Algebra 1

Ian Poon

August 2024

Selected theorems from `Micheal Artin`'s textbook [1] used in `MIT 18.701 Algebra I`

Contents

1 Matrices 1
1.1 Big formula Determinant and Inverse 10

2 Groups 14
2.0.1 groups and subgroups 14
2.0.2 homomorphisms 17
2.0.3 equivlence relations and partitions 19
2.0.4 Cosets 20
2.0.5 The Correspondance Theorem 23

3 Vector Spaces 24

4 Linear Operators 25
4.0.1 The Matrix of a Linear Transformation 25
4.1 Eigenvectors 30
4.2 The Characteristic polynomial 32
4.3 Jordan Form 35

5 Applications of Linear Operators 37

6 Bilinear Forms 38
6.1 symmetric forms 40
6.2 hermitian forms 41
6.3 orthogonality 43
6.4 the spectral theorem 48

1 Matrices

Definition 1 (Identity Matrix)

Let I_n denote the $n \times n$ identity matrix.

$$\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

If A is the $n \times n$ matrix then $AI_n = A$.Product follows $[n \times n] \times [n \times n]$

If A is the $m \times n$ matrix then $I_m A = A$.Product follows $[m \times m] \times [m \times n]$

Proposition 2

Consider 2 square matrices of same dimensions

$$A_1 = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & [B_1] \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & [B_2] \end{bmatrix}$$

Let A_1, A_2 be an indentity matrix with except $[B_1], [B_2]$ which are square block matrix of same dimensions then

$$A_1 A_2 = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & [B_1 B_2] \end{bmatrix}$$

Proof. Consider the matrix product as the sums of elemenatary matrices

$$\sum_{i,k} \sum_{j} e_{ij} e_{jk} (x_{ij})(y_{jk})$$

suppose the top left most element of $[B_1], [B_2]$ is at (m, m) and y_{jk}, x_{ij} is the value of the element at $(i, j), (j, k)$ in A_1, A_2 respectively Hence for elements outside of $[B_1 B_2]$ clearly they are

$$\sum_{i}^{m-1} \sum_{k}^{m-1} e_{ii} e_{ik} = \sum_{i,k}^{m-1} e_{ik}$$

which is essentially the same original identity matrix. But items in $[B_1 B_2]$ we have

$$\sum_{i=m}^{n} \sum_{k=m}^{n} \sum_{j=m}^{n} e_{ij} e_{jk} (x_{ij})(y_{jk})$$

which is essentially just a normal matrix product restricted to that section in the matrix.

Definition 3 (Inverse Matrix)

Let A be a $n \times n$ matrix. If there is matrix B such that

$$AB = I_n \quad \text{and} \quad BA = I_n$$

then B is the inverse of A and A is **invertible**

Lemma 4

If A is a square matrix that has a right inverse R where $AR = I$ and a left inverse where $LA = I$ then $R = L$

Proof. $R = IR = (LA)R = L(AR) = LI = L$

Proposition 5

Let A and B be invertible $n \times n$ matrices. Then

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

Proof. Consider

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

and

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

likewise too.

Remark 6. *by induction we see that*

$$A_1A_2 \dots A_m{}^{-1} = A_m{}^{-1}A_{m-1}{}^{-1} \dots A_1{}^{-1}$$

Lemma 7

A square matrix that has either a row of zeros or a column of zeros is not invertible

Proof. Let A be a $n \times n$ matrix.Let B be any other arbitrary $n \times n$ matrix. If a row of A is zero, then a row of the product AB will be zero as well. Same can be said for a zero column in A. Hence AB is not the identity which has every diagonal element is 1 and thus clearly no zero columns or row. Thus there is no possible right inverse. A similar argument can be made for the left inverse.

Definition 8 (Matrix Units)

The **matrix unit** $e_{i,j}$ of an $m \times n$ matrix M, is also a $m \times n$ matrix but with all positions being 0 except position in row i and column j. Analogously the matrix unit for **column vectors** is just defined by e_i where i is simply row/item i in the column vector.

Fact 9

The set of matrix units is called a basis for the space of all $m \times n$ matrices because every $m \times n$ matrix A is a *linear combination* of $e_{i,j}$

$$A = \sum_{i,j} a_{ij} e_{ij}$$

where $a_{ij} \in \mathbb{K}$ is an element in row i and column j of the matrix A For convenience we can also denote that

$$A = (a_{ij})$$

Fact 10

For the following formulas for multiplying matrix units and standard basis vectors hold

$$e_{ij} e_j = e_i \text{ and } e_{ij} e_k = 0 \text{ if } j \neq k$$

It follows directly from how matrix multiplication is defined

$$e_{ij} e_{j,k} = e_{i,k}$$

Definition 11 (Elementry Row Operations)

There are 3 types of elementary matrices that perform elementary matrix operations. We will first define them for **elementary row operations**

(i) add a·(row j) to X to (row i)

(ii) interchange (row i) and (row j) of X

(iii) multiply (row i) of X with non zero scalar c

We will define for columns later

Lemma 12

Elementry matrices are invertible and their inverses are also elementry matrices

Proof. There exists elementry matrices that are inverse operations to all elementry matrices. The opposite of (i) is "subtract a·(row j) to X to (row i)". To reverse (ii) imply apply (ii) again. For (iii) just divide instead by the same non zero scalar c used to multiply at first.

Proposition 13

The systems $A'X = B'$ and $AX = B$ have the same solutions.

Consider a sequence of elementry matrices that do

$$PAX = PB \Rightarrow A'X = B'$$

From the other direction we have

$$P^{-1}A'X = P^{-1}B \Rightarrow AX = B$$

Definition 14 (row echleon matrix)

A **row echleon matrix** is one that

(a) if (row i) of M is zero, then (row j) is zero for all $j > i$

(b) if (row i) isn't zero, then its first nonzero entry is 1. This entry is called a **pivot**

(c) if (row(i+1)) isn't zero, the pivot in (row(i+1)) is to the right of the pivot in (row i)

(d) the entries above the pivot are zero and by (c) so are the entries below the pivot

Fact 15

We denote a solution X to be trivial if it is equal 0 and non-trivial otherwise. Within non-trivial solutions we distinguish it between unique and infinite solutions(having free variables). It is clear the number of pivots is given by $r = \min(n, m)$ where n is the number of rows and m is the number of columns in the **row echelon matrix**

Example 16 (Free Variables)

Consider a system $AX = B$ in **augmented matrix** form

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

There are clearly no restrictions on what value x_3 can take so we can assign it an arbitrary **non-trivial** solution

Example 17 (Trivial Solution)

Consider a system $AX = B$ in **augmented matrix** form

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

Clearly all $x_i = 0$ so the solution is **trivial**

Example 18 (Unique Solution)

Consider a system $AX = B$ in **augmented matrix** form

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

Clearly $X = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ so the solution is **unique**

Corollary 19

Every system $AX = 0$ of m homogeneous equations in n unknowns with $m < n$ has a solution with X in which some x_i is nonzero

Proof. With m rows there are at most m pivots which is less than n. Then we clearly have free variables we can assign to x_i

Lemma 20

A square matrix is either the indentity matrix I or else its bottom row is zero.

Proof. If there is missing pivot, then there definitely exist a zero row. By the definition of *row echelon matrix*, all subsequent row below will be zero as well.

Theorem 21

Let A be a square matrix. The following statements are equivalent

1. A can be reduced to the indentity by a sequence of elementary row operations
2. A is a product of elementrary matrices
3. A is invertible

Proof. (a) implies (b) as we have proven earlier that the inverse just consists of elementary row operations.(b) implies (c) too since we have also shown earlier that every elementary row operations can be reversed by another elementary row operation.(c) implies (a) since if A is invertible $PA = A'$ must be invertible too since P which is just a sequence of elemantary row operations used during row reduction is invertible. By the above lemma, if A' is invertible it cannot have a zero row thus A' must be the identity matrix showing that A can indeed be reduced to row echleon form as desired for (a).

Theorem 22

The following statements of a square matrix A are equivalent

1. A is invertible
2. the system of equations $AX = B$ has a unique solution for every column vector B
3. The system of homogenous equations $AX = 0$ has only the trivial solution $X = 0$

Proof. (a) implies (b) and that means $A' = I$. so the unique solution $A'X = B'$(after row reduction) is just B' and that we have proven earlier these are the same solutions to $AX = B$. (b) implies (c) obviously. By contradiction suppose (a) is not invertible. Then we know that there is a zero row. However a zero row in a homogenous equation has a non-trivial solution since we can assign a free variable. Hence it can only be that (c) implies (a).

Definition 23 (Matrix Transpose)

The **transpose** of a matrix is a reflection about the diagonal. That is a_{ij} in A is a_{ji} in A^T.

Example 24

Every transpose operation on each row of A looks like so

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

So to get the transpose by hand simply go down the rows of the original matrix and add them as a column in the transpose from right to left.

Proposition 25

The following properties of transpose hold

1. $(AB)^T = B^T A^T$
2. $(A+B)^T = A^T + B^T$
3. $(cA)^T = c(A)^T$
4. $(A^T)^T = A$

Proof. done before pretty straightforward...

Definition 26 (Elementry Column Operations)

Consider the 3 types of **elementry column operations**

(i) add a·(col i) to X to (col j)

(ii) interchange (row i) and (row j) of X

(iii) multiply (row i) of X with non zero scalar c

We can immediately infer the following from the properties of transpose. First recall that elementry row operations are *left operations*. That is if you recall

$$E_k \dots E_2 E_1 A = A'$$

To work with columns with simply want $A^T \to A'^T$ then we see that

$$A^T E_1^T E_2^T \dots E_k^T = A'^T$$

so that we know **elementry column operations** work on the *right* instead, contrary to that elementary row operations.

Remark 27. *Note that due to the orientation of transposes, in (i), the indices i, j are swapped as compared to those in elementary row operations*

Definition 28 (Determinant)

The determinant is a unique function det defined on the space of $n \times n$ matrices with the following properties

1. $\det(I) = 1$
2. (linearity) det is linear in the rowspace of square matrix A
3. if two adjacent rows in matrix A are equal $\det(A) = 0$
4. (multiplicity) $\det(AB) = \det(A)\det(B)$

Theorem 29

The following properties of matrices hold

(a) If A' is obtained from A by adding a multiple of (row j) of A to (row i) and $i \neq j$ then $\det A' = \det A$

(b) If A' is obtained from interchanging (row j) of A with (row i) and $i \neq j$ then $\det A' = -\det A$

(c) If A' is obtained from A by multiplying (row j) by a scalar c then $\det A' = c \det A$

(d) if (row i) of A is equal to a multiple of (row j) and $i \neq j$ then $\det A = 0$

Proof. (c) follows by the linearity of rows. (c) and the definition of determinants imply (d).

$$\det \begin{bmatrix} cR \\ R \end{bmatrix} = c \det \begin{bmatrix} R \\ R \end{bmatrix} = 0$$

(d) implies (a) as

$$\det \begin{bmatrix} R + cS \\ S \end{bmatrix} = \det \begin{bmatrix} R \\ S \end{bmatrix} + c \det \begin{bmatrix} S \\ S \end{bmatrix}$$

It can be shown with repeated use of (a) that (b) is implied

$$\det \begin{bmatrix} R \\ S \end{bmatrix} = \det \begin{bmatrix} R - S \\ S \end{bmatrix} = \det \begin{bmatrix} R - S \\ S + (R - S) \end{bmatrix} = \det \begin{bmatrix} R - S \\ R \end{bmatrix} = \det \begin{bmatrix} -S \\ R \end{bmatrix} = -\det \begin{bmatrix} S \\ R \end{bmatrix}$$

Corollary 30

Consider the space of $n \times n$ matrices.Let E be an elementary matrix. Then for any matrix A, $\det(EA) = \det(E)\det(A)$

1. if E is of the first kind(add a multiple of one row to another), then $\det E = 1$
2. if E is of the second kind(row interchange) then $\det E = -1$
3. if E is of the third kind(multiply a row by c), then $\det E = c$

Proof. Consider

$$\det(EA) = \varepsilon \det A$$

When you let $A = I$. Then from 29 we see

$$\det E = \varepsilon$$

so it follows that ε must correspond to ± 1 and c corresponding to the respective cases.

Corollary 31

The following properties also follow

(a) a square matrix A is invertible if and only if its determinant is different from zero.

(b) $\det(A^{-1}) = \det A^{-1}$

(c) $\det A = \det A^T$

Proof. for (a) consider that if a A is invertible, then A'(row echeolon form) is the identity matrix else it is the identity matrix with zero rows below. Consider a sequence of elementry matrix operations to get to $A \to A'$

$$\frac{\det A'}{(\det E_1)(\det E_2)\dots(\det E_m)} = (\det A)$$

clearly $\det A' = 1$ if invertible and zero otherwise. for(b) consider by *multiplicative property*

$$\det(A^{-1})\det(A) = \det(AA^{-1}) = \det(I) = 1 \quad \Rightarrow \quad \frac{1}{\det A} = \det(A^{-1})$$

. For (c) let $A' = E_m \dots E_1 A$. Then recalling 26 we have

$$A^T E_1^T E_2^T \dots E_m^T = A'^T$$

Firstly for any elementary matrix E_i we have

$$\det E_i = \det E_i^T$$

See the proof below. And moreover $A' = A'^T = I$. If A is not invertible neither is A^T simply consider the following lemma. Because A is not invertible, A^{-1} is not defined and so will $A^{T^{-1}}$. Thus (c) follows.

Proposition 32

The 29 works if *row* is replaced with *column* throughout

Proof. The transpose of each elementary matrix row operation is the equivalent for that of the column. Moreover their determinants are the same as proven below. So this follows

Lemma 33

$(A^T)^{-1} = (A^{-1})^T$

Proof. Taking A^T on both sides

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

immediately implies the concusion

Proposition 34

consider the elementary matrix operations

$$\det E_i = \det E_i^T$$

Proof. Note in the following diagrams let row i be higher above than row j . These are the rows that differ from the identity matrix

Type 1: interchange row i and j

$$\begin{bmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & 0 & & & 1 & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & 1 & & & & 0 \\ & & & & & & & 1 \\ & & & & & & & & \ddots \end{bmatrix}$$

Follows by symmetry Type 2: Add a times of row j to row i

$$\begin{bmatrix} 1 & & & & & \\ & 1 & \dots & a & & \\ & & \ddots & \vdots & & \\ & 0 & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

Note that the diagonal is all 1. Using the determiant cofactor formula on the column containing a we have

$$\det M = (-1)^{i+j} a \det I + (-1)^{j+j} \det I$$

By symmetry the result is the same for M^T Type 3: Scale row i by a

$$\begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & c & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{bmatrix}$$

Follows by symmetry

1.1 Big formula Determinant and Inverse

Fact 35 (Determinant Formula)

The formula for the determinant of a square matrix is given by:

First pick any row i of your choice then

$$\det A = \sum_{j}^{n} a_{ij} C_{ij}$$

where the cofactor C_{ij} is defined by

$$C_{ij} = (-1)^{i+j} \det M_{ij}$$

and M_{ij} is the matrix with ith row and jth column removed

Alternatively pick any column j of your choice

$$\det A = \sum_{i}^{n} a_{ij} C_{ij}$$

Fact 36 (Permutation formula for determinants)

Let A be an $n \times n$ matrix then

$$\det A = \sum_{\sigma} \operatorname{sgn}(\sigma) a_{n,\sigma(n)} \dots a_{1,\sigma(1)}$$

Refer to `manifolds and analysis` to know why this definition makes sense(it has got to do with maps between alternating tensors) noticing that each row defines a linear map $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}$. Note that the co factor definition and other formulas were derived from this definition.

Proposition 37

The determinant of a $n \times n$ **upper triangular matrix** A which are of the form

$$A = \begin{bmatrix} a_{11} & \dots & * \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{bmatrix}$$

where everything below the diagonal is zero while the value of rest of the elements is immaterial is simply the product of diagonal entries that is

$$\det A = a_{11} \dots a_{nn}$$

Proof. Let our induction hypothesis be that for every sub square matrix $A_{i,i}$ where (i, i) are the coordinates of the top leftmost element of this subsquare matrix in A we have

$$\det A_{i,i} = a_{i,i} \det A_{i+1,i+1}$$

Our base case is (n, n) which is obviously true since

$$\det A_{n,n} = a_{n,n}$$

$$A = \begin{bmatrix} \ddots & & \vdots \\ & a_{n-1,n-1} & * \\ 0 & 0 & a_{nn} \end{bmatrix}$$

Let us prove for $\det A_{i-1,i-1}$, it is given by

$$\det A_{i-1,i-1} = a_{i-1,i-1} \det A_{i,i}$$

This is true because if we calculate the determinant by formula(see above) using the leftmost column we have

$$\det A_{i-1,i-1} = \sum_{j}^{i-1} a_{j,i-1} C_{j,i-1}$$

but because the only nonzero term is $a_{i-1,i-1}$(everything below in the same column is zero in upper triangular). And $C_{i-1,i-1} = (-1)^{2i} \det A_{i,i} = \det A_{i,i}$ Therefore proving our induction hypothesis.

Fact 38 (Inverse formula)

The formula for the inverse of a square matrix is given by:

$$A^{-1} = \frac{C^T}{\det A}$$

where C is the cofactor matrix

Proposition 39

The inverse of a diagonal square matrix, that is of the form

$$A = \begin{bmatrix} a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{bmatrix}$$

where the only the values of the diagonal are immaterial but everywhere else is zero is found by simply replacing all diagonal elements with its reciprocal

Proof. Consider that

$$A = \begin{bmatrix} a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 1/a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1/a_{nn} \end{bmatrix} = I$$

where I is the identity matrix. To see why, we have:

$$A = \sum_{i,k} \sum_{j} e_{i,j} e_{j,k} (x_{i,j})(y_{j,k}) = \sum_{i} e_{ii} e_{ii} (a_{ii})(1/a_{ii})$$

where $(x_{i,j})(y_{j,k})$ are simply the values of the elements at those positions in each matrix respectively. This relation follows because those are the only non-zero terms in the matrix product.

Definition 40 (Permutation)

Using notations for elementary matrices e_{ij} and e_i(column vectors) as defined earlier we can express permutation matrices by

$$PX = \left(\sum_i e_{pi,i}\right)\left(\sum_j e_j x_j\right) = \sum_{i,j} e_{pi,i} e_j x_j = \sum_i e_{pi,i} e_i x_i = \sum_i e_{pi} x_i$$

The 3rd and 4th equality follows if you recall 10

You can clearly see what the permutation matrix does

$$\sum_i e_i x_i \to \sum_i e_{pi} x_i$$

or equivalently

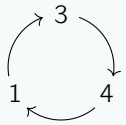
$$P(e_i x_i) = e_{pi} x_i$$

Originally x_i is found in row i in the column vector. Now x_i is found in row p_i of the column vector

Example 41

$$p = (341)(25) \tag{1}$$
$$q = (1452) \tag{2}$$
$$qp = p \circ q = (135) \tag{3}$$
$$pq = q \circ p = (234) \tag{4}$$

In each permutation p, q, pq, qp the brackets represent a cycle. For example (341) in p means



Lets consider $q \circ p$

$$3 \to 4 \to 5 \tag{5}$$
$$4 \to 1 \to 4 \tag{6}$$
$$1 \to 3 \to 3 \tag{7}$$
$$2 \to 5 \to 2 \tag{8}$$
$$5 \to 2 \to 1 \tag{9}$$

To see what composition of permutations means take for example according to (5), 3 goes to 4 in permutation p but 4 goes to 5 in permutation q. So 3 goes to 5 in $q \circ p$. Now (5) and (7) implies 4 and 2 are fixed in permutation pq so we don't write them in any cycle by convention. Now following the arrows for the rest we can tell

$$pq = (135)$$

Proposition 42

The following about permutation matrices are true

(a) A permutation matrix P always has a single 1 in each row and in each column, the rest of its entries are 0. Conversely, any such matrix is a permutation matrix

(b) The determinant of a permutation matrix is ± 1

(c) Let p, q be permutations associated with permutation matrices P, Q respectively. Then the $p \circ q$ is associated with PQ

Proof. (a) is follows 40

$$P = \sum_i e_{pi,i}$$

This means for each column we will have only 1 row with "1", the rest will be zero. Moreover each $p_i \neq p_j$ if $i \neq j$ because that will mean x_i and x_j will both be in row $p_i = p_j$ of the new column matrix which doesn't make any sense.(b) follows from (a) and properties of the determinant. (c) follows if you consider

$$PQ = \sum_i e_{pi,i} \sum_j e_{qj,j} = \sum_{ij} e_{pi,i} e_{qj,j} = \sum_j e_{pqj,qj} e_{qj,j} = \sum_j q_{pqj,j}$$

if you again recall the properties we know from 10. It is evident that $e_{pi,i} e_{p_j,j} = e_{pi,j}$ when $i = p_j$ as desired by the composition $p \circ q$.

Remark 43. *Matrix products correspond to compositions of linear maps in general eg.* $A : \mathbb{R}^n \to \mathbb{R}^k$ *and* $B : \mathbb{R}^m \to \mathbb{R}^n$ *then*

$$AB = \mathbb{R}^m \to \mathbb{R}^k$$

each matrix takes in column vector and outputs a result column vector that the next matrix will work with

Definition 44 (Sign of Permuation)

Every permutation has a **sign**, so we can send $S_n \to \{\pm 1\}$ where we send to 1 is the sign is positive(an **even** permuation) and -1 if the sign is negative(an **odd** permutation)

To see this consider that every $p \in S_n$ corresponds to permuation matrix P. Then the sign is defined to be the determinant of P because permutations essentially swap rows of the indentity matrix. Therefore each swap multiplies the determinant by -1 so the sign of the matrix is $(-1)^n$ where n is the number of transpositions or the order of S_n.

2 Groups

2.0.1 groups and subgroups

Definition 45 (Law of composition)

A **law of composition** on set S is any rule of combining *pairs* $a, b \in S$ to get another element say $p \in S$. Formally it represents the map:

$$S \times S \to S$$

Example 46

$$p = ab, a \times b, a \circ b, a + b$$

are all possible rules defined by the law of composition

Definition 47 (Group)

A **group** is a set G with a law of composition that has the following properties

- The law of composition is **associative**: $(ab)c = a(bc)$ for all $a, b, c \in G$
- G contains an **identity** element 1, such that $1a = a$ and $a1 = a$ for all $a \in G$
- Every element a of G has an **inverse**: an element b such that $ab = 1$ and $ba = 1$

An **abelian group** is a group whose law of composition is **commutative**. We denote the **order** of a group G by

$$|G| = \text{number of elements contained in G}$$

Theorem 48 (Cancellation Law)

Let a, b, c be elements of group G whose law of composition is written multiplicatively.Then if $ab = ac$ or if $ba = ca$ then $b = c$. If $ab = a$ or if $ba = a$ then $b = 1$.

Proof. Simply take inverse(which exists for every element by definition) on both sides for say $a^{-1}ab = a^{-1}ac \Rightarrow b = c$ as desired. Now do this to verify the rest noting that the identity exists in G so $b = 1$ is possible.

Example 49

The $n \times n$ *general linear group* is the group of all invertible $n \times n$ matrices denoted by

$$GL_n = \{n \times n \text{ invertible matrices A}\}$$

We write $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ if we are working with real or complex numbers respectively

Example 50

The group of permutations of the set of *indices* $\{1, 2 \ldots, n\}$ is called the **symmetric group** and is denoted by S_n. It is a finite group of order $n!$. To re-emphasize, there are $n!$ sets of length n each containing *positional indices* not values!

Example 51

To describe say a symmetric group S_3 suppose we have the specific pair of cyclic permutations $x = (1, 2, 3)$ and $y = (1, 2)$ which you generalize for other permutations. Then the group law/law of composition can be easily defined for any pair of x, y by

$$x^3 = 1, \quad y^2 = 1, \quad yx = x^2y$$

That is because x, y are 3 and 2 permutation cycles respectively. For example in this case $x^3 = x \circ x \circ x = x$.The other relations can be verified too similar to how we did in 41. Now, all possible distinct elements in G(in our case permutations in S_3) that we can infer from our group laws are

$$S_3 = \{1, x, x^2, y, xy, x^2y\}$$

By definition of group and clearly from group laws we see that the indentity element 1,obviously x, y themselves must be inside and that yx or x^2y is inside(just take 1 since both same). However our law of composition here is not commutative since we cannot obtain $xy = yx$ from any of our group laws or applying *cancellation law* to simplify them.Thus we add in xy as a distinct element. Also clearly x^2 is a distinct element because x is a 3 cycle. Also $x^3 = x, y^2 = y$, we don't need to add them in since x, y already inside.

Remark 52. *A general strategy is just loop through powers/linear combinations of x, y depending on whether our group law of multiplicative/additive and check if distinct.*

Fact 53

Note that we use $(\mathbb{R}, +)$ to denote the **additive** group of reals to distinguish with $\mathbb{R}^+$, the set of *positive* real numbers .Also, $\mathbb{R}^\times$ denotes the **multiplicative** group of reals

Definition 54 (Order)

An element A of a group has **order** n if A^n is the identity element

Example 55

Consider matrix

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

Since $A^6 = I$ then $A \in GL_2$ with order 6

Definition 56 (subgroup)

A **subgroup** H of group G is a subset of the the group with the *same* law of composition. It is **closed** that is, if $a, b \in H$ then $ab \in H$. In addition 1_G the indentity of G must also be in H and all the inverses a^{-1} of $a \in H$ must also be in H.

Note that group G defines a law of composition on H which then determines the closure conditions for H. We call this the **induced law**. The other conditions are simply to ensure, H is also a group in its own right.

Remark 57. *Note that we have definee this assuming the law of composition is written multiplicatively. Note that we can also define it similary if the law of composition was written additively instead, that is the closure condition is now if $a, b \in H$ then $a + b \in H$*

Example 58

The subset of all multiples of a as denoted by $\mathbb{Z}a$ is a subgroup of the group of $\mathbb{Z}$

Example 59

The **special linear group** SL_n is the set of matrices in GL_n of determinant 1 is a subgroup of GL_n because the determinant is mulitplicative. So clearly for any $a, b \in SL_n$ where $\det(ab) = \det(a)\det(b) = 1$ so $ab \in SL_n$ as well.

2.0.2 homomorphisms

Definition 60 (Homomorphism)

Let G and G' be groups written in multiplicative notation. A **homomorphism** is map from G to G':

$$\varphi : G \to G'$$

where $\forall a, b \in G$

$$\varphi(ab) = \varphi(a)\varphi(b)$$

Similarly if G is written in additive notation while G' in multiplicative we have:

$$\varphi(a + b) = \varphi(a)\varphi(b)$$

We will observe more examples below. Essentially a homorphism is simply a map between 2 groups that is compatible with their **laws of composition**.

Example 61

The following maps are homomorphisms:

(a) the determinant function $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times$

(b) the exponential map exp: $(\mathbb{R}, +) \to \mathbb{R}^\times$

Proof. (a) is clear from the *multipicity* of determinant. However for (b) we have to define our homomorphism such that it is additive in the *domain* and multiplicative in the *range* as required in the question. Hence we have:

$$\varphi(a + b) = \varphi(a)\varphi(b)$$

we can clearly that this is a consequence of

$$e^{a+b} = e^a e^b = e^b e^a$$

Proposition 62

Let $\varphi : G \to G'$ be a group isomorphism where G, G' are written multiplicatively

(a) if $a_1, \ldots, a_k$ are elements of G, then $\varphi(a_1 \ldots a_k) = \varphi(a_1) \ldots \varphi(a_k)$

(b) $\varphi(1_G) = 1_{G'}$

(c) $\varphi(a^{-1}) = \varphi(a)^{-1}$

Proof. (a) follows from induction from the definition. (b) follows from

$$\varphi(1_G)\varphi(1_G) = \varphi(1_G 1_G) = \varphi(1_G)$$

Then applying *cancellation law* on both sides we have

$$\varphi(1_G) = 1_{G'}$$

(b) implies(c), since we have $\varphi(a^{-1})\varphi(a) = \varphi(a^{-1}a) = \varphi(1_G) = 1_{G'}$ Therefore we have

$$\frac{1_{G'}}{\varphi(a)} = \varphi(a^{-1}) \Rightarrow \varphi(a^{-1}) = \varphi(a)^{-1}$$

Definition 63

The **image** of the group homomorphism $\varphi : G \to G'$ is defined by

$$\text{im } \varphi = \{x \in G' | x = \varphi(a), a \in G\}$$

The **kernel** of the group homomorphism $\varphi : G \to G'$ is defined by

$$\ker \varphi = \{a \in G | x = \varphi(a) = 1\}$$

Assuming G, G' are written multiplicatively, the Kernel is then a subgroup of G since if $a, b \in K$,

$$\varphi(ab) = \varphi(a)\varphi(b) = 1$$

so $ab \in K$

Definition 64 (Trivial Homomorphism)

A map between groups defined by $\varphi : G \to G'$ such that every element in group G is mapped to the identity in G' that is

$$\varphi(g) = I$$

for all $g \in G$ and $I \in G'$ is the identity

Definition 65 (Inclusion Map)

A map between groups defined by $\varphi : H \to G$ such that

$$i(x) = x, x \in H$$

where H is a subgroup of G

Definition 66 (Normal Subgroup)

A subgroup N of group G is a **normal subgroup** if for every $a \in N$ and $g \in G$ then the **conjugate** $gag^{-1} \in N$

Definition 67 (Center)

The **center** of group G,denoted by Z is the set of elements in G that commute with every element of G

$$Z = \{z \in G | zx = xz\}$$

for all $x \in G$

Definition 68 (Isomorphism)

A **isomorphism** is homomorphism $\varphi : G \to G'$ defined by a **bijective** map. Suppose this is true then it is clear then there exists a isomorphism $\varphi^{-1} : G' \to G$ too. We then say the 2 groups G, G' are **isomorphic** which we denote by the symbol

$$G \approx hG'$$

Definition 69 (Automorphism)

An isomorphism from a group to itself. That is

$$\varphi : G \to G$$

2.0.3 equivlence relations and partitions

Definition 70 (Partition)

We define a **partition** which we denote as $\prod$ of a set S *subdivision* of S into *non-overlapping,non-empty* subsets. That is

$$S = \text{union of disjoint nonempty subsets as specified by the partition}$$

Definition 71 (Equivalence Relation)

A relation is an equivalence relation if it is

- reflexive: $a \sim a, \forall a \in A$
- symmetric: $a \sim b \to b \sim a, \forall a, b \in A$
- transitive: $(a \sim b)$ and $(b \sim a) \to a \sim c, \forall a, b, c \in A$

Lemma 72 (Equivlance Class)

We define a subset of set S called the equivalence class of element a by

$$C_a = \{b \in S | a \sim b\}$$

Given an equivalence relation on set S, the subsets of S that are equivalence classes partition S

Proof. For any element $a \in S$, the reflexive property tells us equivalence class of a is non-empty as it contains a itself. We now need to prove equivalence classes are disjoint to get the conditions required for being a partition. If two equivalence classes are not disjoint then there exists $x \in S$ such that $a \sim x$ and $b \sim x$ corresponding to C_a and C_b respectively. Then for any element y where $a \sim x$(that is $y \in C_a$) implies $y \sim a \sim x \sim b$ by properties of equivalence relations. Likewise for any element g where $g \sim b$(that is $g \in C_b$) we can get $g \sim a$. Therefore $C_a = C_b$ and $a = b$. Therefore the contra-positive of this means as long as $a \neq b$, they are disjoint. Therefore different equivalence classes are disjoint.

Definition 73 (Inverse Image/Fibre)

Suppose we have any map of sets $f : S \to T$ The **inverse image** of an element t of T is denoted by

$$f^{-1}(t) = \{s \in S | f(s) = t\}$$

The inverse image is also known as the **fibre** of f

Proposition 74

An equivalence relation on a set S determines a partition of S and conversely

Proof. Proving from the backward direction, it is obvious that $a \sim b$ if a and b are in the same subset defined by the partition satisfies all axiom required for an equivalence relation. The forward direction follows from the lemma above.

2.0.4 Cosets

Definition 75 (Left Coset)

If H is a subgroup of G and a is an element in G, then the notation aH will stand for set of all products ah with h in H

$$aH = \{g \in G | g = ah, h \in H\}$$

A set is called a **left coset** of H in G

Proposition 76

Let $\varphi : G \to G'$ be a homomorphism of groups and G, G' be written multiplicatively. Let a and b be the elements of G. Let K be the kernel of φ. The following statements are equivalent

(a) $\varphi(a) = \varphi(b)$

(b) $a^{-1}b$ is in K

(c) b is the coset of aK

Proof. Suppose $\varphi(a) = \varphi(b)$. Then (a) and (b) imply each other bidirectionally since

$$\varphi(a^{-1}b) = \varphi(a^{-1})\varphi(b) = \varphi(a)^{-1}\varphi(b) = 1$$

(b) implies (c) since there exists $b = a(a^{-1}b)$ and $a^{-1}b \in K$.

Theorem 77 (Congruence Relation)
The **cosets** of H in G are equivalence classes for the **congruence relation**

$$a \equiv b$$

if $b = ah$ for some h in H

Proof. (Transitivity) Suppose $a \equiv b$ and $b \equiv c$. Then $b = ah$ and $c = bh$ for some $h, h' \in H$. Then $c = ahh'$. Since hh' is in H by definition of subgroup $a \equiv c$. The prove for the rest is rather trivial...

Corollary 78
Left cosets of a subgroup H of group G partition the group

Proof. Each left cosets are the equivalence classes for the congruence relation. Specially when $a \sim b$ means $a \equiv b$ if $b = ah$ for some $h \in H$ then $b \in C_a$. So for all $a \in G$, G will naturally be partitioned by all unique a(s).

Example 79
Consider the same group $G = S_3 = \{1, x, x^2, y, xy, x^2y\}$ from 51.Let the subgroup to be the permutations $H = \langle y \rangle$ of order 2. We again recall that as a group H must contain the identity too thus $H = \{1, y\}$. Then there are three possible left cosets of H in G:

$$H = \{1, y\} = yH \tag{1}$$
$$xH = \{x, xy\} = xyH \tag{2}$$
$$x^2H = \{x^2, x^2y\} = x^2yH \tag{3}$$

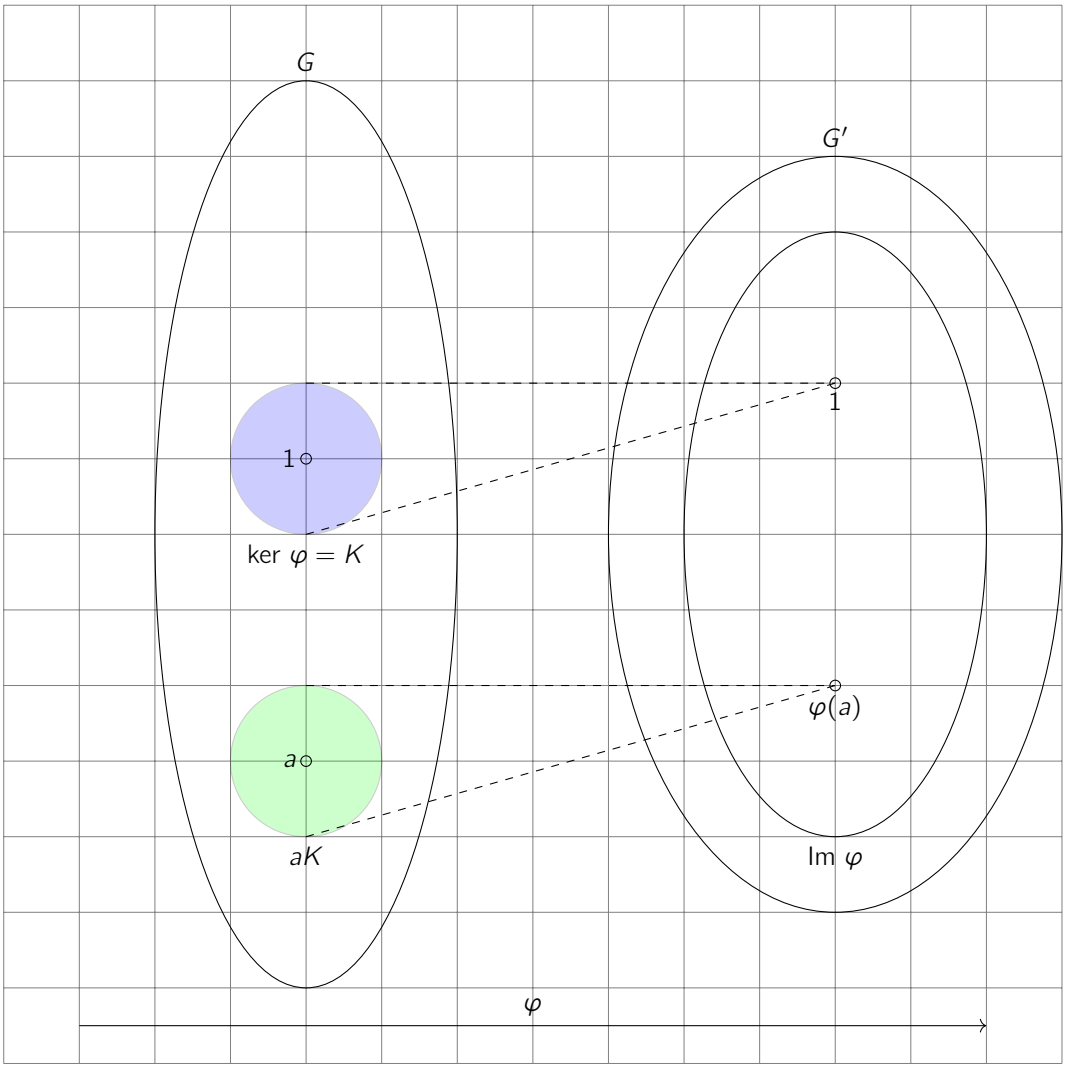
Thus we see that left cosets of H do indeed partition G

Proposition 80
Let K be the kernel of homomorphism $\varphi : G \to G'$ and G, G' are written multiplicatively. The fibre of φ that contains an element a of G is the coset aK of K.These cosets partition the group G and corrersponds to elements in the image of φ

Proof. Recall 63 that the kernel is a subgroup of G in this case. Therefore by 2.0.4 the left cosets of the kernel partition G. Recall from 76, suppose we fix some $a \in G$ and some corresponding $\varphi(a)$ in the image of φ. Then the inverse image of $\varphi(a)$ is defined by

$$\{b : G | \varphi(b) = \varphi(a)\}$$

Then we know that all such b must in the coset aK. Therefore every element in the left coset aK is mapped to an element $\varphi(a)$ in the image of φ



Remark 81. *As you can see if φ is not bijective then φ^{-1} is not necessarily a map since we have 1 element in G' mapped to multiple elements in G*

Definition 82 (Index)
The number of left cosets of a subgroup is called the **index** of H in G which we denote by:

$$[G : H]$$

This means if G is finite, the index may be infinte too. As for the example above

$$[S_3 : \langle y \rangle] = 3$$

Since we have 3 equivalence classes corresponding to 3 congruence classes.

Lemma 83
All left cosets aH of a subgroup H of a group G have the same order

Proof. The map $H \to aH$ is clearly is bijective as the inverse is simply a^{-1} which exists by definition of group for all $a \in G$.

Corollary 84 (Counting Formula)
The order of group G equals the product of the order of subgroup H and the number of left cosets of H

$$|G| = |H|\,[G : H]$$

Proof. Since cosets of subgroup H partition group G and the order of each coset is the same, that is equal H, then the conclusion clearly follows

Theorem 85 (Lagrange Theorem)
Let H be a subgroup of a *finite* group G. The order of H divides the order of G

Proof. Follows very obviously from the counting formula

Corollary 86
Let $\varphi : G \to G'$ be a homomorphism of finite groups and G, G' are written multiplicatively then

1. $|G| = |\ker\varphi|\,|\mathrm{im}\varphi|$
2. $|\ker\varphi|$ divides $|G|$ and
3. $|\mathrm{im}\varphi|$ divides both $|G|$ and $|G'|$

Proof. For (1) it is clear that from 80

$$[G : \ker\varphi] = |\mathrm{im}\varphi|$$

Combining this with the **counting formula**, we have

$$|G| = |\ker\varphi|\,[G : \ker\varphi] = |\ker\varphi|\,|\mathrm{im}\varphi|$$

. (1) clearly implies (2) and the first part of (3). For the second part of (3), given that $\mathrm{im}\varphi$ is a finite subgroup of G', it follows by **langrange theorem** that it divides $|G'|$ too.

Proposition 87 (Multiplicative Property of Index)
Let $G \subset H \subset K$ be subgroups of group G(can be finite or infinite). Then

$$[G : K] = [G : K][H : K]$$

Proof. $G = \bigcup_{n=1}^{\infty} g_n H$ while $H = \bigcup_{m=1}^{\infty} h_m K$. Combining these 2 we have

$$G = \bigcup_{n=1}^{\infty} g_n \left(\bigcup_{m=1}^{\infty} h_m K \right) = \bigcup_{n,m} g_n h_m K$$

2.0.5 The Correspondance Theorem

Lemma 88

Let $\varphi : G \to \mathcal{G}$ be a homomorphism with kernel K and let $\mathcal{H}$ be a subgroup of $\mathcal{G}$. Let $\varphi^{-1}(\mathcal{H}) = H$ Then

1. H is a subgroup of G that contains K
2. If $\mathcal{H}$ is a normal subgroup of $\mathcal{G}$ then H is a normal subgroup of G.
3. If φ is surjective and H is a normal subgroup of G, then $\mathcal{H}$ is a normal subgroup of $\mathcal{G}$.

Theorem 89 (Correspondance Theorem)

Let $\varphi : G \to \mathcal{G}$ be a surjective group homomorphism with kernel K. Then there is a bijective correspondace between subgroups of $\mathcal{G}$ and subgroups of G that contain K

3 Vector Spaces

Definition 90 (Commutative)

If changing the order of operands does not change the result that is

$$a + b = b + a \text{ and } ab = ba \text{ etc}$$

Definition 91 (field)

fields $\mathbb{F}$ are sets together with two laws of composition

$$F \times F \xrightarrow{\times} F \text{ and } F \times F \xrightarrow{+} F$$

that satisfy the following axioms

(i) Addition makes F into an **albenian group** F^+; its identity element is denoted by 0

(ii) Multiplication is commutative, and it makes the set of nonzero elements of F into an **albenian** group $F^\times$; its identity element is denoted by 1

(iii) *Distributive law*: For all a, b, c in F, $a(b + c) = ab + ac$

Lemma 92

Let F be a field

(a) The elements 0 and 1 of F are distinct

(b) For all a in F, $a0 = 0$ and $0a = 0$

(c) Multiplication in F is associative and 1 is an identity element

Proof. axiom (ii) means for any $s \neq 0 \to 1s = s$. Suppose $s = 1$, which is non-zero element. Then $1(1) = 1 \neq 0$. Hence (a) follows. For (b), axiom (1) means $0 + 0 = 0$. Using axiom (iii) we have $a0 + a0 = a(0 + 0) = a0$ for some $a \in F$. Using cacellation law for groups, we have $a0 = 0$. Because multiplication is commutative, $a0 = 0a = 0$. For

(c), consider $a, b, c \in \{F - 0\}$ the albenian group. By definition of group, operations are associative so $a(bc) = (ab)c$. We need to show that this still holds even if one of the elements is zero. In that case by (b), if this is so then $a(bc) = (ab)c = 0$ so multiplication in F is associative. Setting $a = 1$ in (b) shows that 1 is an indentity on all of F.

Fact 93 (matrix product)

Let the **hypervector** $S = (v_1, \ldots, v_n)$ be an ordered set where each v_i is in vector space V. Multiplication between vectors in a vector space is not defined but we do have **scalar multiplication**. Therefore the use of **matrix product** allows us to define scalar multiplication of such hypervectors with a column/coordinate vector X in $\mathbb{F}^n$ in a more convenient way as shown:

$$SX = (v_1, \ldots, v_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} v_1x_1 + \ldots + v_nx_n \end{bmatrix} = \sum_{i=1}^{n} v_ix_i \in V$$

We take the 1×1 matrix to be just $\begin{bmatrix} u \end{bmatrix} = u$ where u is some element. Suppose we had multiple column vectors X, then the result of the above would be another hypervector. You can see it is clearly more convenient to keep track of and save time writing out so many sums this way as you can just express such operations with matrices.

Proposition 94

Prove the following statements regarding base change matrices:

1. Fix *any* two bases $\mathbf{B}$ and $\mathbf{B}'$ of vector space V. The basechange matrix P is an invertible matrix that is determined *uniquely* by the bases $\mathbf{B}$ and $\mathbf{B}'$
2. Pick an *arbitruary* $\mathbf{B} = (v_1, \ldots, v_n)$ basis of vector space V. Any other base are sets of the form $\mathbf{B}' = \mathbf{B}P$ where P is some invertible $n \times n$ matrix.

Proof. For (1) Let $\mathbf{B} = (v_1, \ldots, v_n)$ and $\mathbf{B}' = (v_1', \ldots, v_n')$ be the respective basis expressed in hypervector form. Now suppose we have $\mathbf{B}' = \mathbf{B}P$. We know that each $v_j' = \sum v_i p_{ij}$ where each $p_{ij} \in \mathbb{F}$. Since basis are linear independent, each representation of v_i' using p_{ij} must be unique. Therefore P is unique. To show that P is invertible, first consider there must exist a unique change of basis matrix Q where $\mathbf{B} = \mathbf{B}'Q$ since our goal is to be able to translate between any valid pair of basis for V as stated in the proposition. Then we have

$$\mathbf{B} = \mathbf{B}'Q = \mathbf{B}PQ$$

Then it must be that $PQ = I$ only because the base change matrix between $\mathbf{B}$ and $\mathbf{B}$ must be unique a well.

For (2) we mut show that if $\mathbf{B}$ is a basis and if P is any invertible matrix then $\mathbf{B}' = \mathbf{B}P$ is a basis.Since P is invertible, $\mathbf{B} = \mathbf{B}'P^{-1}$. Therefore, $\mathbf{B}'$ spans V. Since it has the same number of elements/dimension(we are using $n \times n$ matrices) as $\mathbf{B}$, $\mathbf{B}'$ must be a basis as well.

4 Linear Operators

4.0.1 The Matrix of a Linear Transformation

Definition 95 (Linear Transformation)

A **linear transformation** $T : V \to W$ from one vector space over a field F to another is a map that is compatible with addition and scalar multiplication:

$$T(v_1 + v_2) = T(v_1) + T(v_2) \text{ and } T(cv_1) = cT(v_1)$$

"Over a field" is the fact that there is an isomporhism between $F^n \to V$ and $F^m \to W$ respectively. Refer to the commutative diagram and proofs below to learn more. In fact the map $F^n \to F^m$ is also a linear transformation by definition. We will learn that this map is done by matrices which satisfy the requirements for a linear transformation like so

$$A(X_1 + X_2) = AX_1 + AX_2 \text{ and } A(cX) = cAX$$

Remark 96. *This is analogous to homomorphisms of groups. In vector spaces V, W their law of composition is defined additively and with scalar multiplication. Therefore $(v_1 + v_2)$ in the range V and $T(v_1) + T(v_2)$ in the domain W reflects exactly that.*

Lemma 97

Let $T : F^n \to F^m$ be a **linear transformation** between spaces of column vectors and let the coordinate vector of $T(e_j)$ be $A_j = (a_{1j} \dots, a_{mj})^T$. Let A be the $m \times n$ matrix whose columns are $A_1, \dots, A_n$. Then T acts on vectors in F^n by **left multiplication** by A

Proof.

$$T(X) = T(\sum_j e_j x_j) = \sum_j T(e_j)x_j = \sum_j A_j x_j = AX$$

As you can see $\sum_j$ as gone from summing over rows to columns. The end result is still a column vector, just that it has m rows instead of n rows now.

Fact 98 (Hypervector)

The hypervector $T(\mathbf{B})$ where $\mathbf{B} = (v_1, \dots, v_n)$ is a basis of V is represented by

$$T(\mathbf{B}) = (T(v_1), \dots, T(v_n))$$

If $v = \mathbf{B}X = v_1x_1 + \dots + v_nx_n$ then by linearity

$$T(v) = T(v_1)x_1 + \dots T(v_n)x_n = T(\mathbf{B})X$$

Remark 99. *We also note that as seen, scalar multiplication of a hypervector(members in **vector space**) and coordinate/column vector(members in **field**) is done by **left multiplication** by the hypervector by convention. Thus for coordinate vectors and basis(a hypervector) it is done by **left multiplication** by the basis by convention. Same can be said for $T(\mathbf{B})$ and obviously for $\mathbf{B}' = \mathbf{B}P$ where P is the base change matrix which we will learn about next.*

Proposition 100

Let $\mathbf{B} = (v_1, \ldots, v_n)$ and $\mathbf{C} = (w_1, \ldots, w_m)$ be the bases of V and W respectively. Let X be the coordinate vector of an arbituary vector v with respect to basis $\mathbf{B}$ and let Y be the coordinate vector of its image $T(v)$ with respect to basis $\mathbf{C}$. So $v = \mathbf{B}X$ and $T(v) = \mathbf{C}Y$. Then there exists an $m \times n$ matrix A such that

$$T(\mathbf{B}) = \mathbf{C}A \text{ and } AX = Y$$

Proof. Every $T(v_j) \in W$ can be written as a linear combination of the basis $\mathbf{C}$ of W. So we have

$$T(v_j) = w_1 a_{1j} + \ldots + w_m a_{mj}$$

Therefore we get the following hypervector

$$T(\mathbf{B}) = (T(v_1), \ldots, T(v_n)) = (w_1, \ldots, w_m)\begin{bmatrix} & A & \end{bmatrix} = \mathbf{C}A$$

Where A is a $m \times n$ matrix. It is also clear that $A : F^n \to F^m$. Next if $v = \mathbf{B}X$ from 98 we have

$$T(v) = T(\mathbf{B})X = \mathbf{C}AX = \mathbf{C}Y$$

Fact 101

Consider the *isomorphisms*

$$F^n \to V \text{ and } F^m \to W$$

Injectivity each $v \in V$ is unique corresponds to a unique coordinate vector in F^n due to linear independence. Surjectivity because every $v \in V$ has some a representation as coordinate vectors in F^n because the basis spans V.

$$\begin{array}{ccc} F^n & \xrightarrow{A} & F^m \\ {\scriptstyle \mathbf{B}}\downarrow & & \downarrow{\scriptstyle \mathbf{C}} \\ V & \xrightarrow{T} & W \end{array}$$

This is called a **commutative diagram**

Proposition 102

Let A be the matrix of linear transformation T with respect to given bases $\mathbf{B}$ and $\mathbf{C}$

(a) Suppose that new bases $\mathbf{B}'$ and $\mathbf{C}'$ are related by the given matrices P and Q. The matrix of T with respect to the new bases is then

$$A' = Q^{-1}AP$$

(b) The matrices A' that repsent T with respect to other bases are those of the form $A' = Q^{-1}AP$ where Q and P can be any invertible matrices of the appropriate sizes

Proof. For (a) Firstly we obviously need to relate $\mathbf{B}' = \mathbf{B}P$ and $\mathbf{C}' = \mathbf{C}Q$ to some new X' and Y' with respect to the new basis respectively.Consider since our new coordinate vectors must satisfy

$$\mathbf{B}'X' = \mathbf{B}X$$
$$\mathbf{C}'Y' = \mathbf{C}Y$$

Then upon substitution we have $\mathbf{B}PX' = \mathbf{B}X$ and $\mathbf{C}QY' = \mathbf{C}Y$. Hence now susbtitute $X = PX'$ and $Y = QY'$(note the change to right multiplication by the base change matrix P) into $Y = AX$ obtaining $QY' = APX'$. Our objective is to find the new matrix of coordinate vectors A' of image vectors with respect to our new basis that is where $Y' = A'X'$. Thus we have $A' = Q^{-1}AP$. (b) follows because the base exchange matrices Q, P can be any invertile matrix if you recall 94

Fact 103 (Dual Properties of basechange matrix)

From the above, we observe that the **base-change matrix** relates the old and new basis by right multiplication on the *old* basis.

$$\mathbf{B'} = \mathbf{B}P$$

On the other hand it relates the old and new coordinate vectors by left multiplication on the *new* coordinate vector. $PX' = X$

Definition 104 (Linear Operators)

Linear operators are basically linear transformations $T : V \to V$ that map a vector space *to itself* (the exclusive feature)

Fact 105 (Reminder: Disjoint vector spaces)

When we say two subspaces/vector spaces are disjoint we can only say

$$V_1 \cap V_2 = \{0\}$$

which in proper terms is the **trivial intersection** and not

$$V_1 \cap V_2 = \emptyset$$

because every vector space must contain a zero vector(identity inverse) by definition. There is a clear difference between the empty set and the zero vector set

Proposition 106

Let K and W denote the kernel and image respectively of a linear operator T on a finite dimensional vector space V. The following are equivalent

(a) T is bijective

(b) $K = \{0\}$

(c) $W = V$

Proof. Suppose (b) is true. Then by the **rank nullity theorem**,

$$\dim V = \dim \ker T + \dim \operatorname{Im} T$$

So $\dim V = \dim \operatorname{Im} T$ is implied. Because since we are considering *linear operators* we know $\operatorname{Im} T \subset V$. Then we recall from `Axler Linear Algebra Done Right` that the basis for $\operatorname{Im} T$ must span the whole space V now since their

dimension are equal so $V = W$.So (b) implies (c). Recall that $\dim \ker T = 0$ implies injectivity and $\dim \operatorname{Im} V = \dim V$ implies surjectivity. Therefore (b) and (c) implies (a). Also recall that (a) implies (b) and (c) too.

Proposition 107

Let K and W denote the kernel and image respectively of a linear operator T on a finite dimensional vector space V. The following are equivalent

(a) V is the direct sum $K \oplus W$

(b) $K \cap W = \{0\}$

(c) $K + W = V$

Proof. Recall from `Axler Linear Algebra Done Right` that **rank nullity theorem** also applies to linear operators not just linear transformations. Hence

$$\dim V = \dim \ker T + \dim \operatorname{Im} T$$

is certainly still true regardless of (a),(b),(c). We didnt' consider the possibly of the direct sum of kernel and image being the domain for linear transformations because $W \notin V$ in the first place. But for linear operators we now can. Just as long (b) and (c) is satisfied since by definition (a) is true if and only if (b) and (c) are true.

Definition 108 (Invertible vs Singular Operator)

A linear operator is **invertible** if and only if the conditions in 106 are satisfied and **singular** otherwise

Proposition 109

A linear operator is singular if and only if its matrix with respect to an arbitary basis has $\det A = 0$

Proof. Consider the nullspace of linear operator T

$$T(v) = \mathbf{B}AX = 0$$

where to be singular $\ker T$ is non-empty. Because a zero vector cannot be a basis we must have $AX = 0$. Therefore it implies $\ker A$ must also be non-emtpy. That directly implies that then A must be non-invertible and hence $\det A = 0$ You can clearly see the bidirectional nature of the proof why A must be invertible for T to be invertible.

Proposition 110

Let A be the matrix of a linear operator T with respect to a basis $\mathbf{B}$

(a) Suppose that new basis $\mathbf{B}'$ is decribed by $\mathbf{B}' = \mathbf{B}P$. The matrix that represents T with respect to this basis is $A' = P^{-1}AP$

(b) The matrices A' that represent the operator T for different bases are the matrices of the form $A' = P^{-1}AP$ where P can be any invertible matrix

Proof. we have every

$$T(v_i) = v_1 a_{1i} + \ldots v_n a_{ni}$$

Therefore

$$T(\mathbf{B}) = (T(v_1), \ldots, T(v_n)) = \mathbf{B}A$$

Then if $v = \mathbf{B}X$ we have

$$T(v) = T(\mathbf{B}X) = x_1 T(v_1) + \ldots + x_n T(v_n) = T(\mathbf{B})X$$

So combining with the above we have

$$T(\mathbf{B})X = \mathbf{B}AX$$

Let $T(v) = \mathbf{B}Y$ where Y is the coordinate vector of its image with respect to basis $\mathbf{B}$. Then

$$T(v) = \mathbf{B}Y = \mathbf{B}AX$$

which implies $Y = AX$. Also we must have

$$\mathbf{B}'X' = \mathbf{B}X$$
$$\mathbf{B}'Y' = \mathbf{B}Y$$

and subbing $\mathbf{B}' = \mathbf{B}P$ into the above we have

$$\mathbf{B}PX' = \mathbf{B}X$$
$$\mathbf{B}PY' = \mathbf{B}Y$$

Hence, $PX' = X$ and $PY' = Y$ so subbing into the below we have

$$Y = AX \Rightarrow PY' = APX' \Rightarrow Y' = P^{-1}APX'$$

So for $Y' = A'X'$ which is the analogous of $Y = AX$ previously we must have by comparison $A' = P^{-1}AP$. And clearly the base change matrix P must be invertible for $\mathbf{B}' = \mathbf{B}P$ to be valid so (b) follows.

Definition 111 (Similar)
Square matrix A is **similar** to another matrix A' if

$$A' = P^{-1}AP$$

for some invertible matrix P. We say matrix A' is obtained from A by **conjugating** by P^{-1}

This definition is cleary motivated by 110

4.1 Eigenvectors

Definition 112 (Invariant subspaces)
A subspace W of V is **invariant**, or more precisely **T-invariant** if it is carried to itself by the operator that is

$$TW \subset W$$

Fact 113

Suppose W has a basis $(w_1, \ldots, w_k)$ and we extend it to the basis of V of dimension n by

$$\mathbf{B} = (w_1, \ldots, w_k, v_1, \ldots, v_{n-k})$$

Then the matrix of the coordinate vectors of the image vectors is given by

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

Where A is a $k \times k$ matrix. This is because when we write $T(w_j)$ in terms of basis $\mathbf{B}$, all the coefficients of $v_1, \ldots, v_{n-k}$ must be zero since W is T-invariant

Example 114

Similar if $V = W_1 \oplus W_2$ where W_1, W_2 are T-invariant we have

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

Where A_1, A_2 are square matrices corresponding to the restriction of T to W_1 and W_2 respectively.

Fact 115

To check if whether a vector v is an eigenvector that is

$$T(v) = \lambda v$$

Simply consider that this is also equivalent to

$$T(v) = \mathbf{B}(AX) = \mathbf{B}(\lambda X) = \lambda(\mathbf{B}X) = \lambda v$$

so we simply need to check if its corresponding coordinate vector X is a multiple AX. Therefore we have:

$$AX = \lambda X$$

It follows from the idea the coordinate vectors are with respect to the same base $\mathbf{B}$ anyway since we are considering *linear operators* which maps a space to itself.

Proposition 116

Similar matrices which are of the form $A' = P^{-1}AP$ where P is some abitruary invertible matrix have the same eigenvalues.

Proof. Knowing that P is invertible here, we can conjugate both sides by:

$$A' = P^{-1}AP$$
$$PA' = AP$$
$$PA'P^{-1} = A$$

Now subbing $PA'P^{-1} = A$ as well as the fact that $X = PX'$ we have:

$$\begin{aligned} AX &= \lambda X \\ (PA'P^{-1})(PX)' &= \lambda(PX') \\ PA'X' &= P\lambda X' \end{aligned}$$

Then taking P^{-1} on both sides we see that

$$A'X' = \lambda X'$$

This is to be expected as $A' = P^{-1}AP$ is simply a change of how v is represented. The same linear transformation will not change the properties of subspaces.

Proposition 117

Let T be a linear operator on a vector space V. The matrix of T with respect to a basis $\mathbf{B}$ is diagonal if and only if each of the basis vectors v_i is an eigenvector.

Proof. The proof is trivial. Just consider how the values of the matrix will look like if we have

$$T(v_j) = \lambda_j v_j$$

□

Proposition 118

Let $v_1, \ldots, v_r$ be eigenvectors of linear operator T with distinct eigenvalues $\lambda_1, \ldots, \lambda_r$. Then $(v_1, \ldots, v_r)$ is independent

Proof. Refer to `Axler Linear Algebra Done Right`

4.2 The Characteristic polynomial

Proposition 119

If T is a singular operator then T has an eigenvalue equal to zero.

Proof. Singular implies null space contains non-zero vector. Then

$$Tv = 0$$

for some non-zero v. But because we know that the nullspace is an invariant subspace, recall `Axler`, the fact that it contains a non-zero vector shows that it must be spanned by a non-zero vector. That implies there exists

$$Tv = \lambda v$$

where $\lambda = 0$ and v is a non-zero eigenvector □

Proposition 120

The following propositions motivate our formulation of the characteristic polynomial below:

1. A nonzero vector v is an eigenvector with eigenvalue λ if and only if it is in the kernel of $\lambda I - T$.
2. $\lambda I - T$ is singular if and only $\det(\lambda I - A) = 0$

Proof. (1) is trivial.For (2) recall 109. The exact same logic applies.

Definition 121

The **characteristic polynomial** of a linear operator T is the polynomial

$$p(t) = \det(tI - A)$$

Proposition 122

Let A be an upper or lower triangular $n \times n$ matrix with diagonnal entries $a_{11}, \ldots, a_{nn}$. The characteristic polynomial of A is

$$(t - a_{11}) \ldots (t - a_{nn})$$

where the diagonal entries are the eigenvalues

Proof. Consider when A is in upper triangular form. Then we have

$$(\lambda I - A) = \begin{bmatrix} \lambda - a_{11} & \ldots & * \\ & \ddots & \vdots \\ 0 & & \lambda - a_{nn} \end{bmatrix}$$

Now $p(\lambda) = \det(\lambda I - A) = \prod_i (\lambda - a_{ii})$ recall 37. We see that the diagonal elements are indeed of the roots of the characteristic equation. Then applying knowledge on **polynomial theory**(see later), we can express $p(\lambda)$ in sum of roots(given by the trace clearly) and product of roots(determinant) form. That is

$$p(\lambda) = \lambda^n - (\text{trace} A)\lambda^{n-1} + (\text{intermediate powers of } \lambda) + (-1)^n(\det A)$$

Proposition 123

Let T be a linear operator on a finite dimensional *complex* vector space V. Then there is a basis $\mathbf{B}$ of V such that the matrix of T with respect to that basis is *upper triangular*

Proof. Recall from **fundamental theorem of algebra**(see later) a complex polynomial has at least one root and thus 1 eigenvalue. Recall 113, then there exists a $A' = P^{-1}AP$ such that

$$A_{n-1} = \begin{bmatrix} \lambda_n & * \\ 0 & [D] \end{bmatrix}$$

where D is a $(n-1) \times (n-1)$ matrix. Let our induction hypoethesis be that for all subsequent subsquare matrices D

we can do the same. Let

$$Q_j = \begin{bmatrix} \ddots & & \\ & 1 & * \\ & 0 & [Q] \end{bmatrix}$$

where the apart from the bottom right $j \times j$ square invertible matrix $[Q]$(recall 111) , the rest is the indentity matrix. Then Q_j^{-1} is simply

$$Q_j^{-1} = \begin{bmatrix} \ddots & & \\ & 1 & * \\ & 0 & [Q^{-1}] \end{bmatrix}$$

Recall 2, we can then let

$$A_{j-1} = Q_j^{-1} A_j Q_j = \begin{bmatrix} \ddots & & \\ & 1 & * \\ & 0 & [Q^{-1}] \end{bmatrix} \begin{bmatrix} \ddots & & \\ & \lambda_j & * \\ & 0 & [D] \end{bmatrix} \begin{bmatrix} \ddots & & \\ & 1 & * \\ & 0 & [Q] \end{bmatrix}$$

$$= \begin{bmatrix} \ddots & & \\ & \lambda_j & * \\ & 0 & [Q^{-1}DQ] \end{bmatrix}$$

Which exists again because there must 1 eigenvalue for the matrix in which our linear operator is restricted to

Proposition 124

Let T be a linear operator on F^n with a corresponding transformation matrix A.

(a) If $\mathbf{B} = (v_1, \ldots, v_n)$ is a basis of *eigenvectors* of T. Then $\Lambda = P^{-1}AP$ is diagonal where $P = [\mathbf{B}]$

(b) Suppose we have the same A from (a).If we have a polynomial function on matrix A, that is

$$f(A) = a_o I + a_1 A + \ldots a_n A^n$$

then we also have

$$f(A) = P f(\Lambda) P^{-1}$$

Proof. By assumption we have $T(v) = IAX$ where essentially A is a transformation matrix of T with respect to the standard basis. Thus recall by 117, if let our matrix of transformation be defined with respect to the basis of eigenvectors instead we would be able to get a diagonal matrix. Consider where we recall that we let our basis change matrix P be any invertible matrix of apropriate size.

$$\mathbf{B} = IP \quad \Rightarrow \quad \mathbf{B}P^{-1} = I$$

It is from this result that $P = \mathbf{B}$. Hence (a) follows.

For (b) the first step is to naturally try to express powers of A with respect to our diagonal matrices. Now consider that

$$\Lambda = P^{-1}AP \quad \Rightarrow \quad A = P\Lambda P^{-1}$$

therefore we have

$$A^k = (P\Lambda P^{-1})^k = (P\Lambda P^{-1})(P\Lambda P^{-1}) \dots \text{k times}$$

but notice "$\dots P^{-1})(P \dots$" between brackets cancel to get the identity vector.Therefore we have

$$A^k = P\Lambda^k P^{-1}$$

Hence using linearity of linear transformations we have

$$\begin{aligned} f(A) &= a_o I + a_1 A + \dots a_n A^n \\ &= a_o I + a_1 P\Lambda P^{-1} + \dots a_n P\Lambda^n P^{-1} \\ &= P f(\Lambda) P^{-1} \end{aligned}$$

□

Remark 125. *As we know from 39, the powers of such diagonal matrices will give:*

$$\Lambda^k = \begin{bmatrix} \lambda_1 & 0 & \\ 0 & \lambda_2 & \\ & & \ddots \end{bmatrix}^k = \begin{bmatrix} \lambda_1^k & 0 & \\ 0 & \lambda_2^k & \\ & & \ddots \end{bmatrix}$$

4.3 Jordan Form

Definition 126 (Generalized Eigenvectors)

A **generalized eigenvector** with eigenvalue λ of a linear operator T is a non-zero vector x such that

$$(T - \lambda)^k x \begin{cases} = 0 & k \geq d \\ \neq 0 & k < d \end{cases}$$

for some $k > 0$. d is known as the **exponent**

Proposition 127

Let x be a generalized eigenvector of T with eigenvalue λ and exponent d for $j \geq 0$, let $u_j = (T - \lambda)^j x$. Let $\mathbf{B} = (u_0, \dots, u_{d-1})$ and let $X = \text{Span}(\mathbf{B})$. Then X is a T-invariant subspace and $\mathbf{B}$ is a basis of X.

Proof. We note that We know that $u_j = 0$ for $j \geq d$ and $u_j \neq 0$ otherwise.

$$u_{d-1+k} = (T - 1)^{d-1+k} x \tag{1}$$

$$(T - \lambda) u_{j-1+k} = (T - \lambda)^{d+k} x \tag{2}$$

$$T u_{d-1+k} = \lambda u_{d-1+k} + (T - \lambda)^{d+k} x \tag{3}$$

$$= \lambda u_{d-1+k} + u_{d+k} \tag{4}$$

Therefore the cases for $k < 0, k = 0, k > 0$ correspond to the 3 cases below in that order.

$$Tu_j = \begin{cases} \lambda u_j + u_{j+1} & j < d-1 \\ \lambda u_j & j = d-1 \\ 0 & j > d-1 \end{cases} \tag{5}$$

$$\tag{6}$$

Therefore we have shown that X is invariant under T. Now the prove that $\mathbf{B}$ is indeed a basis, since we already know that is spans X by assumption, it remains to show that it is independent. By 128, we see that every non-trivial(that is $c_j, \ldots, c_{d-1}$ all not equal zero) linear combination of vectors in $\mathbf{B}$ is a generalized eigenvector which is non-zero. Hence the contrapositive of this is that all zero vectors correspond to only trivial linear combination of vectors in $\mathbf{B}$. In other words, this is a direct statement that $\mathbf{B}$ is indeed independent as desired.

Lemma 128

With u_j above, a linear combination $y = c_j u_j + \ldots + c_{d-1} u_{d-1}$ with $j \leq d-1$ and $c_j \neq 0$, then y is a generalized eigenvector with eigenvalue λ and exponent $d - j$

Proof. Consider that y is a sum of $d-j$ terms since $j \leq d-1$. Again note that $u_j = 0$ for $j \geq d$ and $u_j \neq 0$ otherwise.

$$y = c_j(T-\lambda)^j x + c_{j+1}(T-\lambda)^{j+1} x + \ldots + c_{d-1}(T-\lambda)^{d-1} x \tag{1}$$

$$(T-\lambda)^{d-1-j} y = (T-\lambda)^{d-1-j}\left(c_j(T-\lambda)^j x + c_{j+1}(T-\lambda)^{j+1} x + \ldots + c_{d-1}(T-\lambda)^{d-1} x\right) \tag{2}$$

$$= (T-\lambda)^{d-1}\left(c_j x + c_{j+1}(T-\lambda)^1 x + \ldots c_{d-1}(T-\lambda)^{d-1-j} x\right) \tag{3}$$

$$= (T-\lambda)^{d-1} c_j x + 0 \tag{4}$$

$$(T-\lambda)^{d-j} y = (T-\lambda)^{d}\left(c_j x + c_{j+1}(T-\lambda)^1 x + \ldots c_{d-1}(T-\lambda)^{d-1-j} x\right) \tag{5}$$

$$= 0 \tag{6}$$

Therefore we have shown that y is a generalized vector and $d-j$ is the exponent because anything smaller is non-zero.

Definition 129

A linear operator on T on a vector space V is **nilpotent** if for some positive integer r the operator T^r is zero

That is for any element in $v \in V$, there must exist r such that $T^r v = 0$. In particular we have shown above that linear opeator is nilpotent on the space generated by generalized vectors like above which we refer to **Jordan generators**. Moreover it is also obvious the linear operator is nilpotent on space spanned by Jordan generators as well.

Definition 130 (Jordan block)

A **Jordan block** is any of the following matrices for a fixed λ

$$\begin{bmatrix} \lambda \end{bmatrix}; \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}; \begin{bmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{bmatrix} \ldots$$

Essentially square matrices with λ on the diagonal, 1 right below and zero everywhere else

Theorem 131 (Jordan Decomposition Theorem)

Any complex $n \times n$ matrix is similar to a matrix J made up of diagonal **jordan blocks** - that is it has the **Jordan form**

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_\ell \end{bmatrix}$$

where $J_i = J_{\lambda_i}$ for some λ_i. Each J_i is a $d_i \times d_i$ matrix with $\sum_i d_i = n$. The characteristic polynomial is given by:

$$p(t) = (t - \lambda_1)^{d_1}(t - \lambda_2)^{d_2} \dots (t - \lambda_\ell)^{d_\ell}$$

5 Applications of Linear Operators

Definition 132 (Dot Product)

The **dot product** of *column vectors* $X = (x_1, \dots, x_n)^t$ and $Y = (y_1, \dots, y_n)^t$ in $\mathbb{R}^n$ is defined to be

$$(X \cdot Y) = x_1 y_1 + \dots x_n y_n$$

Writing as a matrix product we have:

$$(X \cdot Y) = X^t Y$$

Definition 133 (Orthgonal Vector)

A vector X is **orthogonal** to another vector Y(written as $X \perp Y$ if and only if $X^T Y = 0$)

Note that orthogonal matrix is not the same as orthogonal vectors in a sense the column vectors of an orthogonal matrix must not be orthogonal but orthonormal.

Example 134

The matrix $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ is not orthogonal since

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}^T = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \neq I$$

But its column vectors are orthogonal since

$$[2\ 0][0\ 3]^T = [0\ 0]^T$$

Definition 135 (Orthonormal Basis)

An **orthonormal basis** $\mathbf{B} = (v_1, \ldots, v_n)$ of $\mathbb{R}^n$ is a basis of orthogonal **unit vectors**(that is $||v_i|| = 1$)

$$(v_i \cdot v_j) = \delta_{ij}$$

where δ_{ij}, the **Kronecker delta** is 1 if $i = j$ and 0 otherwise.

Definition 136 (Orthogonal Matrix)

A real $n \times n$ matrix A is **orthogonal** if $A^t A = I$ which is say A is invertible and its inverse is A^t

It is clear that A is orthogonal if and only if its columns form an orthonormal basis of $\mathbb{R}^n$

Corollary 137

The determinant of an orthogonal matrix is ± 1

Proof. Since $\det^2 M = \det M^T M = \det I = 1$.

Fact 138

Note that we call orthogonal matrices with $\det M = -1$ a **reflection** while $\det M = 1$ a **rotation**(or proper orthgonal). $|\det M| = 1$ implies preservation of lenghts/volume. But the change in sign implies a change in orientation.

Definition 139

An orthogonal operator T on $\mathbb{R}^n$ is a linera operator that preserves the dot product that is

$$(TX \cdot TY) = (X \cdot Y)$$

6 Bilinear Forms

Definition 140

A **bilinear form** on V is a real-valued function of two vector variables that is: a map $V \times V \to \mathbb{R}$. We denote the real number the form returns as $\langle v, w \rangle$ and it must, as the name implies, be **linear** in each variable that is:

$$\langle rv_1, w_1 \rangle = r\langle v_1, w_1 \rangle \text{ and } \langle v_1 + v_2, w_2 \rangle = \langle v_1, w_1 \rangle + \langle v_2, w_1 \rangle$$

$$\langle w_1, rv_1 \rangle = r\langle w_1, v_1 \rangle \text{ and } \langle w_1, v_1 + v_2, w_2 \rangle = \langle w_1, v_1 \rangle + \langle w_1, v_2 \rangle$$

for all $v_i, w_i \in V$ and all real numbers r

Fact 141

Like how we defined the billinear form on V, the form on $\mathbb{R}^n$ is defined by

$$\langle X, Y\rangle = X^t AY$$

where A is $n \times n$ matrix and X, Y are coordinate/column vectors in $\mathbb{R}^n$ Note that $A = I$ corresponds to the case of the dot product. One always assumes this form unless specified otherwise

Proposition 142

Let $\langle , \rangle$ be a bilinear form on a vector space V with basis $\mathbf{B} = (v_1, \ldots, v_n)$. Let X and Y are the coordinate vectors(which are column vectors by convention) of the vectors $v, w \in V$ respectively. If we define

$$a_{ij} = \langle v_i, v_j\rangle$$

where $a_{i\,j}$ is simply the real number our form returns then we can express our form as:

$$\langle v, w\rangle = X^t AY = \langle X, Y\rangle$$

where A is known as the **matrix of the form**

Proof. Let $v = \mathbf{B}X$ and $w = \mathbf{B}Y$. Therefore by properties of bilinear form we have

$$\langle v, w\rangle = \langle \sum_i v_i x_i, \sum_j v_j y_j\rangle = \sum_{i,j} x_i \langle v_i, v_j\rangle y_j = \sum_{i,j} x_i a_{ij} y_j$$

To express this sum of real numbers x_i, a_{ij}, y_j in terms of matrix product, recall 93 we can do:

$$\sum_{i,j} x_i a_{ij} y_j = X^t AY$$

where $A = (a_{ij})$

Definition 143 (Symmetric)

A bilinear form is **symmetric** if

$$\langle v, w\rangle = \langle w, v\rangle$$

and **skew-symmetric** if:

$$\langle v, w\rangle = -\langle w, v\rangle$$

for all v and w in V

Proposition 144

Let A be an $n \times n$ matrix. The form $X^t AY$ is symmetric that is

$$\langle X, Y\rangle = X^t AY = Y^t AX = \langle Y, X\rangle$$

for all X and Y if and only if the matrix A is symmetric : $A^t = A$

Proof. Assume A is symmetric

$$X^tAY = (X^tAY)^t = Y^tA^tX = Y^tAX$$

For proof from the other direction consider assume

$$X^TAY = Y^TAX$$

then taking transpose on both sides we have

$$Y^TA^TX = X^TA^TY$$

but recall that the a bilinea form results in a scalar so we must have that

$$X^TAY = X^TAY \quad \text{and} \quad Y^TA^TX = Y^TAX$$

too which proves the proposition as desired.

Proposition 145

Let A and A' be the matrix of a bilinear form with respect to basis **B** and **B**' respectively. The matrices that represent the same form with respect to different bases are

$$A' = P^tAP$$

where P can be any invertible matrix and that $\mathbf{B}' = \mathbf{B}P$

Proof. Firstly recall for P to be a base change matrix, it has to be invertible and that $X = PX'$. Then we have

$$\langle X, Y\rangle = X^tAY = (PX')^tA(PY') = (X')^t(P^tAP)(Y')$$

where recall that the second equality follows from the property of transposes

Remark 146. *If A is symmetric then so is A'. Consider*

$$(A')^t = (P^tAP)^t = P^tA^tP = P^tAP = A'$$

6.1 symmetric forms

Definition 147 (definite)

A symmetric form on V is **positive definite** if $\langle v, v\rangle > 0$ and **positive semi-definite** if $\langle v, v\rangle \geq 0$ for all non-zero vectors v. Negative definite and negative semi-definite is defined analogously

Equivalently a matrix M(more precisely the matrix of the form) is positive definite if $x^TMx > 0$ for all x. For positive definite we have $x^TMx \geq 0$ instead. Where as usual x is the coefficient column vector of v,

Corollary 148

The matrices A that represent the form $\langle X, Y\rangle = X^tAY$ equivalent to dot product are those that can be written as a product P^tP for some invertible matrix P

Proof. come back later

6.2 hermitian forms

Definition 149 (Hermitian Form)

A **hermitian form** on a complex vector space V is a map

$$V \times V \to \mathbb{C}$$

and again denoted by $\langle v, w \rangle$ that is conjugate linear in the first variable, linear in the second variable and **hermitian symmetric**:

$$\langle cv_1, w_1 \rangle = \overline{c}\langle v_1, w_1 \rangle \text{ and } \langle v_1 + v_2, w_2 \rangle = \langle v_1, w_1 \rangle + \langle v_2, w_1 \rangle$$

$$\langle w_1, rv_1 \rangle = r\langle w_1, v_1 \rangle \text{ and } \langle w_1, v_1 + v_2 \rangle = \langle w_1, v_1 \rangle + \langle w_1, v_2 \rangle$$

$$\langle w_1, v_1 \rangle = \overline{\langle v_1, w_1 \rangle}$$

In the case where we have the same first and second variable, due to hermitian symmetry we must have

$$\langle v, v \rangle = \overline{\langle v, v \rangle}$$

hence $\langle v, v \rangle$ must be a real number for all vectors v. If not we might get some $a + bi \neq a - bi$

Definition 150 (Standard Hermitian Form)

The **standard hermitian form** on $\mathbb{C}^n$ is

$$\langle X, Y \rangle = X^*Y = \overline{x}_1 y_1 + \ldots + \overline{x}_n y_n$$

where X^* stands for the conjugate transpose that is:

$$X^* = (\overline{x}_1, \ldots, \overline{x}_n) \text{ and } X = (x_1, \ldots, x_n)^t$$

From the defintion we also clearly see

$$\langle X, X \rangle$$

is a *positive* real number for any non-zero X That means it is not zero, since zero is neither positive or negative if you recall.

Definition 151 (adjoint)

The **adjoint** of A^* of a complex matrix $A = (a_{ij})$ is the complex conjugate of the transpose matrix A^t. That is the i,j entry of A^* equals $\overline{a}_{ji}$. The square matrix A is **hermitian**(or **self-adjoint**) if

$$A^* = A$$

Fact 152 (Rules for adjoint matrices)

Here are some rules for computing with adjonit matrices

$$(cA)^* = \overline{c}A^*, \quad (A+B)^* = A^* + B^*, \quad (AB)^* = B^*A^*, \quad A^{**} = A$$

In fact they follow directly from linearity and transpose properties from before. The only new thing to note is the 2nd and 3rd relations, which show that the conjugate is distributive. Consider individual pairs elements during matrix multiplication

$$\overline{(e^{i\theta})}\ \overline{(e^{i\theta})} = e^{-i\theta}e^{-i\theta} = e^{-i2\theta} = \overline{(e^{i\theta}e^{i\theta})}$$

Proposition 153

Let A be the matrix of a hermition form $\langle , \rangle$ on a complex vector space V with respect to basis $\mathbf{B}$. If X, Y are coodinate vectors of the vectorss v, w respectively. If we define

$$a_{ij} = \langle v_i, v_j \rangle$$

where $a_{i,j}$ is simply the complex number our form returns then we can express our form as:

$$\langle v, w \rangle = X^*AY = \langle X, Y \rangle$$

where A is a **hermitian matrix**

Proof. It is an analogous to 142. Let $v = \mathbf{B}X$ and $w = \mathbf{B}Y$. Since conjugate linear in 1st variable we have:

$$\langle v, w \rangle = \langle \sum_i v_i x_i, \sum_j v_j y_j \rangle = \sum_{i,j} \overline{x_i} \langle v_i, v_j \rangle y_j = \sum_{i,j} \overline{x_i} a_{ij} y_j$$

To express this sum these complex numbers x_i, a_{ij}, y_j in terms of matrix product, recall 93 we can do:

$$\sum_{i,j} \overline{x_i} a_{ij} y_j = X^*AY$$

where $A = (a_{ij})$. Now A is a hermitian matrix because our form needs to satisfy hermitian symmetry. That is:

$$\langle v, w \rangle = X^*AY = (X^*AY)^* = Y^*A^*X = Y^*AX = \overline{\langle w, v \rangle}$$

Remark 154. *We can see that the case when $A = I$ will be the standard hermitian form. This is analagous to the dot product case we defined for bilinear forms previously*

Proposition 155

Let A and A' be the matrix of a hermitian form with respect to basis $\mathbf{B}$ and $\mathbf{B}$' respectively. The matrices that represent the same form with respect to different bases are

$$A' = P^*AP$$

where P can be any invertible matrix and that $\mathbf{B}' = \mathbf{B}P$

Proof. Firstly recall for P to be a base change matrix, it has to be invertible and that $X = PX'$. Then we have

$$\langle X, Y\rangle = X^*AY = (PX')^*A(PY') = (X')^*(P^*AP)(Y')$$

where recall that the second equality follows from the property of adjoints

Theorem 156
The eigenvalues, the trace and the determinant of a hermitian matrix A are real numbers

Proof. Recall from the chapter on **characteristic polynomials** that the trace and determinant are represented in terms of eigenvalues. Hence we just need to prove that eigenvalues are real. Let us make use of the fact that $\langle X, X\rangle$ is a positive real number for *any* non-zero X and let X be an eigenvector of A with eigenvalue λ.

$$X^*AX = X^*(AX) = X^*(\lambda X) = \lambda X * X \in \mathbb{R}^+$$

. Like how we showed that $\langle X, X\rangle$ was positive we try to obtain the form $\lambda = \overline{\lambda}$. We see that $(\lambda X)^* = \overline{\lambda}X^*$. So we repeat the steps like we did above, noting that $A^* = A$ since it is a hermitian matrix.

$$X^*AX = (X^*A)X = (X^*A^*)X = (AX)^* = (\lambda X)^*X = \overline{\lambda}X^*X \in \mathbb{R}^+$$

Since clearly X^*X is non-zero too then we have

$$\lambda = \overline{\lambda}$$

as desired. λ cant be complex or the conjugate will be different.

Corollary 157
The eigenvalues of a real symmetric matrix are real numbers

Proof. a real symmetric matrix is Hermitian. So the above follows

Definition 158 (Unitary)
A matrix P such that

$$P^*P = I$$

or $P^* = P^{-1}$

In particular orthogonal matrices(which include both rotation(proper orthogonal) and reflections) M are unitary since by definition 138 they satisfy

$$M^TM = MM^T = I$$

which implies $M^T = M^{-1}$

6.3 orthogonality

Definition 159

Two vectors v and w are **orthogonal** written($v \perp w$) if

$$\langle v, w \rangle = 0$$

note that $v \perp w$ if and only if $w \perp v$

Definition 160 (Orthogonal Space)

The **orthogonal space** to a subspace W of V, often denoted by $W^{\perp}$ is the subspace of vectors v that are orthogonal to every vector in W or symbolically such that $v \perp W$

$$W^{\perp} = \{v \in V | \langle v, w \rangle = 0, \forall w \in W\}$$

Definition 161 (Orthogonal Basis)

An **orthogonal basis** $\mathbf{B} = (v_1, \ldots, v_n)$ of V is a basis whose vectors are mutually orthogonal that is

$$\langle v_i, v_j \rangle = 0$$

for all $i \neq j$ If additionally

$$\langle v_i, v_j \rangle = 1$$

for all $i = j$ then we say such a basis is an **orthonormal basis**

Definition 162 (Null Vector)

A **null vector** of V is a vector orthogonal to ever vector in V and the **nullspace** N of the form is the set of null vectors. The nullspace can be described as the **orthogonal space** to the *whole* space V

$$N = \{v | v \perp V\} = V^{\perp}$$

Definition 163 (Nondegenerate)

The form on V is **nondegenerate** if its nullspace is the zero space $\{0\}$. That is to say for every non-zero $v \in V$ there exists a $v' \in V$ such that $\langle v, v' \rangle \neq 0$(or equivalently there are no non zero vectors in V that are orthogonal to all other vectors in V)

Fact 164

The form on V is nondegenerate on a subspace W if its restriction to W is a non-degeneate form. That is to say for every non-zero $w \in W$ there exists a $w' \in W$ such that $\langle w, w' \rangle \neq 0$ or equivalently

$$W \cap W^T = \{0\}$$

Proposition 165

Let $\langle , \rangle$ be a nondegenerate symmetric or Hermitian form on V, and let v, v' be vectors in V. If $\langle v, w \rangle = \langle v', w \rangle$ for all vectors $w \in V$ then $v = v'$

Proof. Consider

$$\langle v, w\rangle = \langle v', w\rangle$$
$$\langle v - v', w\rangle = 0$$

since $(v - v') \perp w$ for all $w \in W$ and that being nondegenerate, only the zero vector satisfy this. Hence

$$v - v' = 0$$
$$v = v'$$

Proposition 166

Let $\langle , \rangle$ be a symmetric form on a real vector space or a hermitian form on a complex vector space, and let A be its matrix with respect to a basis.

(a) A vector v is a null vector if and only if its coordinate vector Y solves the homogenous equation $AY = 0$

(b) The form is nondegenerate if and only if the matrix A is invertible

Proof. (a) v with respect to a basis has a form that corresponds to X^*AY where Y is the coordinate vector of v.If $AY = 0$ then $X^*AY = 0$ for all X so Y is the the null vector. Conversely if Y is the null vector, $X^*AY = 0$ for all X.Let the coordinate vector X be the unit matrix e_i. This is certainly possible because recall that the elements of coordinate vectors X, Y are simply the scalars muliplying the basis vectors in the linear combination. These scalars can take any value.e_iAY picks out the ith coordinate of AY. Thus if there exists any $e_i^*AY \neq 0$ then $AY \neq 0$. So it must be that $AY = 0$. As for (b) if invertible then the only possible Y that solves $AY = 0$ is when Y is the zero vector. Therefore from (a) it follows that the zero vector is the only vector in the nullspace. That is for all X, the form $X^*AY = 0$ only if $Y = 0$

Theorem 167

Let $\langle , \rangle$ be a symmetric form on a real vector space V or a hermitian form on a complex vector space V and let W be an *arbitrary* subspace of V

(a) the form is nondegenerate on W if and only if V is the direct sum $W \oplus W^\perp$

(b) the form is nondegenerate on V and on W then it is nondegenerate on $W^\perp$

Proof. For (a) From the forward direction $W \oplus W^\perp$ implies $V = W + W^\perp$ and $W \cap W^\perp = \{0\}$ thus the conclusion follows. From the other direction suppose we have a matrix of the form with respect to the basis $\mathbf{B} = (w_1, \ldots, w_k, v_1, \ldots, v_{n-k})$ where $(w_1, \ldots, w_k)$ is the basis of W and the rest is the extension of the basis to $\mathbf{B}$. So we have a matrix in the form

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where A is a $k \times k$ matrix while D is a $(k-n) \times (k-n)$ matrix. We simply only need to show that there exists a basis in which B is zero block since it represents $\langle w_i, v_j\rangle$ for $i : 1 \ldots k$ and $j : 1 \ldots n-k$ And because M is symmetric/hermitian, so too will C automatically. Recall 145 that a change of base doesn't not affect the symmetry property of the matrix

of form. Now consider $\mathbf{B}' = \mathbf{B}P$ and

$$P = \begin{bmatrix} I & Q \\ C & I \end{bmatrix}$$

which is a valid change of basis matrix because it is inertible. Determinant of this upper triangular matrix is equal one which is non-zero. Hence we have:

$$M' = P^* M P = \begin{bmatrix} I & 0 \\ Q^* & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & AQ+B \\ * & * \end{bmatrix}$$

Then we can get the upper right block to be zero when $Q = -A^{-1}B$. This exists because A is invertible since the form on W is non-degenerate. Again we are guaranteed the bottom left block will also go to zero with this, so we don't have to care about them. For (b) Since W non-degenerate we have the matrix of the form

$$\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$$

as proven above and that A is invertible. Because V is invertible, this whole matrix must be invertible. Therefore D must be invertible as well which implies the form on $W^\perp$ is nondegenerate.

Lemma 168

If a symmetric or hermitian form is not identically zero(basically every $\langle v, w\rangle \neq 0$ or equivalently the matrix of form is the zero matrix), there is a vector v in V such that $\langle v, v\rangle \neq 0$

Proof. Because not identically zero there exists non zero $\langle x, y\rangle$. Since symmetric/hermition we have $\langle x, y\rangle = \langle y, x\rangle$. If hermition we can just replace y with a cy where the non-zero scalar $c \in \mathbb{C}$ that makes $\langle x, cy\rangle = c\langle x, y\rangle \in \mathbb{R}$. Now we attempt to find 3 possiblities of $\langle v, v\rangle$ using x, y. It turns out we can relate them like so:

$$\langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle$$

Because $2\langle x, y\rangle \neq 0$, we must have at least one of the other 3 terms be non-zero

Theorem 169

Let $\langle . \rangle$ be a symmetric or hermitian form on V. Then there exists an orthogonal basis for V.

Proof. If the form on the space is identically zero then every basis is orthogonal since by definition $\langle v_i, v_j\rangle = 0$ if $i \neq j$ is cleary satisfied. The defintion doesn't say anything about when $i = j$. Otherwise from the above lemma, we can select a v_1 such that $\langle v_1, v_1\rangle \neq 0$. Therefore we can let this be a 1×1 cell in our matrix of form. Since it is obviously invertible it is a non-degenerate form restricted to a space say W. Therefore by 167 we must have $V = W \oplus W^T$ So our matrix of the form looks like

$$\begin{bmatrix} W & 0 \\ 0 & W^T \end{bmatrix}$$

Now do induction on W^T(repeat the above steps for it). Then in the end we will get some diagonal matrix which clearly reflects an orthogonal basis(and if we normalize them we get orthonormal basis)

Remark 170. *Intuitively can we also then conclude that all symmetric/hermitian matrices are diagonalizable?(well also note that the zero matrix is a diagonal matrix) I mean any symmetric/hermitian matrix could correspond to the*

matrix of some symmetric/hermitian form(recall 144) Answer is yes and in fact this is exactly how it is reasoned. For reference see this in spectral theorem below, which assuming an orthonormal basis, connects operators to matrices of forms and then finally to matrices in general.

Example 171

positive definite does not imply symmetric. Consider

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

To check whether this matrix is positive definite, we need to examine the quadratic form $\mathbf{x}^\top A\mathbf{x}$:

Let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$:

$$\mathbf{x}^\top A\mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 2x_1x_2 + x_2^2.$$

This expression can be rewritten as:

$$\mathbf{x}^\top A\mathbf{x} = (x_1 + x_2)^2.$$

The expression $(x_1 + x_2)^2 > 0$ for all non-zero vectors $\mathbf{x}$ means that the matrix A is indeed positive definite. However, A is not symmetric because $A \neq A^\top$:

$$A^\top = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Theorem 172 (Projection formula)

Let a symmetric/hermitian form on V be nondegenerate on a subspace W of V with an orthogonal basis $(w_1, \ldots, w_n)$. Consider a the **orthogonal projection** which is a map $\pi : V \to W$ which is defined by

$$\pi(v) = w_1c_1 + \ldots w_kc_k$$

for $v \in V$ and $c_i = \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle}$

Proof. It is sufficient to show that $\pi(w_j) = w_j$ and $\pi(u) = 0$ for all $u \in W^T$. If u is orthgonal to W then clearly $\langle u, w_i \rangle$ for all i so $\pi(u) = 0$. On the other hand if we plug w_j in the formula above we can easily get $c_i = w_j$.

Corollary 173

Let $\langle , \rangle$ be a symmetric/hermitian form on a vector space V. Then there is an orthogonal basis $\mathbf{B} = (v_1, \ldots, v_n)$ such that for each i,

$$\langle v_i, v_i \rangle = 1 \text{ or } -1 \text{ or } 0$$

Proof. Consider any non-zero real number c we have

$$\langle cv, cv \rangle = c^2 \langle v, v \rangle$$

Since by property of hermitian/symmetry, $\langle v_i, v_i \rangle \in \mathbb{R}$ which could be positive, zero or negative we can always scale the basis the basis vector by some c such that we get a magnitude of 1.

Corollary 174

If A is the matrix of a symmetric/hermitian form and positive definite then the matrix of the represents the dot product with respect to some basis of $\mathbb{R}^n$.

Proof. Consider from the previous theorems there exists a orthogonal basis and that we can scale it such that the magnitude of all non-zero elements are 1. Because positive definite $\langle cv, cv\rangle$ can only be 1. If we arrange our basis vectors suitably we can have our matrix of the form to be the indentity matrix and therefore we have the dot product.

Fact 175 (the angle)

We can relate familiar ideas of length and angle in $\mathbb{R}^2$ learnt in high school to our knowledge on forms. The **lenghth** $|v|$ of a vector v is defined by

$$|v|^2 = \langle v, v\rangle = x^t x$$

where x is the coordinate vector of v. The the **angle** is defined by **law of cosines** which states:

$$(x \cdot y) = |x|\,|y| \cos\theta = \langle v, w\rangle$$

so essentially the trigonometric function is simply used to define our form in $\mathbb{R}^2$.

6.4 the spectral theorem

With our understanding from above we can make sense of the following definition.

Definition 176 (Euclidean space)

The space $\mathbb{R}^n$ equipped with the dot product is the called the **standard euclidean space**. An orthonormal basis for any euclidean space will refer the space back to the standard euclidean space

Definition 177 (Hermitian space)

The space $\mathbb{C}^n$ equipped with the standard hermititian form is called the **standard hermitian space**. An orthonormal basis for any hermitian space will refer the space back to the standard hermitian space

Definition 178 (adjoint operator)

Let $T : V \to V$ be a linear operator on a hermitian space V and let A be the matrix of T with respect to basis $\mathbf{B}$. The **adjoint operator** $T^* : V \to V$ is the operator whose matrix with respect to the *same basis* is the adjoint matrix A^*. In which case we write

$$T = T^*$$

We will explain this notation more below

Remark 179. *The* **hermitian operator** *is also known as the* **self-adjoint operator**

Proposition 180

Assume that we are working with linear operators on hermitian spaces with respect to orthonormal bases. Then base change matrix between 2 orthonormal bases is unitary

Proof. A hermition space is equipped with a hermition form by definition recall 177. Therefore there must exist a matrix of the form M and M' corresponding to the chosen orthonormal basis $\mathbf{B}$ and $\mathbf{B}'$. (Please see 184, do not confuse the matrix of form with matrix of linear operator). Hence we must have $\mathbf{B}^*\mathbf{B} = M' = I$ and also $\mathbf{B}'^*\mathbf{B}' = M' = I$. For example we could have

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

and

$$\mathbf{B}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so expressing the matrix of the linear map in terms of our new basis we must have

$$M' = P^*MP \quad \Rightarrow \quad I = P^*IP$$

and because P is invertible we must have that $P^* = P^{-1}$ so we have

$$P^*P = PP^* = I$$

Definition 181 (Normal Matrix)
A **normal matrix** is complex matrix A that commutes with its adjoint that is

$$A^*A = AA^*$$

Lemma 182
Let A be a complex $n \times n$ matrix and let P be an $n \times n$ unitary matrix. If A is normal, hermitian or unitary, so is P^*AP

Proof. Consider if A hermitian($A^* = A$) and P unitary($P^*P = PP^* = I$)

$$(A')^* = (P^*AP)^* = P^*A^*P = P^*AP = A'$$

Consider if A unitary($A^*A = AA^* = I$) and P unitary($P^*P = PP^* = I$)

$$(P^*AP)^*(P^*AP) = (P^*A^*P)(P^*AP) = I$$

Consider if A normal($A^*A = AA^*$) and P unitary($P^*P = PP^* = I$)

$$(P^*AP)^*(P^*AP) = (P^*A^*P)(P^*AP) = (P^*A^*AP) = (P^*AA^*P) = (P^*AP)(P^*A^*P) = (P^*AP)(P^*AP)^*$$

Now we now show just like how define a general square *matrix* to be hermitian,normal or unitary, we can also define linear operators similarly too if we assume their matrices of linear transformation are with respect to orthonormal basis.(we require extra conditions because we need to consider the basis with respect to the matrix of linear transformation not just the matrix itself. Recall $T(B) = BA$)

Theorem 183

Let T be a linear operator on a hermitian space. Assume its matrix be with respect to an *orthonormal basis*

T is **normal** if its matrix of linear transformation is normal in which we denote

$$T^*T = TT^*$$

T is **hermitan** if its matrix of linear transformation is hermitian in which we denote

$$T = T^*$$

T is **unitary** if its matrix of linear transformation is unitary in which we denote

$$T^*T = I$$

Show that this is well defined(i.e it is independent of the choice of orthornormal basis)

Proof. recall that $T(\mathbf{B}) = \mathbf{B}A$. But by definition 178 we are working with orthonormal bases(refer to 180 for examples). Therefore we know the matrix of the form M must be the identity matrix regardless which implies P is unitary by the previous theorems. Then we know that this definition makes sense since suppose we made a change of basis to another orthonormal basis, so the matrix of the linear operator becomes $P^{-1}AP$. But since P is unitary $P^{-1} = P^*$. Hence by the previous lemma the relations in our theorem follows

Remark 184. *Do not confuse the matrix of the form M with the matrix of the linear operator(map between same vector space) A. The matrix of the form M relates*

$$M = \mathbf{B}^T\mathbf{B} \quad \Rightarrow \quad I = P^*IP \quad \Rightarrow$$

for a given orthonormal basis $\mathbf{B}$. The linear operator maps a set of coeffient vector with respect to $\mathbf{B}$ to another coeffient vector with respect to the same basis $\mathbf{B}$. That is

$$\mathbf{B}AX = \mathbf{B}Y \quad \Rightarrow \quad \mathbf{B}P(P^{-1}AP)X' = \mathbf{B}PY'$$

So in summary $M \neq A$!!!

Proposition 185

Let T be a linear operator on hermitian space V and let T^* be the adjoint operator

(a) For all v and w in V, $\langle Tv, w\rangle = \langle v, T^*w\rangle$ and $\langle v, Tw\rangle = \langle T^*v, w\rangle$

(b) T is normal if and only i for all v and w in V, $\langle Tv, Tw\rangle = \langle T^*v, T^*w\rangle$

(c) T is hermitian if and only if for all v and w in V $\langle Tv, w\rangle = \langle v, Tw\rangle$

(d) T is unitary if and only if for all v and w in V $\langle Tv, Tw\rangle = \langle v, w\rangle$

Proof. for (a) consider that $v = \mathbf{B}X$ and $w = \mathbf{B}Y$ and so analagous to $\langle v, w\rangle = \langle \mathbf{B}X, \mathbf{B}Y\rangle = \langle X, Y\rangle$ like we have proven in the past we have:

$$\langle Tv, w\rangle = \langle \mathbf{B}AX, \mathbf{B}Y\rangle = (AX)^*MY = (AX)^*Y = X^*A^*Y = \langle v, T^*w\rangle$$

where our matrix of form M here is the identity matrix if you recall 180, therefore we basically are using standard hermitian form. The proof for the other formula is similar. For (b) consider from (a)

$$\langle TT * v, w\rangle = \langle T^*v, T^*w\rangle$$

since normal we have this equal to

$$\langle T^*Tv, w\rangle = \langle Tv, Tw\rangle$$

For (c) To show that T is Hermitian if and only if for all v and w in V, $\langle Tv, w\rangle = \langle v, Tw\rangle$. Assume T is Hermitian. By definition of the adjoint operator, we have:

$$\langle Tv, w\rangle = \langle v, T^*w\rangle.$$

Since T is Hermitian, it follows that $T^* = T$. Thus, we can rewrite the inner product:

$$\langle Tv, w\rangle = \langle v, Tw\rangle.$$

This proves the "if" part. Now, assume that $\langle Tv, w\rangle = \langle v, Tw\rangle$ for all v and w in V. We want to show that T is Hermitian, i.e., $T^* = T$. To show this, let $w = Tv$ for any $v \in V$. Then we have:

$$\langle Tv, Tv\rangle = \langle v, T(Tv)\rangle = \langle v, T^*(Tv)\rangle.$$

By the assumption, we can conclude:

$$\langle Tv, Tv\rangle = \langle v, T^*Tv\rangle.$$

Since this holds for all v, we have:

$$\langle v, T^*w\rangle = \langle Tv, w\rangle \Rightarrow T^* = T,$$

and thus T is Hermitian.
For (d) To show that T is unitary if and only if for all v and w in V, $\langle Tv, Tw\rangle = \langle v, w\rangle$. Assume T is unitary. By the definition of a unitary operator, we have:

$$\langle Tv, Tw\rangle = \langle v, w\rangle$$

for all v and w in V. Next, we show the converse. Assume that $\langle Tv, Tw\rangle = \langle v, w\rangle$ for all v and w in V. We want to show that T is unitary. To show this, we can use the properties of adjoint operators. Using the adjoint relation from (a):

$$\langle Tv, w\rangle = \langle v, T^*w\rangle.$$

Now, since T is defined such that:

$$\langle Tv, Tw\rangle = \langle v, w\rangle,$$

this leads us to conclude that the inner products maintain the same structure. Particularly, taking $w = v$ gives:

$$\langle Tv, Tv\rangle = \langle v, v\rangle,$$

indicating that T preserves the norm.We also know:

$$\langle Tv, w\rangle = \langle v, T^*w\rangle,$$

which leads us to:

$$\langle Tv, Tw\rangle = \langle v, w\rangle \implies T^* = T^{-1},$$

showing that T is unitary.

□

We now move on to the main applications of **spectral theorem** For the rest of the section we assume that V is positive definite.

Proposition 186
Let T be a linear operator on a hermitian space V and let W be a subspace of V. If W is T-invariant then the orthogonal space $W^{\perp}$ is T^*-invariant. If W is T^* invariant then W^T is T-invariant

Proof. Suppose that W is T invariant. To show that $W^{\perp}$ is T^* invariant we must show that if $u \in W^T$ then $T^*u \in W^{\perp}$. By definition $W^{\perp}$ means that $\langle w, T^*u\rangle = 0$ for all $w \in W$. However by 185 we know that $\langle w, T^*u\rangle = \langle Tw, u\rangle$. Since W is T invariant , Tw is in W then since $u \in W^{\perp}$, $\langle Tw, u\rangle = 0$ so $\langle w, T^*u\rangle = 0$ as desired. Since $T^{**} = T$ one obtains the second assertion by interchanging the roles of T and T^*

Lemma 187
Let T be a normal operator on a Hermitian space v and let v be an eigenvector of T with eigenvalue λ. Then v is also an eigenvector of T^* with eigenvalue $\overline{\lambda}$

Proof. Consider

(Case: i) $\lambda = 0$. Then $Tv = 0$ and we must show $T^*v = 0$ too. Since the form is positive definite it suffices to show that $\langle T^*v, T^*v\rangle = 0$. By 185 we know that

$$\langle T^*v, T^*v\rangle = \langle Tv, Tv\rangle = \langle 0, 0, \rangle = 0$$

(Case: ii) λ is arbitrary. Let S denote the linear operator $T - \lambda I$. Then v is an eigenvector for S with eigenvalue zero. Moreover, $S^* = T^* - \overline{\lambda} I$. Therefore by case 1 v is an eigen value for S^* with eigenvalue 0 which implies

$$S^*v = T^*v - \overline{\lambda}v = 0$$

so the proposition follows as desired

Theorem 188 (Spectral theorem for normal operators)
Consider

1. Let T be a normal operator on a hermitian space V. There is an orthonomrla basis of V consisting of eigenvectors for T
2. Matrix form: Let A be a normal matrix. There is a unitary matrix P such that P^*AP is diagonal

Theorem 189 (Spectral Theorem for Hermitian Operators)

Consider

1. Let T be a Hermitian operator on a Hermitian space V
   (a) There is an orthonormal basis of V consisting of eigenvectors of T
   (b) the eigenvalues of T are real numbers
2. Matrix form: Let A be a hermitian matrix
   (a) there is a unitary matrix P such that P^*AP is a real diagonal matrix
   (b) the eigenvalues of A are real numbers

Theorem 190 (Spectral Theorem for Unitary Matrices)

Consider

1. let A be a unitary matrix. There is a unitary matrix P such that P^*AP is diagonal
2. Every conjugacy class in the unitary group U_n contains a diagonal matrix

Proposition 191

Let T be a linear operator on a euclidean space V

1. T is symmetric if and only if for all v and w in V, $\langle Tv, W\rangle = \langle v, Tw\rangle$
2. T is orthogonal if and only if for all v and w in V, $\langle Tv, Tw\rangle = \langle v, w\rangle$

Theorem 192 (Spectral Theorem for symmetric operators)

1. Let T be a symmetric operator on a euclidean space V
   (a) there is an orthonormal basis of V consisting of eigenvectors of T
   (b) The eigenvalues of T are real numbers
2. Matrix form: Let A be a real symmetric matrix
   (a) there is an orthonormal matrix P such that P^TAP is a real diagonal matrix
   (b) The eigenvalues of A are real numbers

References

[1] Michael Artin. *Algebra*. eng. Second edition. Pearson modern classic. New York, NY: Pearson, 2018. ISBN: 978-0-13-468960-9.