non-diagonal elements of A are complex conjugates of each other and the diagonal elements are real numbers (since $a_{ii} = \overline{a_{ii}}$).

Theorem 3.22. *Let $U: V \to V$ be a Unitary operator on a complex inner product space then*

i) $\langle Ux, Uy\rangle = \langle x, y\rangle \; \forall\, x, y \in V$

ii) $\|Ux\| = \|x\| \; \forall\, x \in V$

iii) *The rows(columns) of U are mutually orthonormal*

PROOF. i) $\langle Ux, Uy\rangle = \langle x, U^\dagger Uy\rangle = \langle x, Iy\rangle = \langle x, y\rangle$

ii) $\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, U^\dagger Ux\rangle = \langle x, Ix\rangle = \langle x, x\rangle = \|x\|^2$

iii) The inner product of the j^{th} column with the k^{th} column of U is given by

$$\begin{aligned}\langle U_{*j}, U_{*k}\rangle &= \sum_i U_{ij}\bar{U}_{ik} \\ &= \sum_i U_{ij}U^\dagger_{ki} \\ &= \sum_i U^\dagger_{ki}U_{ij} = \delta_{kj} \quad \text{Since } U^\dagger U = \mathbb{1}\end{aligned}$$

For the rows

$$\begin{aligned}\langle U_{j*}, U_{k*}\rangle &= \sum_i U_{ji}\bar{U}_{ki} \\ &= \sum_i U_{ji}U^\dagger_{ik} = \delta_{jk} \quad \text{Since } UU^\dagger = \mathbb{1}\end{aligned}$$

□

THEOREM 3.23. *Let $P : V \to V$ be a Orthogonal operator on a real inner product space then*

i) $\langle Px, Py\rangle = \langle x, y\rangle \ \forall\, x, y \in V$

ii) $\|Px\| = \|x\| \ \forall\, x \in V$

iii) *The rows(columns) of U are mutually orthonormal*

PROOF. Is similar to the proof for the Unitary operators replacing the conjugate transpose (†) operator by the transponse. □

The orthogonal operators on real inner products are the rotations ans reflections. In $\mathbb{R}^2$ a rotation anticlock-wise by an angle θ is

$$\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$$

The rotations by angle θ

$$O_x = \begin{pmatrix}1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta\end{pmatrix} \quad O_y = \begin{pmatrix}\cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta\end{pmatrix} \quad O_z = \begin{pmatrix}\cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1\end{pmatrix}$$

Coming back to Orthogonal Projections, let P_u be a orthogonal projection onto a unit vector u, then $P_u = uu^\dagger$ since

$$P_u(v) = uu^\dagger(v) = u^\dagger(v)u = \langle v, u\rangle u$$

Extending this to a subspace U spanned by mutually orthonormal vectors $\{u_1, u_2, \ldots, u_k\}$

$$P_U(v) = \sum_{i=1}^{k}\langle v, u_i\rangle u_i$$

Hence $P_U = \sum_{i=1}^{k} u_i u_i^\dagger$ since

$$P_U(v) = \sum_{i=1}^{k}\langle v, u_i\rangle u_i = \sum_{i=1}^{k} u_i u_i^\dagger(v)$$

We have the following theorem

THEOREM 3.24. *Let $P : V \to V$ be a linear operator on a vector space V then P is a projection if and only if $P^2 = P$ and $P^\dagger = P$*

PROOF. $\Longrightarrow$ Let P be an orthogonal projection, therefore $V = \text{Range}(P) \oplus^\perp \text{Null}(P)$. If $v \in V$ then $v = u + w$ where $u \in \text{Range}(P)$ and $w \in \text{Null}(P)$ and $P(v) = u$. Since $u \in \text{Range}(P)$ $Pu = u$. Therefore,

$$P^2(v) = P(P(v)) = Pu = u = P(v)$$

Hence $P^2 = P$. Now if $x, x' \in V$ then $x = y + z$ and $x' = y' + z'$ for some $y, y' \in \text{Range}(P)$ and some $z, z' \in \text{Null}(P)$. So,

$$\langle Px, x'\rangle = \langle y, y' + z'\rangle = \langle y, y'\rangle = \langle y + z, y'\rangle = \langle x, Px'\rangle$$

' Hence $P = P^\dagger$ $\Longleftarrow$ Let $P^2 = P$ and $P^\dagger = P$ and let $x \in V$ then

$$x = Px + (x - Px)$$

$Px \in \text{Range}(P)$ and $P(x - Px) = Px - P^2x = Px - Px = 0$, so $x - Px \in \text{Null}(P)$. Now if $y \in \text{Range}(P)$ then there exist $x \in V$ such that $Px = y$. If $z \in \text{Null}(P)$ then

$$\langle y, z\rangle = \langle Px, z\rangle = \langle x, P^\dagger z\rangle = \langle x, Pz\rangle = \langle x, 0\rangle = 0$$

So $\text{Range}(P) \perp \text{Null}(P)$. Therefore $V = \text{Range}(P) \oplus^\perp \text{Null}(P)$ and P is a projection.

EXERCISE 3.25. Let P_U be the orthogonal projection onto the subspace U. Show that $\mathbb{1} - P_U$ is the orthogonal projection onto $U^\perp$

$\square$

Let u be a unit vector, then

$$R_u := 2uu^\dagger - \mathbb{1} = uu^\dagger - (\mathbb{1} - uu^\dagger)$$

As noted earlier $uu^\dagger$ is projection onto the unit vector u and $\mathbb{1} - uu^\dagger$ is projection onto $U^\perp$. One observes that $R_u(u) = u$ and if $w \in \text{span}\{u^\perp\}$ then $R_u(w) = -w$. In general one notes that if $x \in V$ then $x = u + w$ where $w \in \text{span}\{u^\perp\}$ then $R_u(x) = u - w$ (see figure).

EXERCISE 3.26. Show that $\mathbb{1} - 2uu^\dagger$ is reflection about $\text{span}\{u^\perp\}$

3.5. Discrete Fourier Transform

The solutions to the equation $z^n = 1$ over the complex field are called the n^{th} roots of unity. The roots of unity are given by $\{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$ where $\omega = e^{2\pi i/n} = \cos\frac{2\pi i}{n} + i\sin\frac{2\pi i}{n}$. Let $\zeta = \bar{\omega} = e^{-2\pi i/n}$.

The Discrete Fourier transform is a linear transformation $F_n : \mathbb{C}^n \to \mathbb{C}^n$,

$$F_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta & \zeta^2 & \cdots & \zeta^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \zeta^{n-1} & \zeta^{n-2} & \cdots & \zeta \end{pmatrix}$$

LEMMA 3.27. *If k is an integer then*

$$1 + \zeta^k + \zeta^{2k} + \cdots + \zeta^{(n-1)k} = 0 \quad \textit{if } \zeta^k \neq 1$$

PROOF.

$$\zeta^k(1+\zeta^k+\zeta^{2k}+\cdots+\zeta^{(n-1)k}) \quad = \quad \zeta^k+\zeta^{2k}+\cdots+\zeta^{(n-1)k}+1$$

So,

$$(1+\zeta^k+\zeta^{2k}+\cdots+\zeta^{(n-1)k})(1-\zeta^k)=0$$

Since $\zeta^k \neq 1$, so

$$1+\zeta^k+\zeta^{2k}+\cdots+\zeta^{(n-1)k}=0$$

□

THEOREM 3.28. *The columns of F_n are orthogonal*

PROOF.

$$\begin{aligned}
\langle F_{*r}, F_{*s}\rangle &= \sum_{j=0}^{n-1} F_{jr}F_{js}^{\dagger} \\
&= \sum_{j=0}^{n-1} \bar{F}_{sj}F_{jr} \\
&= \sum_{j=0}^{n-1} \zeta^{\bar{s}j}\zeta^{jr} \\
&= \sum_{j=0}^{n-1} \zeta^{j(r-s)} = n\delta_{rs} \quad \text{Due to lemma}
\end{aligned}$$

□

So, $\frac{1}{\sqrt{n}}F_n$ is Unitary. Also notice that F_n is symmetric therefore

$$\left(\frac{1}{\sqrt{n}}F_n\right)^{-1} = \left(\frac{1}{\sqrt{n}}F_n\right)^{\dagger} = \frac{1}{\sqrt{n}}\bar{F}_n$$

So,

$$F_n^{-1} = \frac{1}{n}\bar{F}_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \omega^{n-1} & \omega^{n-2} & \cdots & \omega \end{pmatrix}$$

The Fourier matrices of order 2 and 4 are

$$F_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad F_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

Let $\hat{a} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{2n\times 1}$ and $\hat{b} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{2n\times 1}$ then we define

$$\hat{a} \odot \hat{b} = \begin{pmatrix} \alpha_0\beta_0 \\ \alpha_0\beta_1 + \alpha_1\beta_0 \\ \vdots \\ \alpha_{n-1}\beta_0 + \alpha_{n-2}\beta_1 + \cdots \alpha_1\beta_{n-2} \\ \vdots \\ \alpha_{n-1}\beta_{n-1} \\ 0 \end{pmatrix} \quad \hat{a} \times \hat{b} = \begin{pmatrix} \alpha_0\beta_0 \\ \alpha_1\beta_1 \\ \vdots \\ \alpha_{n-1}\beta_{n-1} \\ \vdots \\ 0 \end{pmatrix}$$

THEOREM 3.29 (Convolution).

$$F_{2n}(\hat{a} \odot \hat{b}) = F_{2n}(\hat{a}) \times F_{2n}(\hat{b})$$

PROOF. $\square$

The convolution theorem along with the fast Fourier transform is used to for fast multiplication of integers.

Exercise 4

1. Let V be an inner product space, then show that $\|x\|= \sqrt{\langle x, x\rangle}$ is a defines a norm on V.

2. Parallelogram law
   Let V be a real or complex vector space with an inner product. Show that the norm defined by the inner product satisfies the parallelogram law

$$\|x+y\|^2+\|x-y\|^2= 2\|x\|^2+2\|y\|^2$$
$$t$$

3. In $M_2(\mathbb{R})$ we can define an inner product $\langle A, B\rangle = \mathrm{Tr}(AB^T)$. Verify that the set

$$\mathcal{B} = \left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \right\}$$

   is an orthonormal basis. Compute the Fourier expansion of $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ with respect to $\mathcal{B}$

4. Consider the space of real valued integrable functions on the interval $(-\pi, \pi)$ with inner product
$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(t)g(t)dt$$
Verify that the set of trigonometric
$$\mathcal{B} = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos t}{\sqrt{\pi}}, \frac{\cos 2t}{\sqrt{\pi}}, \dots, \frac{\sin t}{\sqrt{\pi}}, \frac{\sin 2t}{\sqrt{\pi}}, \dots\right\}$$
is an orthonormal basis of this space. Find the expansion of the square wave function
$$f(t) = \begin{cases} -1 \text{ when } -\pi < t < 0 \\ 1 \text{ when } 0 < t < \pi \end{cases}$$

5. Apply the Gram-Schmidt procedure for $\mathbb{C}^3$ to the vectors $\left\{\begin{pmatrix} i \\ i \\ i \end{pmatrix}, \begin{pmatrix} 0 \\ i \\ i \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ i \end{pmatrix}\right\}$

6. Consider the vector space P_3 of polynomials with rational coefficients that are of degree less than or equal to 3 on the domain $[-1, 1]$. Check that $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$ is an inner product on P_3. The monomials $\{1, x, x^2, x^3\}$ form a basis of P_3. Use the Gram-Schmidt orthogonalization process to produce an ONB of P_3. Express the polynomial $x^3 + 3x^2 + 2x + 3$ as a linear combination of the ONB vectors.

7. Let $A : V \to W$ be a linear transformation. Show that
a. $(A^\dagger)^\dagger = A$ b. $(A + B)^\dagger = A^\dagger + B^\dagger$ c. $(AB)^\dagger = B^\dagger A^\dagger$

8. Let $u = \begin{pmatrix} -2 \\ 1 \\ 3 \\ -1 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ 4 \\ 0 \\ -1 \end{pmatrix}$. Find
a. Orthogonal Projection of u onto span$\{v\}$
b.Orthogonal Projection of v onto span$\{u\}$
c.Orthogonal Projection of u onto $v^\perp$
d.Orthogonal Projection of v onto $u^\perp$

9. Determine the orthogonal projection of the vector $b = \begin{pmatrix} 5 \\ 2 \\ 5 \\ 3 \end{pmatrix}$ on to the Subspace M where $M = \text{span}\left\{\begin{pmatrix} -3/5 \\ 0 \\ 4/5 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 4/5 \\ 0 \\ 3/5 \\ 0 \end{pmatrix}\right\}$. What matrix representation of the operator P_M that projects onto the M in the standard basis. Find a basis and representation of P_M in this basis which is very convenient.

10. Let a solid unit cube cube be placed such that one on of the vertex is at the origin and the diagonally opposite vertex v is at the point $(1,1,1,)$. The cube is rotated first 90^o anticlockwise around the x-axis, followed by 45^o anticlockwise around the y-axis followed by 60^o anticlockwise around the z-axis. Find the location of the vertex v at the end of the three rotations

11. Let R be the reflection about the vector $u = \frac{1}{\sqrt{3}}(1,1,1)$ in $\mathbb{R}^3$. Find action of the reflection about u and on $u^\perp$ on the vector $v = (1,0,0)$

12. The Discrete Fourier transform is a linear transformation $F_n : \mathbb{C}^n \to \mathbb{C}^n$,
$$F_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta & \zeta^2 & \cdots & \zeta^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \zeta^{n-1} & \zeta^{n-2} & \cdots & \zeta \end{pmatrix}$$
i) Show that the columns of F_n are orthogonal
ii)$F_n^{-1} = \frac{1}{n}\bar{F}_n$

13. Let $\hat{a} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{2n\times 1}$ and $\hat{b} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{2n\times 1}$ then we define

$$\hat{a} \odot \hat{b} = \begin{pmatrix} \alpha_0\beta_0 \\ \alpha_0\beta_1 + \alpha_1\beta_0 \\ \vdots \\ \alpha_{n-1}\beta_0 + \alpha_{n-2}\beta_1 + \cdots \alpha_1\beta_{n-2} \\ \vdots \\ \alpha_{n-1}\beta_{n-1} \\ 0 \end{pmatrix} \quad \hat{a} \times \hat{b} = \begin{pmatrix} \alpha_0\beta_0 \\ \alpha_1\beta_1 \\ \vdots \\ \alpha_{n-1}\beta_{n-1} \\ \vdots \\ 0 \end{pmatrix}$$

Covolution theorem:

$$F_{2n}(\hat{a} \odot \hat{b}) = F_{2n}(\hat{a}) \times F_{2n}(\hat{b})$$

Use the convolution theorem to multiply $48_{10} \times 64_{10}$

14. Fast Fourier Transform (FFT)
Verify that

$$F_{2n} = \begin{pmatrix} F_n & D_nF_n \\ F_n & -D_nF_n \end{pmatrix} P_n$$

$$\text{where } D_n = \begin{pmatrix} 1 & & & & \\ & \zeta & & & \\ & & \zeta^2 & & \\ & & & \ddots & \\ & & & & \zeta^{n-1} \end{pmatrix} \text{ and } P_n^T = [e_0 e_2 \dots e_{2n-2} | e_1 e_3 \dots e_{2n-1}]$$

CHAPTER 4

Determinants

4.1. Determinants

Let $\sigma \in \mathcal{S}_n$ be a permutation on n elements for some $n \in \mathbb{Z}^+$, where $\mathcal{S}_n$ is the symmetric group (also called permutation group) on n elements. We define the sign of a permutation as

$$\text{sign}(\sigma) = \begin{cases} 1 & ; \text{ if } \sigma \text{ is even} \\ -1 & ; \text{ if } \sigma \text{ is odd} \end{cases}$$

For example, $\text{sign}((1,2,3)) = -1$ and $\text{sign}((1,2,3,4) = 1$. In other words, a permutation is said to be even if it can be written as a product of even number of transpositions and odd otherwise.

Definition 4.1. Let $A \in \mathbb{F}^{n\times n}$, for some $n \in \mathbb{Z}^+$, then the determinant of the square matrix A is defined as

$$det(A) = \sum_{\sigma \in \mathcal{S}_n} \text{sign}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$$

where $A_{i,j}$ is the (i,j)-th of A.

Example 4.2. Consider a 3×3 matrix $A \in \mathbb{F}^{3\times 3}$. Since

$$\mathcal{S}_3 = \{e, (1,2), (1,3), (2,3), (1,2,3), (1,3,2)\}$$

we can compute the determinant of A as

$$\begin{aligned} det(A) = {} & A_{1,1}A_{2,2}A_{3,3} \ + \ A_{1,2}A_{2,1}A_{3,3} \ + \ A_{1,3}A_{2,2}A_{3,1} \\ & + \ A_{1,1}A_{2,3}A_{3,2} \ - \ A_{1,2}A_{2,3}A_{3,1} \ - \ A_{1,3}A_{2,1}A_{3,2} \end{aligned}$$

Proposition 4.3. Consider $\mathbb{F}^{n\times n}$ for some $n \in \mathbb{Z}^+$. Then,

1. $det(O_n) = 0$, where $O_{i,j} = 0 \ \forall \ 1 \le i, j \le n$. Here O_n is $n \times n$ zero matrix.
2. $det(I_n) = 1$, where for $1 \le i, j \le n$,

$$I_{i,j} = \begin{cases} 0; & \text{if } i \neq j \\ 1; & \text{if } i = j \end{cases}$$

   Here I_n is $n \times n$ identity matrix.
3. If D is an $n \times n$ diagonal matrix then

$$det(D) = \prod_{i}^{n} D_{i,i}.$$

4. If T is an $n \times n$ triangular matrix, then

$$det(T) = \prod_{i}^{n} T_{i,i}.$$

T is called upper triangular matrix if $T_{i,j} = 0,\ \forall i > j$ and lower traingular if $T_{i,j} = 0,\ \forall i < j$.

5. $det(A^T) = det(A)$.

PROOF. abcd □

Let E be an $n \times n$ elementary matrix. There are three types of elementary matrices each corresponding to row operations. We say E is of

(1) Type-I if $E_{i,j} = I_{i,j}\ \forall\ 1 \leq i, j \leq n$ except for some $1 \leq r, s \leq n$ such that $E_{r,s} = 1, E_{s,r} = 1$ and $E_{r,r} = E_{s,s} = 0$.
(2) Type-II if $E_{i,j} = I_{i,j}\ \forall\ 1 \leq i, j \leq n$ except for some $1 \leq r \leq n$ such that $E_{r,r} = \lambda \neq 0$ $(\lambda \in \mathbb{F})$.
(3) Type-III if $E_{i,j} = I_{i,j}\ \forall\ 1 \leq i, j \leq n$ except for some $1 \leq r, s \leq n$ such that $E_{r,s} = \lambda$ $(\lambda \in \mathbb{F})$.

Type-I matrix interchanges the rows r and s, Type-II multiplies the r-th row and Type-III adds λ times row s to the row r of a matrix.

PROPOSITION 4.4. Let $A \in \mathbb{F}^{n\times n}$ for some $n \in \mathbb{Z}^+$ and E be an $n\times n$ elementary matrix. Then,

1. $det(EA) = -det(A)$, if E is of Type-I,
2. $det(EA) = \lambda\, det(A)$, if E is of Type-II,
3. $det(EA) = det(A)$, if E is of Type-III.

REMARK 4.5. Since elementary matrices are also square matrices so their determinant can be computed. We can easily check that

(1) $det(E) = -1$, if E is of Type-I,
(2) $det(E) = \lambda$, if E is of Type-II,
(3) $det(E) = 1$, if E is of Type-III.

Thus, the previous proposition can be restated as

COROLLARY 4.6. Let A be an $n \times n$ matrix and E be an elementary matric then

$$det(EA) = det(E)det(A).$$

REMARK 4.7. Let A be an $n \times n$ matrix and R be its row-reduced matrix. Then, $A = E_1E_2 \dots E_kR$ for some elementary operations $E_1, E_2, \dots, E_k$. We can compute the determinant of A as

$$det(A) = det(E_1)det(E_2) \dots det(E_k)det(R).$$

PROPOSITION 4.8. An $n \times n$ matrix is invertible (also called non-singular) if and only if $det(A) \neq 0$. Equivalently, a matrix is singular iff its determinant is zero.

PROPOSITION 4.9. Let $A, B \in \mathbb{F}^{n\times n}$ then

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

DEFINITION 4.10. Let $A \in \mathbb{F}^{n\times n}$. The determinant of a $k \times k$ submatrix of A is called a minor determinant of order k of A.

DEFINITION 4.11. (**Cofactors**) For an $n \times n$ matrix A we define the cofactor corresponding to (i, j)-th position of A as

$$A(i, j) = (-1)^{i+j} A(\bar{i}, \bar{j})$$

where $A(\bar{i}, \bar{j})$ is the minor of A obtained by deleting i-th row and j-th column.

The matrix of cofactors is denoted by $\mathring{A}$. So, $\mathring{A}_{i,j} = A(i,j)$.

PROPOSITION 4.12. **Cofactor expansions:** Given an $n \times n$ matrix A, its determinant can be computed as

1. about the i-th row
$$det(A) = \sum_{j}^{n} A_{i,j} \mathring{A}_{i,j}.$$
2. about the j-th column
$$det(A) = \sum_{i}^{n} A_{i,j} \mathring{A}_{i,j}.$$

CHAPTER 5

Eigenvalues and Eigenvectors

5.1. Determinants

Let $\sigma \in \mathcal{S}_n$ be a permutation on n elements for some $n \in \mathbb{Z}^+$, where $\mathcal{S}_n$ is the symmetric group (also called permutation group) on n elements. We define the sign of a permutation as

$$\text{sign}(\sigma) = \begin{cases} 1 & ;\text{ if } \sigma \text{ is even} \\ -1 & ;\text{ if } \sigma \text{ is odd} \end{cases}$$

For example, $\text{sign}((1,2,3)) = -1$ and $\text{sign}((1,2,3,4) = 1$. In other words, a permutation is said to be even if it can be written as a product of even number of transpositions and odd otherwise.

Definition 5.1. Let $A \in \mathbb{F}^{n\times n}$, for some $n \in \mathbb{Z}^+$, then the determinant of the square matrix A is defined as

$$det(A) = \sum_{\sigma \in \mathcal{S}_n} \text{sign}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$$

where $A_{i,j}$ is the (i,j)-th of A.

Example 5.2. Consider a 3×3 matrix $A \in \mathbb{F}^{3\times 3}$. Since

$$\mathcal{S}_3 = \{e, (1,2), (1,3), (2,3), (1,2,3), (1,3,2)\}$$

we can compute the determinant of A as

$$\begin{aligned} det(A) = {} & A_{1,1}A_{2,2}A_{3,3} \ + \ A_{1,2}A_{2,1}A_{3,3} \ + \ A_{1,3}A_{2,2}A_{3,1} \\ & + \ A_{1,1}A_{2,3}A_{3,2} \ - \ A_{1,2}A_{2,3}A_{3,1} \ - \ A_{1,3}A_{2,1}A_{3,2} \end{aligned}$$

Proposition 5.3. Consider $\mathbb{F}^{n\times n}$ for some $n \in \mathbb{Z}^+$. Then,

1. $det(O_n) = 0$, where $O_{i,j} = 0 \ \forall \ 1 \le i,j \le n$. Here O_n is $n \times n$ zero matrix.
2. $det(I_n) = 1$, where for $1 \le i,j \le n$,

$$I_{i,j} = \begin{cases} 0; & \text{if } i \ne j \\ 1; & \text{if } i = j \end{cases}$$

   Here I_n is $n \times n$ identity matrix.
3. If D is an $n \times n$ diagonal matrix then

$$det(D) = \prod_{i}^{n} D_{i,i}.$$

4. If T is an $n \times n$ triangular matrix, then

$$det(T) = \prod_{i}^{n} T_{i,i}.$$

T is called upper triangular matrix if $T_{i,j} = 0,\ \forall i > j$ and lower traingular if $T_{i,j} = 0,\ \forall i < j$.

5. $det(A^T) = det(A)$.

PROOF. abcd □

Let E be an $n \times n$ elementary matrix. There are three types of elementary matrices each corresponding to row operations. We say E is of

(1) Type-I if $E_{i,j} = I_{i,j}\ \forall\ 1 \leq i, j \leq n$ except for some $1 \leq r, s \leq n$ such that $E_{r,s} = 1, E_{s,r} = 1$ and $E_{r,r} = E_{s,s} = 0$.
(2) Type-II if $E_{i,j} = I_{i,j}\ \forall\ 1 \leq i, j \leq n$ except for some $1 \leq r \leq n$ such that $E_{r,r} = \lambda \neq 0\ (\lambda \in \mathbb{F})$.
(3) Type-III if $E_{i,j} = I_{i,j}\ \forall\ 1 \leq i, j \leq n$ except for some $1 \leq r, s \leq n$ such that $E_{r,s} = \lambda\ (\lambda \in \mathbb{F})$.

Type-I matrix interchanges the rows r and s, Type-II multiplies the r-th row and Type-III adds λ times row s to the row r of a matrix.

PROPOSITION 5.4. Let $A \in \mathbb{F}^{n\times n}$ for some $n \in \mathbb{Z}^+$ and E be an $n\times n$ elementary matrix. Then,

1. $det(EA) = -det(A)$, if E is of Type-I,
2. $det(EA) = \lambda\ det(A)$, if E is of Type-II,
3. $det(EA) = det(A)$, if E is of Type-III.

REMARK 5.5. Since elementary matrices are also square matrices so their determinant can be computed. We can easily check that

(1) $det(E) = -1$, if E is of Type-I,
(2) $det(E) = \lambda$, if E is of Type-II,
(3) $det(E) = 1$, if E is of Type-III.

Thus, the previous proposition can be restated as

COROLLARY 5.6. Let A be an $n \times n$ matrix and E be an elementary matric then

$$det(EA) = det(E)det(A).$$

REMARK 5.7. Let A be an $n \times n$ matrix and R be its row-reduced matrix. Then, $A = E_1E_2 \ldots E_kR$ for some elementary operations $E_1, E_2, \ldots, E_k$. We can compute the determinant of A as

$$det(A) = det(E_1)det(E_2)\ldots det(E_k)det(R).$$

PROPOSITION 5.8. An $n \times n$ matrix is invertible (also called non-singular) if and only if $det(A) \neq 0$. Equivalently, a matrix is singular iff its determinant is zero.

PROPOSITION 5.9. Let $A, B \in \mathbb{F}^{n\times n}$ then

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

DEFINITION 5.10. Let $A \in \mathbb{F}^{n\times n}$. The determinant of a $k \times k$ submatrix of A is called a minor determinant of order k of A.

DEFINITION 5.11. (**Cofactors**) For an $n \times n$ matrix A we define the cofactor corresponding to (i, j)-th position of A as

$$A(i, j) = (-1)^{i+j} A(\bar{i}, \bar{j})$$

where $A(\bar{i}, \bar{j})$ is the minor of A obtained by deleting i-th row and j-th column.

The matrix of cofactors is denoted by $\mathring{A}$. So, $\mathring{A}_{i,j} = A(i,j)$.

PROPOSITION 5.12. **Cofactor expansions:** Given an $n \times n$ matrix A, its determinant can be computed as

1. about the i-th row
$$det(A) = \sum_{j}^{n} A_{i,j} \mathring{A}_{i,j}.$$
2. about the j-th column
$$det(A) = \sum_{i}^{n} A_{i,j} \mathring{A}_{i,j}.$$

5.2. Diagonalization

Let V be a vector space over the field $\mathbb{F}$.

DEFINITION 5.13. Let $A : V \longrightarrow V$ be a linear transformation then $\lambda \in \mathbb{F}$ is called an eigenvalue of A if there exists a non-zero vector $v \in V$ such that $Av = \lambda v$. In this case the vector v is called an eigenvector pertaining to the eigenvalue λ.

We wish to obtain conditions under which there exists a basis where in the matrix representation of a linear operator is diagonal. Let A be the matrix representation of the linear operator in a basis $\mathcal{B}$, then diagonalizing A means to find a basis $\mathcal{B}'$ in which the operator A is diagonal. If P is the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$ then $[A]_{\mathcal{B}'} = P^{-1}[A]_{\mathcal{B}}P$.

THEOREM 5.14. *Let $A : V \longrightarrow V$ be a linear operator and dim $V = n$, then A is diagonalizable iff A has n linearly independent eigenvectors.*

PROOF. ($\Leftarrow$) Let $\{v_1, v_2, \dots, v_n\}$ be a set of linearly independent set of eigenvectors of A then $Av_j = \lambda_j v_j$, where λ_j is the eigenvalue corresponding to eigenvector v_j. Then, in the basis $\{v_1, v_2, \dots, v_n\}$ the matrix representation of A is **EMPTY**. Hence A can be diagonalized.
($\Rightarrow$) If A can be diagonalized then $\exists$ a basis $\mathcal{B}' = \{v_1, v_2, \dots, v_n\}$ such that $[A]_{\mathcal{B}'}$ =**EMPTY** for some $\lambda_i \in \mathbb{F}$. Due to this we see that $Av_j = \lambda_j v_j$. Hence v_j's are the eigenvectors of A and since they form a basis of V they are linearly independent. $\square$

Now, if v is an eigenvector of A with eigenvalue λ then $Av = \lambda v \Rightarrow (A - \lambda I)v = 0 \Rightarrow v \in null(A - \lambda I)$. Hence if $v \neq 0$ is a eigenvector of A with eigenvalue λ then $v \in null(A - \lambda I)$. But the existence of a non-zero vector in the null space of a matrix is equivalent to the matrix being singular (non-invertible) which is equivalent to having a zero determinant. Hence we have the following theorem.

THEOREM 5.15. *The followings are equivalent*

i) *v is an eigenvector of A with eigenvalue λ*
ii) *$v \in null(A - \lambda I)$*
iii) *$A - \lambda I$ is singular (non-invertible)*
iv) *$det(A - \lambda I) = 0$*

We denote $V_\lambda = null(A - \lambda I)$ and we call this subspace the eigenspace associated with the eigenvalue λ. Any vector in this space satisfies $Av = \lambda v$. Moreover if $\lambda \neq \mu$ then $V_\lambda \cap V_\mu = \emptyset$ since $Av = \lambda v$ and $Av = \mu v$ implies that $\lambda v = \mu v \Rightarrow (\lambda - \mu)v = 0 \Rightarrow v = 0$ since $\lambda \neq \mu$. This also implies the following theorem.

THEOREM 5.16. *$A : V \longrightarrow V$ is diagonalizable iff $V = V_{\lambda_1} \bigoplus V_{\lambda_2} \bigoplus \ldots \bigoplus V_{\lambda_r}$ where $\lambda_1, \lambda_2, \ldots, \lambda_r$ are the distinct eigenvalues of A.*

Lets look at a few examples.

EXAMPLE 5.17.

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

Can A be diagonalized? If so find a basis in which A is diagonal.
To find the eigenvalues of A we solve the equation $det(\lambda I - A) = 0$.

$$\begin{vmatrix} \lambda - 5 & 6 & 6 \\ 1 & \lambda - 4 & -2 \\ -3 & 6 & \lambda + 4 \end{vmatrix} = 0 \Rightarrow (\lambda - 2)^2(\lambda - -1) = 0.$$

So, the eigenvalues are 2 with multiplicity 2 and 1 with multiplicity 1.

To find the eigenvectors we look a th the null spacees of $A - \lambda I$ for $\lambda = 1$ and $\lambda = 2$.

$$A - I = \begin{vmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{vmatrix}$$

We proceed to reduce this matrix to row-reduced echelon form.

$$A - I \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

which gives $x_1 - x_3 = 0$ and $x_2 + \frac{x_3}{3} = 0$. Thus the null space of $A - I$ is one dimensional and consists of all vectors that are scalar multiples of $\begin{pmatrix} 1 \\ \frac{-1}{3} \\ 1 \end{pmatrix}$. Next,

$$A - 2I = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which gives $x_1 - 2x_2 - 2x_3 = 0$. And the nullspace is spanned by vectors of the form $\begin{pmatrix} 2x_2 + 2x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$. Thus the nullspace of $A - 2I$ is two dimensional and spanned by the two linearly independent vectors $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ and

$\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$. The eigenvectors of A are $\begin{pmatrix} 1 \\ \frac{-1}{3} \\ 1 \end{pmatrix}$ corresponding to eigenvalue 1 and $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ corresponding to eigenvalue 2. Since there are three linearly independent eigenvalues of A therfore A can be diagonalized. The change of basis matrix from the basis in which A was originally represented to the basis of eigenvectors

$$P = \begin{pmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \text{ and } P^{-1}AP = D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

EXAMPLE 5.18.

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

In this case $det(\lambda I - A) = \begin{vmatrix} \lambda - 3 & -1 & 1 \\ -2 & \lambda - 2 & 1 \\ -2 & -2 & \lambda \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2$ gives $\lambda = 1$ and $\lambda = 2$. Now,

$$A - I = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & \frac{1}{2} & \frac{-1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which gives $x_2 = 0$ and $x_1 - \frac{x_3}{2} = 0$. So null space of $A - I$ is spanned by $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$.

Next,

$$A - 2I = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & \frac{-1}{2} \\ 0 & 1 & \frac{-1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

which gives $x_1 - \frac{x_3}{2} = 0$ and $x_2 - \frac{x_3}{2} = 0$. So, $null(A - 2I)$ is also one dimensional. In this case there are only two linearly independent eigenvectors. Hence, A is not diagonalizable.

EXAMPLE 5.19.

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

Here $det(\lambda I - A) = \lambda^2 + 1$. Now, if the field was the set of real number then one cannot factorize $\lambda^2 + 1$ hence there are no eigenvalues and eigenvectors. However, if the field was the set of complex numbers then $\lambda = i$ and $\lambda = -i$ are the two eigenvalues. One can check that $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ i \end{pmatrix}$ are the corresponding eigenvectors that span $\mathbb{C}^2$. Hence over the field of complex number this matrix is diagonalizable.

DEFINITION 5.20. Let $T : V \to V$ be a linear operator over a field $\mathbb{F}$ and $dim(V) = n$ then the characteristic polynomial of T splits if it can be written as linear factors over $\mathbb{F}$, that is,

$$\det(T - \lambda \mathbb{1}) = \alpha(\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_m)^{k_m}$$
$$\text{where} \quad \sum_i k_i = n \quad \text{and } \alpha \in \mathbb{F}$$

Definition 5.21 (Geometric Multiplicity). Let $T : V \to V$ be a linear operator with characteristic polynomial $\det(T-\lambda\mathbb{1}) = \alpha(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k-2}\cdots(\lambda-\lambda_m)^{k_m}$, then the geometric multiplicity of an eigenvalue λ_i of T is the dimension of $null(T-\lambda_i\mathbb{1})$.

Definition 5.22 (Algebraic Multiplicity). Let $T : V \to V$ be a linear operator with characteristic polynomial $\det(T-\lambda\mathbb{1}) = \alpha(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k-2}\cdots(\lambda-\lambda_m)^{k_m}$, then the algebraic multiplicity of eigenvalue λ_i is k_i.

Lemma 5.23. *Let $T : V \to V$ be a linear operator with eigenvalue λ_0 then Geometric Multiplicity$(\lambda_0) \le$ Algebraic Multiplicity(λ_0)*

Proof. Let $v_1, v_2, \ldots, v_p$ be a basis for $V_{\lambda_0} := null(T - \lambda_0\mathbb{1})$, then we can extend this to a basis of $\mathcal{B} = \{v_1, v_2, \ldots, v_p, v_{p+1}, \ldots, v_n\}$ of V. In the basis $\mathcal{B}$ we can write

$$[T]_{\mathcal{B}} = \left(\begin{array}{cccc|c} \lambda_0 & 0 & \cdots & 0 & \\ 0 & \lambda_0 & \cdots & 0 & \text{B} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & \lambda_0 & \\ \hline & & 0 & & \text{C} \end{array}\right)$$

Then,

$$\begin{aligned} \det(T-\lambda\mathbb{1}) &= \det \left(\begin{array}{cccc|c} \lambda_0-\lambda & 0 & \cdots & 0 & \\ 0 & \lambda_0-\lambda & \cdots & 0 & \text{B -}\lambda\mathbb{1} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & \lambda_0-\lambda & \\ \hline & & 0 & & \text{C -}\lambda\mathbb{1} \end{array}\right) \\ &= (\lambda-\lambda_0)^p \det(C-\lambda\mathbb{1}) \end{aligned}$$

Therefore the algebraic multiplicity of λ_0 is at least p. Hence the geometric multiplicity of λ_0 is less than or equal to the algebraic multiplicity. □

Next we prove a nice criterion for the an operator to be diagonalized whose characteristic polynomial splits.

Theorem 5.24. *Let $T : V \to V$ be a linear operator on a vector space V over a field $\mathbb{F}$ such that the characteristic polynomial of T splits into distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ then T can be diagonalized $\iff$ Algebraic Multiplicity (λ_i) = Geometric Multiplicity (λ_i) for $i = i \ldots k$*

Proof. Let the characteristic polynomial of T be

$$P_T(\lambda) = (\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_m)^{k_m}$$

Since the characteristic polynomial splits $\sum_i k_i = dim(V) = n$. Let $V_{\lambda_i} = null(T-\lambda_i\mathbb{1})$ and $d_i = dim(V_{\lambda_i})$ is the geometric multiplicity of eigenvalue λ_i. Let $\mathcal{B}_i$ be the basis V_{λ_i}.
$\Leftarrow$ If Algebraic Multiplicity of λ_i = Geometric Multiplicity of λ_i for all i then since $\mathcal{B}_l \cap \mathcal{B}_j = \emptyset$ therefore $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cdots \cup \mathcal{B}_m$ is a basis of V consisting of eigenvectors. Therefore T can be diagonalized.

$\Rightarrow$ Suppose T can be diagonalized, let $\mathcal{B}$ be the basis of linearly independent eigenvectors of T. Let $\mathcal{B}_i = \mathcal{B} \cap V_{\lambda_i}$ and $dim(\mathcal{B}_i) = n_i$. It is clear that $n_i \leq d_i$ Now since $V_{\lambda_i} \cap V_{\lambda_i} = \emptyset$ for $i \neq j$, therefore $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$ for $i \neq j$ and hence $\sum_i n_i = n$. Also $\sum_i k_i = n$ and $d_i \leq k_i$ which gives

$$n = \sum_i n_i \leq \sum_i d_i \leq \sum_i k_i = n$$

Therefore we get

$$\begin{aligned} \sum_i k_i &= \sum_i d_i \\ \sum_i (k_i - d_i) &= 0 \end{aligned}$$

Since $k_i \geq d_i$ we get $k_i = d_i$ for all i □

As we have seen that not all operators can be diagonalized. Then, what is the best that we can do in terms of writing the operator in an acceptable form over some chosen basis. Well, if the operator is acting over a vector space over $\mathbb{C}$ then we have a theorem due to Schur

THEOREM 5.25 (Schur). *Let $T : V \to V$ be a linear operator over a complex vector space ($dim(V) = n$) then there exists an ONB in which T is upper triangular, that is*

$$[T]_{\mathcal{B}} = \begin{pmatrix} \times & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \times \end{pmatrix}$$

PROOF. We will proceed by induction. Case $n = 1$ is obvious. Assume that we can write a $n-1 \times n-1$ matrix in triangular form. Now since we are working over the complex field T has at least one eigenvalue corresponding to a eigen vector. By lemma **??** we have that $T^\dagger$ also must have an an eigenvector. Let w be an eigenvector of $T^\dagger$ with eigenvalue λ, that is $T^\dagger w = \lambda w$ and let $W = span w$.
Claim: $W^\perp$ is a T-invariant subspace.
Indeed, if $v \in W^\perp$ then

$$\langle Tv, w\rangle = \langle v, T^\dagger w\rangle = \langle v, \lambda w\rangle = \bar{\lambda}\langle v, w\rangle = 0$$

Now $dim(W^\perp) = n-1$ and hence by our induction hypothesis $T \restriction_{W^\perp}$ can be written in upper triangular form in some ONB basis $\{v_1, v_2, \ldots, v_{n-1}$. We can extend this to a basis $\mathcal{B} = \{v_1, v_2, \ldots, v_{n-1}, w\}$ of V which is an ONB of V, and the matrix representation of T in this basis is upper triangular □

LEMMA 5.26. *If $T : V \to V$ is a linear operator and v is an eigenvector with eigenvalue λ then $T^\dagger$ has an eigenvector with eigenvalue $\bar{\lambda}$*

PROOF. Let $x \in V$ then

$$\begin{aligned} 0 = \langle 0, x\rangle &= \langle (T - \lambda \mathbb{1})v, x\rangle \\ &= \langle v, (T - \lambda \mathbb{1})^\dagger x\rangle \\ &= \langle v, (T^\dagger - \bar{\lambda}\mathbb{1})x\rangle \end{aligned}$$

So, $v \in Range(T^\dagger - \bar{\lambda}\mathbb{1})^\perp$ and hence by rank-nullity theorem there exists a non-zero vector in the $Ker(T^\dagger - \bar{\lambda}\mathbb{1})$ □

5.3. Systems of Differential Equations

5.4. Normal matrices and Spectral Theorem

Definition 5.27 (Normal Operator). $T : V \to V$ is called a Normal operator if $TT^\dagger = T^\dagger T$

Examples of Normal operators:

(1) Unitary operators. $(U^\dagger = U^{-1})$
$UU^\dagger = UU^{-1} = \mathbb{1} = U^{-1}U = U^\dagger U$
(2) Hermitian operators. $(A^\dagger = A)$
$A^\dagger A = AA^\dagger = A^2$
(3) Skew Hermitian operators. $(A^\dagger = -A)$. $A^\dagger A = AA^\dagger = A^2$.
(4) Orthogonal operators on real vector spaces. $(P^T = P^{-1})$.
$PP^\dagger = PP^T = \mathbb{1} = P^{-1}P = P^TP = P^\dagger P$
(5) Symmetric operators on real vector spaces. $(A^T = A)$.
$AA^T = A^TA = A^2$

We have the following theorem

Theorem 5.28. *Let $T : V \to V$ be a Normal operator on a complex inner product space then*
(1) $\|Tv\| = \|T^\dagger v\|$
(2) $T - c\mathbb{1}$ *is normal for all* $c \in \mathbb{C}$
(3) *If* $Tv = \lambda v$ *then* $T^\dagger v = \bar{\lambda} v$
(4) *If* $\lambda_1 \neq \lambda_2$ *are eigenvalues of* T *with corresponding eigenvectors* v_1 *and* v_2 *then* $v_1 \perp v_2$

Theorem 5.29. *(Spectral theorem) If A is a normal matrix then $\exists$ a unitary (orthogonal in case A is real) matrix that diagonalizes A, i.e. $\exists$ a unitary matrix U such that $U^\dagger AU = D$.*

Proof. By Schur's theorem there exists an ONB $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ such that $[T]_\mathcal{B}$ is upper triangular. Since T is upper triangular we have that $Tv_1 = \lambda v_1$. Assume by the way of induction that $\{v_1, v_2, \ldots v_{k-1}\}$ are eigenvectors of T, that is $Tv_j = \lambda_j v_j$ for $j = 1, \ldots k-1$. Now we have

$$Tv_k = T_{1k}v_1 + T_{2k}v_2 + \cdots + Tkkv_k$$

where we recall that the matrix elements in the basis B

$$\begin{aligned} T_{jk} = \langle Tv_k, v_j\rangle &= \langle v_k, T^\dagger v_j\rangle \\ &= \langle v_k, \bar{\lambda}_j v_j\rangle \text{ From lemma} \\ &= \lambda_j \langle v_k, v_j\rangle = 0 \end{aligned}$$

Therefore,

$$Tv_k = T_{kk}v_k$$

and hence v_k is also an eigenvector of T. Our induction hypothesis is true and hence T is diagonal in the basis $\mathcal{B}$ □

5.5. Singular Value Decomposition

As we have learned that even over the complex field not all operators can be diagonalized. The theorem of Schur shows that every operator over a complex field can be brought into upper triangular form. If we have a linear transformation from an n dimensional space to a m dimensional space then its matrix representation is rectangular $m \times n$ matrix. Here our goal is to show the most general form of decomposition of of a rectangular matrix with complex entries called the singular value decomposition. The SVD has found its use in image and signal processing, data analysis and quantum mechanics. We begin with the following lemma.

LEMMA 5.30. *Let $A : V \to V$ be self ad joint, then all the eigenvalues A are real*

PROOF. Let v be an eigenvector of A corresponding to eigenvalue λ.

$$\begin{aligned} \langle Av, v\rangle &= \langle \lambda v, v\rangle = \lambda\langle v, v\rangle \\ &= \langle v, A^\dagger v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle \\ &= \bar{\lambda}\langle v, v\rangle \end{aligned}$$

Therefore $\lambda = \bar{\lambda}$ and the eigenvalues of A are real. □

THEOREM 5.31. *Let A be a symmetric operator over a real vector space then there exists an ONB in which A is diagonal.*

PROOF. Consider A as a linear operator over the complex vector field. The characteristic polynomial of A splits over $\mathbb{C}$. According to lemma (5.30) sll eigenvalues are real and therefore the characteristic polynomial splits over $\mathbb{R}$. Also, by Schur's theorem we get that there exists a basis in which

$$[A]_{\mathcal{B}} = \begin{pmatrix} \times & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \times \end{pmatrix}$$

$A = A^T$ implies that all the non-diagonal entries are zero. Hence A is diagonal in basis $\mathcal{B}$. □

DEFINITION 5.32. An operator $A : V \to V$ is called positive semi definite if

- A is self adjoint
- $\langle Av, v\rangle \geq 0$ for all $v \in V$

With this background in place we are ready for the statement of the singular value decomposition.

THEOREM 5.33. *Let A be a $m \times n$ matrix with complex entries, then there exist unitary matrices $U_{m\times m}$ and $V_{n\times n}$ and a matrix $\Sigma_{m\times n}$ such that*

$$\Sigma_{ij} = \begin{cases} \sigma_i \textit{ if } i = j \\ 0 \textit{ else} \end{cases}$$

and $A = U\Sigma V^\dagger$

PROOF. Need to show that $AV = U\Sigma$. A is an linear transformation from an n dimensional complex vector space V to an m dimensional complex vector space W. From the lemma we have ONB $\{v_1, v_2, \dots, v_n\}$ of V and a ONB $\{u_1, u_2, \dots, u_m\}$ of W. Consider the $n \times n$ matrix

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}_{n \times n}$$

where $\mathbf{v}_i$ is a $n \times 1$ vector representation of v_i in the standard basis. Similarly let

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{bmatrix}_{m \times m}$$

and

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & 0 \\ 0 & 0 & \cdots & \sigma_r & 0 \end{pmatrix}_{m \times n}$$

Now Av_j is the j^{th} column of the product AV. But by the lemma

$$Av_j = \sigma_j u_j$$

Next, look at the jth column of $U\Sigma$

$$U\Sigma = U \begin{bmatrix} \sigma_1 \mathbf{e}_1 & \sigma_2 \mathbf{e}_2 & \cdots & \sigma_r \mathbf{e}_r & \cdots & 0\mathbf{e}_n \end{bmatrix}$$

The j^{th} column of $U\Sigma$ is

$$U[\sigma_j \mathbf{e}_j] = \sigma_j U[\mathbf{e}_j] = \sigma_j \mathbf{e}_j$$

$\square$

LEMMA 5.34. *Let $A : V \to W$ be a linear transformation with rank$(A) = r$, then there exists ONB $\{v_1, v_2, \dots v_n\}$ of V and ONB $\{u_1, u_2, \cdots, u_m\}$ of W and positive numbers $\sigma_1 \geq \sigma_2 \cdots \sigma_r$ such that*

$$Av_i = \begin{cases} \sigma_i u_i \ for 0 \leq i \leq r \\ 0 \ for i > r \end{cases}$$

PROOF. Consider $A^{dagger} A$. It is easy to see that $A^\dagger A$ is positive semi definite. Since $A^\dagger A$ is self adjoint, it can be diagonalized. Let $\lambda_1, \lambda_2, \dots \lambda_r, 0, \dots 0$ be the eigen values corresponding to the orthonormal basis of eigen vectors $\{v_1, v_2, \dots v_n\}$, that is

$$A^\dagger A v_i = \lambda_i v_i$$

Let

$$\begin{aligned} \sigma_i &= \sqrt{\lambda_i} \quad \text{and} \\ u_i &= \frac{Av_i}{\lambda_i} \end{aligned}$$

Claim: $\langle u_i, u_j \rangle = \delta_{ij}$
Indeed,

$$\begin{aligned}\langle u_i, u_j \rangle &= \langle \frac{Av_i}{\sigma_i}, \frac{Av_j}{\sigma_j} \rangle \\ &= \frac{1}{\sigma_i \sigma_j} \langle v_i, A^\dagger A v_j \rangle \\ &= \frac{1}{\sigma_i \sigma_j} \langle v_i, \lambda_j v_j \rangle \\ &= \frac{\lambda_j}{\sigma_i \sigma_j} \langle v_i, v_j \rangle = \delta_{ij}\end{aligned}$$

Finally extend $\{u_1, u_2, \ldots, u_r\}$ to an ONB $\{u_1, u_2, \ldots, u_m\}$. □

THEOREM 5.35 (Polar Decomposition). *Let $A : V \to V$ be a linear operator on a vector space V then A can be decomposed as $A = WP$, where W is a unitary matrix and W positive semi definite matrix.*

PROOF. From the singular value decomposition

$$\begin{aligned}A &= U\Sigma V^\dagger \\ &= UV^\dagger V \Sigma V^\dagger\end{aligned}$$

Set $UV^\dagger = W$ and $V\Sigma V^\dagger = P$. W is unitary since it is a product of unitaries.

$$\begin{aligned}\langle V\Sigma V^\dagger x, x \rangle &= \langle \Sigma V^\dagger x, V^\dagger x \rangle \\ &= \langle \Sigma y, \Sigma y \rangle \geq 0 \ V^\dagger x = y\end{aligned}$$

where the last inequality follows because $\sigma_i's$ are positive. □