Eigen value and Eigen vectors

Let $A \in M_{n\times n}(F)$ and $\lambda \in F$. We say that λ is an eigen value of the matrix A if $\exists$ $v(\neq 0) \in F^n$ such that $Av = \lambda v$. Such a non-zero $v \in F^n$ corresponding to that $\lambda \in F$ is called an eigen vector of A corresponding to the eigen value λ.

Examples ① Let $A = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{n\times n} \in M_{n\times n}(F)$.

Note that 0 is the only eigen value of A and every non-zero vector $v \in F^n$ is an eigenvector of A corresponding to the eigen value 0.

② Let $A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in M_{n\times n}(F)$.

Note that 1 is the only eigen value of A and every non-zero vector $v \in F^n$ is an eigen vector of A corresponding to 1.

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

Let $\lambda \in \mathbb{R}$. If $A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ for some $(x_1, x_2) \neq (0,0)$,

then we have
$$\lambda x_1 = x_2,$$
$$\lambda x_2 = -x_1.$$

As $(x_1, x_2) \neq (0,0)$, then $x_1 \neq 0$ or $x_2 \neq 0$.

Without loss of generality, suppose $x_1 \neq 0$.

Then $\lambda^2 x_1 = -x_1$ gives $\lambda^2 = -1$, which is not possible.

So the matrix A has no real eigen value.

④ Consider $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$.

Then for $\lambda \in \mathbb{C}$, we know $\lambda^2 = -1$ has solution.

In particular $\lambda = i, -i$.

So, $i, -i$ are possible complex eigen values of A.

Note that $(1, i)$ is an eigen vector of A corresponding to the eigen value i and $(1, -i)$ is an eigen vector of A corresponding to the eigen value $-i$.

The notion of eigen value and eigen vector can be generalized to the notion of eigen value and eigen vector of linear operator also.

Let V be a vector space over F. Let $T: V \longrightarrow V$ be a linear operator. A scalar $\lambda \in F$ is called an eigen value of T if $\exists\ v(\neq 0) \in V$ such that $T(v) = \lambda v$. Such non-zero v is called an eigen vector of T corresponding to the eigen value λ.

Remark Suppose $\dim_F V = n$.

Let $T: V \to V$ be a linear operator and $\lambda \in F$ an eigen value of T. So, $\exists\ v(\neq 0) \in V$ such that $T(v) = \lambda v$.

Let B be a basis of V.

Then $[T(v)]_B = [\lambda v]_B = \lambda [v]_B$.

Also we know, $[T(v)]_B = [T]_B [v]_B$.

So, $[T]_B [v]_B = \lambda [v]_B$.

So λ is an eigen value of $[T]_B$ as $[v]_B \neq 0$.

Also if $A \in M_{n\times n}(F)$ and λ is an eigen value of A. Then note that λ is an eigen value for $T : F^n \longrightarrow F^n$ defined by $T(x_1, \dots, x_n) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Examples

① Recall we discussed real and complex eigen values of $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

We now view it through linear operators.

Define, $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

by $T(x_1, x_2) = (x_2, -x_1) \; \forall (x_1, x_2) \in \mathbb{R}^2$.

This operator has no eigen value.

Whereas if we define $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$

by $T(x_1, x_2) = (x_2, -x_1) \; \forall (x_1, x_2) \in \mathbb{C}^2$,

then this operator has exactly two eigen values $i, -i$.

② Define $T : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ by $T(P(x)) = x P(x)$ $\forall \; P(x) \in \mathbb{R}[x]$.

If possible, let $\exists \; \lambda \in \mathbb{R}$ such that $T(P(x)) = \lambda P(x)$ for some $P(x) \neq 0$. Then $x P(x) = \lambda P(x)$. Clearly $\lambda \neq 0$. Now $\deg(x P(x)) > \deg P(x) = \deg(\lambda P(x))$. So T has no eigen value.

(3) Define $T : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$

by $T(a_0 + a_1 x + \cdots + a_n x^n) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$.

Note that for any $\lambda \in \mathbb{R}$,

$$T(P(x)) = \lambda P(x)$$

$$\Rightarrow a_1 + 2a_2 x + \cdots + n a_n x^{n-1} = \lambda a_0 + \lambda a_1 x + \cdots + \lambda a_n x^n$$

So, $\lambda a_n = 0$. If $\lambda \neq 0$, then $a_n = 0$.

Then arguing same way we have, $a_i = 0 \ \forall \ 0 \leq i \leq n$.

So, $P(x) = 0$.

Hence $\lambda = 0$.

So, 0 is the only possible eigen value of T.

Note that for any $c (\neq 0) \in \mathbb{R} \subseteq \mathbb{R}[x]$,

$$T(c) = 0 = 0 \cdot c.$$

So, any non-zero constant polynomial is an eigen vector of T corresponding to the eigen value 0.

<u>Eigen space</u>

Let $A \in M_{n \times n}(F)$ and $\lambda \in F$.

Denote $E(A, \lambda) := \{ v \in F^n : Av = \lambda v \}$.

Note that, $E(A, \lambda) \neq \Phi$ as $0 \in E(A, \lambda)$.

For $v_1, v_2 \in E(A, \lambda)$ and $c \in F$,

$$\begin{aligned} A(cv_1 + v_2) &= Acv_1 + Av_2 \\ &= cAv_1 + Av_2 \\ &= c\lambda v_1 + \lambda v_2 \\ &= \lambda(cv_1 + v_2) \end{aligned}$$

So, $cv_1 + v_2 \in E(A, \lambda)$.

This proves that $E(A, \lambda)$ is a subspace of F^n.

If λ is an eigen value of A, then we call the subspace $E(A,\lambda)$ the eigen space of λ.

Similarly for a linear operator $T: V \to V$ and an eigen value λ of T, the subspace

$$E(T,\lambda) := \{v \in V : T(v) = \lambda v\} \text{ of } V \text{ is}$$

called the eigen space of λ.

<u>Note</u> $E(T,\lambda) = \operatorname{Ker}(\lambda I - T)$.

<u>Observe</u>

$\lambda \in F$ is an eigen value of A (or of T)

$\Leftrightarrow$ $E(A,\lambda)$ (or $E(T,\lambda)$) contains a non-zero vector.

How to find the eigen values of a matrix?

For $A \in M_{n\times n}(F)$ and $c \in F$, we know

λ is an eigen value of A

$\Leftrightarrow$ $E(A,\lambda) \neq \{0\}$

$\Leftrightarrow$ $\{v \in F^n : Av = \lambda v\} \neq \{0\}$.

$\Leftrightarrow$ $\{v \in F^n : (\lambda I - A)v = 0\} \neq \{0\}$

$\Leftrightarrow$ $(\lambda I - A)x = 0$ has a non-trivial solution

Here I denotes the $n\times n$ identity matrix

$\Leftrightarrow$ The columns of $(\lambda I - A)$ are linearly dependent i.e. $(\lambda I - A)$ is not invertible.

$\Leftrightarrow$ $\det(\lambda I - A) = 0$.

Consider the polynomial $\det(xI - A) \in F[x]$. This polynomial is called the characteristic polynomial of A. So, $c \in F$ is an eigen value of A if and only if it is a root of the characteristic polynomial of A.

Remarks ① Let $A \in M_{n\times n}(F)$ be such that all its eigen values are in F. Then $\det A =$ product of all the eigen values of A.

char. poly $= \det(xI - A)$.
Now, the constant term of $\det(xI-A) = (-1)^n \det A$.
$\det(xI - A) = (x-\lambda_1)\cdots(x-\lambda_n)$ So, the constant term $= (-1)^n \lambda_1 \cdots \lambda_n$

② Let $A = (a_{ij}) \in M_{n\times n}(F)$. Then A, $A^t = (a_{ji})$ have same eigen values as $\det(xI - A) = \det(xI-A)^t = \det(xI - A^t)$

Recall we considered matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

$$xI - A = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix}$$

So, $\det(xI - A) = x^2 + 1$.

The polynomial x^2+1 has no root in $\mathbb{R}$ and has two roots $\pm i$ in $\mathbb{C}$.

Some more examples

① $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ $\quad xI - A = \begin{pmatrix} x-1 & -2 \\ -3 & x-4 \end{pmatrix}$

$$\det(xI - A) = (x-1)(x-4) - 6$$
$$= x^2 - x - 4x + 4 - 6$$
$$= x^2 - 5x - 2$$
$$x = \frac{5 \pm \sqrt{25+8}}{2} = \frac{5 \pm \sqrt{33}}{2}$$

So, $\frac{5 \pm \sqrt{33}}{2}$ are two real eigen values of A.

If we consider this matrix A over $\mathbb{Q}$, then there is no eigen value of A.

② $A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$. So, $xI - A = \begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix}$.

$$\det(xI - A) = (x-3)(x^2-2x+2) + (-2x+2) + (4+2x-4)$$
$$= x^3 - 3x^2 - 2x^2 + 6x + 2x - 6 - 2x + 2 + 2x$$
$$= x^3 - 5x^2 + 8x - 4.$$
$$= x^3 - x^2 - 4x^2 + 4x + 4x - 4$$
$$= x^2(x-1) - 4x(x-1) + 4(x-1)$$
$$= (x-1)(x-2)^2$$

So, $1, 2, 2$ are the eigen values of A.

__Thm__ Let $A, B \in M_{n\times n}(F)$ be two similar matrices. Then the characteristic polynomials of A and B respectively are the same and hence A, B have same eigen values.

__Pf__ Since A, B are similar, $\exists$ P invertible in $M_{n\times n}(F)$ such that $B = P^{-1}AP$.

$$xI - B = xI - P^{-1}AP$$
$$= P^{-1}P\,xI - P^{-1}AP$$
$$= P^{-1}xIP - P^{-1}AP$$
$$= P^{-1}(xI - A)P$$

So, $\det(xI-B) = \det\left(P^{-1}(xI-A)P\right)$
$= (\det P^{-1})(\det(xI-A))(\det P)$
$= \det(xI-A)$

How to find eigen values of $T: V \longrightarrow V$ where $\dim_F V < \infty$?

Let B be a basis of V and $\dim_F V = n$

consider $A = [T]_B$.

If B' is another basis of V, then

$A' := [T]_{B'} = P^{-1}AP$ for some invertible matrix $P \in M_{n\times n}(F)$.

So, A, A' are similar matrices.

So, $\det(xI-A) = \det(xI-A')$.

The characteristic polynomial of the matrix of T with respect to any basis is called the characteristic polynomial of T.

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

$\Leftrightarrow [T(v)]_B = \lambda [v]_B$

$\Leftrightarrow [T]_B [v]_B = \lambda [v]_B$

So the eigen values of T and the eigen values of $[T]_B$ are the same.

Now λ is an eigen value of $[T]_B$

$\Leftrightarrow \lambda$ is a root of $\det(xI-[T]_B)$.

So, λ is an eigen value of T if and only if λ is a root of the characteristic polynomial of T.

Algebraic and geometric multiplicity

Let $A \in M_{n \times n}(F)$ and λ be an eigen value of A.
We know, $E(A, \lambda)$ is a non-zero subspace of F^n
The dimension of $E(A, \lambda)$ over F is called the geometric multiplicity of λ.

Clearly, $1 \leq$ geometric multiplicity of $\lambda \leq n$.

because $E(A,\lambda) \neq \{0\}$

because $E(A,\lambda)$ is a subspace of F^n

By the algebraic multiplicity of λ we mean the multiplicity of λ in $\det(xI - A)$ i.e. the power of the factor $(x - \lambda)$ in $\det(xI - A)$.

Clearly, $1 \leq$ algebraic multiplicity of $\lambda \leq n$

because λ is a root of the polynomial $\det(xI - A)$

because $\det(xI - A)$ is an n-degree polynomial

Example For $A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$, we have seen that $2, 2, 1$ are the eigen values.

Here, the algebraic multiplicity of $1 = 1$,
and the algebraic multiplicity of $2 = 2$.

$$E(A, 1) = \{ v \in \mathbb{R}^3 : Av = v \}$$

$$A\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \Rightarrow \begin{array}{l} 3v_1+v_2-v_3=v_1, \\ 2v_1+2v_2-v_3=v_2, \\ 2v_1+2v_2 \quad = v_3. \end{array}$$

So, $v_2=0$ and $v_3=2v_1$.

So, $E(A,1) = \langle (1,0,2) \rangle$.

So, the geometric multiplicity of $1 = 1$.

$E(A,2) = \{ v \in F^3 : Av = 2v \}$.

Solving
$$\begin{array}{l} 3v_1+v_2-v_3=2v_1, \\ 2v_1+2v_2-v_3=2v_2, \\ 2v_1+2v_2 \quad = 2v_3, \end{array}$$
we have,
$$v_3 = 2v_2$$
and $v_1 = v_3 - v_2 = v_2$.

So, $E(A,2) = \langle (1,1,2) \rangle$.

So, geometric multiplicity of $2 = 1$.

<u>Relation between the geometric and the algebraic multiplicities of an eigen value</u>

Let $A \in M_{n\times n}(F)$ and λ an eigen value of A.

<u>claim</u> geometric multiplicity of λ
$\leq$ algebraic multiplicity of λ

We know, $T: F^n \longrightarrow F^n$ defined by $T(v) = Av$ is linear. Note that, $[T]_{\{e_1,\dots,e_n\}} = A$.

So the characteristic polynomial of T and of A are same.

Consider $E(T, \lambda) \subseteq F^n$.

Suppose, $\dim_F E(T, \lambda) = m$.

Let $\{v_1, \dots, v_m\}$ be a basis of $E(T, \lambda)$ over F.

We extend it to a basis B of F^n over F.

We write down $[T]_B$.

As $T(v_i) = \lambda v_i \;\; \forall \; 1 \leq i \leq m$, we have

$$[T]_B = \begin{pmatrix} \lambda I_{m\times m} & C_{m\times(n-m)} \\ O_{(n-m)\times m} & D_{(n-m)\times(n-m)} \end{pmatrix}$$

$$O_{(n-m)\times m} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & \cdots & 0 \end{pmatrix}_{(n-m)\times m}$$

So, $\det(xI - [T]_B) = (x-\lambda)^m \det(xI - D)$

So, the algebraic multiplicity of $\lambda \geq m$

m = the geometric multiplicity of λ.

Examples

(1) $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ A has only eigen value 0 with algebraic multiplicity 3.

$$E(A, 0) = \{v \in \mathbb{R}^3 : Av = 0\}$$
$$= \mathbb{R}^3$$

So geometric multiplicity of 0 is 3.

② $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ A has only eigen value 0 with alg. mult. 3.

$$E(A,0) = \{v \in \mathbb{R}^3 : Av = 0\}$$
$$= \{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_3 = 0\}$$
$$= \{(v_1, v_2, 0) \in \mathbb{R}^3\}$$
$$= \langle (1,0,0), (0,1,0) \rangle$$

So, geometric multiplicity of 0 is 2.

③ $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ A has only eigen value 0 with alg. mult. 3.

$$E(A,0) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 : A\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = 0\}$$
$$= \{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_2 = v_3 = 0\}$$
$$= \langle (1,0,0) \rangle$$

so the geometric multiplicity of 0 is 1.

Diagonalizable matrix

Let $A \in M_{n\times n}(F)$. We say that A is diagonalizable over F if $\exists$ an invertible matrix $P \in M_{n\times n}(F)$ such that $P^{-1}AP$ is a diagonal matrix.

In other words, We say, A is diagonalizable over F if A is similar to a diagonal matrix.

Examples

① $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$.

If A is digonalizable then $\exists$ P invertible in $M_{2\times 2}(\mathbb{R})$ so that $P^{-1}AP$ is a diagonal matrix.

Now the eigen values of $P^{-1}AP$ are the diagonal entries of $P^{-1}AP$. So $P^{-1}AP$ has eigen values in $\mathbb{R}$.

We know, the eigen values of A are same as the eigen values of $P^{-1}AP$ as they are similar. But we know A has no eigen value in $\mathbb{R}$. Hence A is not diagonalizable over $\mathbb{R}$.

Question Is A diagonalizable over $\mathbb{C}$?

Hint Find $P \in M_{2\times 2}(\mathbb{C})$ such that P is invertible and $P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

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

The eigen values of A are $1, 3$.

Claim A is similar to $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$.

To find invertible $P \in M_{2\times 2}(\mathbb{R})$ such that

$$P\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} P.$$

Write, $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

So,
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} a & 3b \\ c & 3d \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+2c & b+2d \\ 3c & 3d \end{pmatrix}.$$

$\therefore$ $3c = c$ and so $c = 0$.

$3b = b + 2d$ and so $b = d$.

Take $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Note In this example, for both the eigen values, the algebraic and the geometric multiplicities are the same.

(3) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$

A has eigen values 1, 1 i.e. 1 is the eigen value with algebraic multiplicity 2.

Claim A is not diagonalizable over $\mathbb{R}$.

If A is diagonalizable then A is similar to the diagonal matrix
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Let $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$ be such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ c & d \end{pmatrix}.$

So, $c = 0, d = 0.$

So, $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is not invertible.

<u>Note</u> Here the geometric multiplicity of $1 = 1$ as

$$E(A, 1) = \left\{ (v_1, v_2) \in \mathbb{R}^2 : \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\}$$

$$= \left\{ (v_1, v_2) \in \mathbb{R}^2 : \begin{pmatrix} v_1 + v_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\}$$

$$= \left\{ (v_1, 0) \in \mathbb{R}^2 \right\}$$ has dimension 1.

So here the geometric multiplicity of 1 $<$ the algebraic multiplicity of 1.

Next we shall try to understand diagonalizability via linear operators.

<u>Definition</u> Let V be a vector space over F and $T: V \to V$ a linear operator. We say that T is diagonalizable if $\exists$ a basis of V consisting of eigen vectors of T.

We will focus on linear operators on finite dimensional vector spaces.

Let $\dim_F V < \infty$. Say $\dim_F V = n$ and $T: V \to V$ is diagonalizable. So $\exists$ a basis $B = \{v_1, \dots, v_n\}$ of V such that $v_1, \dots, v_n$ are eigen vectors of T i.e.

$$T(v_i) = \lambda_i v_i \quad \forall \ 1 \le i \le n \text{ for some scalars } \underbrace{\lambda_1, \dots, \lambda_n}_{\text{need not be distinct}}.$$

Now,
$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \cdots & \lambda_n \end{pmatrix}.$$

So $[T]_B$ is a diagonal matrix.

On the other hand, if for any ordered basis B' of V, $[T]_{B'}$ is diagonal then B' consists of eigen vectors of T.

Let $B' = \{v_1', \dots, v_n'\}$ and
$$[T]_{B'} = \begin{pmatrix} \lambda_1' & & O \\ & \ddots & \\ O & & \lambda_n' \end{pmatrix}.$$

Then $\quad [T(v_i')]_{B'} = [T]_{B'} [v_i']_{B'}$

$$= \lambda_i \quad \forall \; 1 \leq i \leq n.$$

So, $\quad T(v_i') = \lambda_i v_i' \quad \forall \; 1 \leq i \leq n.$

So for linear operators on finite dimensional vector spaces we can use the following definition of diagonalizability.

Definition Let V be a finite dimensional vector space and T a linear operator on V. We say that, T is diagonalizable if $\exists$ an ordered basis B such that $[T]_B$ is a diagonal matrix. (In other words the matrix of T with respect to any basis of V is similar to a diagonal matrix).

Observe A matrix $A \in M_{n \times n}(F)$ is diagonalizable $\Leftrightarrow T : F^n \longrightarrow F^n$ defined by $T(v) = Av$ is diagonalizable.

Thm Let V be a finite dimensional vector space over F. Let $T : V \longrightarrow V$ be a diagonalizable linear operator. Then for every eigen value λ of T, the algebraic multiplicity of λ is same as the geometric multiplicity of λ.

Pf Let $\dim_F V = n$. Since T is diagonalizable, $\exists$ a basis $B = \{v_1, \dots, v_n\}$ of V consisting of eigen vectors.

Let $\lambda_1, \dots, \lambda_k$ be all the distinct eigen values of T. Suppose, among $v_1, \dots, v_n$,

d_1 many are eigen vectors corresponding to λ_1,

d_2 many are eigen vectors corresponding to λ_2,

$\vdots$

d_k many are eigen vectors corresponding to λ_k.

So after suitable ordering of B, we have an ordered basis B' of V such that

$$[T]_{B'} = \begin{pmatrix} \left.\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}\right\} d_1 \text{ times} & & \\ & \left.\begin{matrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_2\end{matrix}\right\} d_2 \text{ times} & \\ & & \ddots \\ & & & d_k \text{ times}\left\{\begin{matrix}\lambda_k & & \\ & \ddots & \\ & & \lambda_k\end{matrix}\right. \end{pmatrix}$$

The characteristic polynomial of $[T]_{B'}$

$= (x - \lambda_1)^{d_1} \cdots (x - \lambda_k)^{d_k}$

So, d_i = algebraic multiplicity of $\lambda_i \ \forall\ 1 \leq i \leq k$.

Note that, for each $1 \leq i \leq k$,

$\dim_F E(T, \lambda_i) \geq d_i$.

$\therefore$ alg. mult. of $\lambda_i \leq$ geo. mult. of λ_i.

In general, geo. mult. of $\lambda_i \leq$ alg. mult. of λ_i

So, alg. mult. of λ_i = geo. mult. of $\lambda_i \ \forall\ 1 \leq i \leq k$.

<u>Question</u> How about the converse?

Infact, the converse is also true. For that first we observe the following:

Let λ_1, λ_2 be two distinct eigen values of $T: V \longrightarrow V$. Let $T(v_1) = \lambda_1 v_1$ and $T(v_2) = \lambda v_2$ where $v_1 \neq 0$, $v_2 \neq 0$.

Then v_1, v_2 are linearly independent.

<u>Pf</u> If $c_1 v_1 + c_2 v_2 = 0$, then $c_1 T(v_1) + c_2 T(v_2) = 0$.

So, $c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 = 0$.

Now $\lambda_1 \neq \lambda_2$. So at least one of λ_1, λ_2 is non-zero.

Without loss of generality, let $\lambda_1 \neq 0$.

Then, $c_1 v_1 = -c_2 \frac{\lambda_2}{\lambda_1} v_2$.

So, $-c_2 v_2 = -c_2 \frac{\lambda_2}{\lambda_1} v_2$ as $c_1 v_1 = -c_2 v_2$.

$$\Rightarrow c_2 v_2 \left(\frac{\lambda_2}{\lambda_1} - 1 \right) = 0$$

Now, $\frac{\lambda_2}{\lambda_1} \neq 1$.

So, $c_2 v_2 = 0$.

$\Rightarrow c_2 = 0$ as $v_2 \neq 0$ (because if $c_2 \neq 0$, then $0 = c_2^{-1}(c_2 v_2) = v_2$)

$\therefore c_1 = 0$

Hence v_1, v_2 are l.I.

In general, we have the following lemma:

Lemma Let $T: V \longrightarrow V$ be linear and $\lambda_1, \dots, \lambda_k$ are some distinct eigen values of T. For each $1 \leq i \leq k$, let B_i be a l.I subset of $E(T, \lambda_i)$.

Then the disjoint union $\bigsqcup_{i=1}^{k} B_i$ is a l.I subset of V.

Pf Let $c_1 v_1 + \cdots + c_m v_m = 0$ where $c_j \in F$ and $v_j \in \bigsqcup_{i=1}^{k} B_i$ $\forall$ $1 \leq j \leq m$.

To show $c_j = 0$ $\forall$ $1 \leq j \leq m$.

We write,

This is by clubbing the $c_j v_j$'s together which belong to the same eigen space.

$c_1 v_1 + \cdots + c_m v_m$ as $\omega_1 + \cdots + \omega_k$ where $\omega_i \in E(T, \lambda_i)$ $\forall$ $1 \leq i \leq k$.

So, $\omega_1 + \cdots + \omega_k = 0$.

We show, $\omega_i = 0$ $\forall$ $1 \leq i \leq k$.

consider, $f_i(x) = \prod_{\substack{j \neq i \\ 1 \leq j \leq k}} \frac{(x - \lambda_j)}{(\lambda_i - \lambda_j)} \in F[x]$ for each $1 \leq i \leq k$

Note that, $f_i(\lambda_l) = \delta_{il}$, where $1 \leq i \leq k$, $1 \leq l \leq k$.

Now, $f_i(T): V \longrightarrow V$ is linear $\forall$ $1 \leq i \leq k$.

So, $f_i(T)(\omega_1 + \cdots + \omega_k) = 0$

$\Rightarrow f_i(T)(\omega_1) + \cdots + f_i(T)(\omega_k) = 0$

$\Rightarrow f_i(\lambda_1)\omega_1 + \cdots + f_i(\lambda_k)\omega_k = 0$

So, $\sum_{l=1}^{k} f_i(\lambda_l)\, w_l = 0$.

So, $\sum_{l=1}^{k} \delta_{il}\, w_l = 0$. So, $w_i = 0$.

Now, $w_i = 0 \Rightarrow$ corresponding c_j's are 0 as w_i is a linear combination of elements of B_i and B_i is l.I.

Another proof of the lemma (using PMI)

Statement to be proved using PMI

If sum of the vectors each coming from distinct eigen spaces is 0, then each vector is 0, i.e. if $w_1 + \cdots + w_t = 0$, where $w_1, \ldots, w_t$ each of them belongs to distinct eigen space. Then $w_i = 0 \ \forall 1 \leq i \leq t$.

Base If $t = 1$, then the statement is clearly true.

Induction hypothesis Assume the statement to be true for $t-1$.

Inductive step Let $w_1 + \cdots + w_t = 0$ where each w_i belongs to distinct eigen space.

Without loss of generality, let $T(w_i) = \lambda_i w_i$ $\forall\, 1 \leq i \leq t$.

So, $T(w_1) + \cdots + T(w_t) = 0$ gives $\sum_{i=1}^{t} \lambda_i w_i = 0$.

Write $\lambda_1 w_1 = -(\lambda_2 w_2 + \cdots + \lambda_t w_t)$.

$\Rightarrow -\lambda_1 (w_2 + \cdots + w_t) = -(\lambda_2 w_2 + \cdots + \lambda_t w_t)$

$\Rightarrow (\lambda_1 - \lambda_2) w_2 + \cdots + (\lambda_1 - \lambda_t) w_t = 0$.

Note that each $(\lambda_1 - \lambda_i) w_i \; \forall\, 2 \leq i \leq t$ belongs to distinct eigen space. This is a sum of $(t-1)$ many eigen vectors from $(t-1)$ many distinct eigen spaces. Hence by induction hypothesis $(\lambda_1 - \lambda_i) w_i = 0 \; \forall\, 2 \leq i \leq t$. Now, $\lambda_1 - \lambda_i \neq 0$, so $w_i = 0 \; \forall\, 2 \leq i \leq t$. So $w_1 = 0$.

Thm Let V be a finite dimensional vector space over F and $T: V \longrightarrow V$ is linear such that all the eigen values of T are in F. Let $\lambda_1, \ldots, \lambda_k$ be all the distinct eigen values of T such that for each $1 \leq i \leq k$, the alg. mult. of λ_i = the geo. mult. of λ_i.

(Equivalently, $\dim V = \sum_{i=1}^{k} \dim E(T, \lambda_i)$)

Then T is diagonalizable.

$$\dim V = \sum_{i=1}^{k} \dim E(T, \lambda_i) \leq \sum_{i=1}^{k} (\text{alg mult. of } \lambda_i) \leq \dim V$$

So, $\sum_{i=1}^{k} \dim E(T, \lambda_i) = \sum_{i=1}^{k} (\text{alg. mult. of } \lambda_i)$

Also, $\dim E(T, \lambda_i) \leq$ alg. mult. of $\lambda_i \; \forall\, 1 \leq i \leq k$.

So, $\dim E(T, \lambda_i) =$ alg. mult. of $\lambda_i \; \forall\, 1 \leq i \leq k$.

If $\dim E(T, \lambda_i) =$ alg. mult. of $\lambda_i \; \forall\, 1 \leq i \leq k$, then clearly $\sum_{i=1}^{k} \dim E(T, \lambda_i) = \sum_{i=1}^{k} (\text{alg. mult. of } \lambda_i)$

$= \dim V$.

Proof of the thm

To show T is diagonalizable i.e. $\exists$ a basis of V containing eigen vectors.

Let B_i be a basis of $E(T, \lambda_i)$ $\forall\ 1 \leq i \leq k$.

We know, $\bigsqcup_{i=1}^{k} B_i$ is a l.I subset of V. (by the lemma)

Given that $\sum_{i=1}^{k} |B_i| = \sum_{i=1}^{k} (\text{alg mult. of } \lambda_i)$

$$= \dim V$$

Then consider, $B = \bigsqcup_{i=1}^{k} B_i \subseteq V$.

So B is a basis of V containing eigen vectors of T.

Corollary Let $A \in M_{n \times n}(F)$ be such that it has n distinct eigen values. Then A is diagonalizable.

A has n distinct eigen values

$\Rightarrow$ alg. mult. of each eigen value is 1

$\Rightarrow$ geo. mult. of each eigen value is 1

$\Rightarrow$ T is diagonalizable as for each eigen value alg mult. = geo. mult.

Applications ① $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ is diagonalizable.

Using the fundamental thm of algebra } ② If $A \in M_{3\times3}(\mathbb{R})$ has at least one non-real eigen value, then A is diagonalizable over $\mathbb{C}$.

Fundamental thm of algebra (FTA)

Any non-constant polynomial in $\mathbb{C}[x]$ has all its roots in $\mathbb{C}$.

A corollary to FTA

Every $n\times n$ matrix over $\mathbb{C}$ has all its eigen values in $\mathbb{C}$.

Example consider,

$$A_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

characteristic polynomials of A_1, A_2, A_3 are same. In fact,

$$\text{char. poly } A_1 = \text{char. poly } A_2 = \text{char. poly } A_3$$
$$= x(x-1)^2.$$

alg. mult. of $0 = 1$.

so, geo. mult. of $0 = 1$.

alg. mult. of $1 = 2$.

so we shall check geo. mult. of 1 for each A_i.

$$E(A_1, 1) = \left\{ (x_1, x_2, x_3) \in F^3 : \begin{array}{r} x_1 + x_3 = x_1, \\ x_2 = x_2, \\ 0 = x_3. \end{array} \right\}$$

$$= \{ (x_1, x_2, 0) \in F^3 \}$$

$\dim E(A_1, 1) = 2$. So A_1 is diagonalizable.

$$E(A_2, 1) = \left\{ (x_1, x_2, x_3) \in F^3 : \begin{array}{l} x_1 = x_1, \\ x_2 + x_3 = x_2, \\ 0 = x_3 \end{array} \right\}$$

$$= \{ (x_1, x_2, 0) \in F^3 \}$$

$\dim E(A_2, 1) = 2$. So A_2 is diagonalizable

$$E(A_3, 1) = \left\{ (x_1, x_2, x_3) \in F^3 : \begin{array}{l} x_1 + x_2 = x_1, \\ x_2 = x_2, \\ 0 = x_3 \end{array} \right\}$$

$$= \{ (x_1, 0, 0) \in F^3 \}$$

$\dim E(A_3, 1) = 1 <$ alg. mult. of 1 for A_3

So, A_3 is not diagonalizable.

Recall

Let $A \in M_{n \times n}(F)$ be a diagonalizable matrix. Then we know A is similar to a diagonal matrix.

i.e. $\exists$ invertible $P \in M_{n \times n}(F)$ such that $P^{-1} A P$ is diagonal.

How to find such a P?

As A is diagonalizable, $\exists$ a basis of F^n over F consisting of eigen vectors. To get such a basis, from each distinct eigen space we take a basis and then take their union.

Now there are n l.I vectors of F^n in that basis. writing each of those vectors as column vectors we name them $c_1, \ldots, c_n$. Let c_i correspond to eigen value λ_i, $1 \leq i \leq n$.

$\left\{ \lambda_1, \ldots, \lambda_n \right.$ need not be distinct

consider,

$$P = (c_1 \;\; c_2 \;\; \cdots \;\; c_n) \in M_{n \times n}(F).$$

Now,
$$\begin{aligned} AP &= (Ac_1 \;\; Ac_2 \;\; \cdots \;\; Ac_n) \\ &= (\lambda_1 c_1 \;\; \lambda_2 c_2 \;\; \cdots \;\; \lambda_n c_n) \\ &= (c_1 \;\; c_2 \;\; \cdots \;\; c_n) \underbrace{\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}}_{= D} \\ &= PD \end{aligned}$$

Since columns of P are l.I, P is invertible.

<u>Recall</u> we have discussed the following matrix is diagonalizable.

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

char. poly $A = x(x-1)^2$.

$$E(A, 0) = \left\{ (x_1, x_2, x_3) \in F^3 : \begin{array}{l} x_1 + x_3 = 0, \\ x_2 = 0. \end{array} \right\}$$

$$= \{ (x, 0, -x) \in F^3 \} = \langle (1, 0, -1) \rangle$$

$$E(A, 1) = \left\{ (x_1, x_2, x_3) \in F^3 : \begin{array}{r} x_1 + x_3 = x_1, \\ x_2 = x_2, \\ 0 = x_3. \end{array} \right\}$$

$$= \{ (x_1, x_2, 0) \in F^3 \}$$

$$= \langle (1,0,0), (0,1,0) \rangle$$

Consider, $P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}$. Clearly P is invertible as the columns are l.I.

$$AP = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=D}$$

$$= PD$$

Cayley-Hamilton theorem

Let $A \in M_{n\times n}(F)$ and $f(x) = a_0 + a_1 x + \cdots + x^n \in F[x]$ be its characteristic polynomial. Then $f(A) = a_0 I + a_1 A + \cdots + A^n = O_{n\times n}$.

Caution $f(x) = \det(xI - A) \in F[x]$.

Do not put A in place of x to get a proof of Cayley-Hamilton theorem. Because we are taking determinant of the matrix $xI - A$ where $x \in F$. Then by taking determinant we get a polynomial in x which annihilates A due to Cayley-Hamilton theorem.

Application of Cayley-Hamilton theorem

For $A \in M_{n\times n}(F)$ invertible, we can find out the inverse of A from powers of A.

We know, A is invertible

$\Leftrightarrow AX = 0$ has only trivial solution.

$\Leftrightarrow 0$ is not an eigen value of A.

So, $f(x) = \det(xI - A)$ has non-zero constant term.

By Cayley-Hamilton theorem we get,

$$a_0 I = -a_1 A - a_2 A^2 - \cdots - A^n$$

Multiplying both sides by $a_0^{-1} A^{-1}$ we get,

$$\boxed{A^{-1} = -\frac{a_1}{a_0} I - \frac{a_2}{a_0} A - \cdots - \frac{1}{a_0} A^{n-1}.}$$

Examples

(1) $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

Char. poly. of $A = (x-1)(x-3)$
$= x^2 - 4x + 3$

So, $A^2 - 4A + 3I = 0$.

So, $A^{-1} = \frac{1}{3}(4I - A)$.

(2) $A = \begin{pmatrix} 1 & 3 & -5 \\ 3 & -1 & 5 \\ -5 & 5 & -5 \end{pmatrix}$

Char. poly. of $A = x^3 + 5x^2 - 60x + 100$.

So, $A^3 + 5A^2 - 60A + 100I = 0$.

So, $A^{-1} = \frac{1}{100}(60I - 5A - A^2)$.

Minimal polynomial of $A \in M_{n\times n}(F)$

The least degree monic polynomial annihilating A is called the minimal polynomial of A.

(Existence and uniqueness of such a poly. requires a proof that we shall skip for this course

Remark Note that the characteristic polynomial of A annihilates A. Also the characteristic polynomial of A is monic. So the degree of the minimal poly. $\leq$ the degree of the characteristic polynomial.

Important fact — Statement only for this course: The characteristic and the minimal polynomial of A have same roots and min. poly. of A divides char. poly. of A.

Example

$$A_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

char. poly. of $A_1 = x(x-1)^2$.

Min. poly. of $A_1 = x(x-1)$ as $A_1(A_1 - I) = 0$

$$A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

char. poly. of $A_2 = x(x-1)^2$.

Min poly of $A_2 = x(x-1)$ as $A_2(A_2 - I) = 0$.

$$A_3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

char. poly. of $A_3 = x(x-1)^2$. So the possibilities of min. poly. of A_3 are $x(x-1)$, $x(x-1)^2$.

Now, $A_3(A_3 - I) \neq 0$. So. min. poly. of A_3
$= x(x-1)^2$
$=$ char. poly. of A_3.

<u>Definition</u> Let $A \in M_{n\times n}(\mathbb{R})$. We say that,
A is symmetric if $A = A^t$. A is skew-symmetric if $A = -A^t$. A is orthogonal if $AA^t = A^tA = I_{n\times n}$.
i.e. A is invertible and $A^t = A^{-1}$.

Let $A \in M_{n\times n}(\mathbb{C})$. We say that,
A is Hermitian if $A = A^*$, where A^* denotes $\bar{A}^t$
A is skew-Hermitian if $A = -A^*$
A is unitary if $AA^* = I = A^*A$.
A is normal if $AA^* = A^*A$.

<u>Observe</u> ① If $A \in M_{n\times n}(\mathbb{C})$ is unitary, then $\det(AA^*) = 1$ i.e. $(\det A)(\overline{\det A}) = 1$. So, $|\det A|^2 = 1$. So, $|\det A| = 1$. If $A \in M_{n\times n}(\mathbb{R})$, then $\det A = 1$ or -1 here.

② If $A \in M_{n\times n}(\mathbb{C})$ is skew-Hermitian, then $a_{ii} = -\overline{a_{ii}}$ $\forall\ 1 \le i \le n$. So a_{ii} is purely imaginary $\forall\ 1 \le i \le n$. If $A \in M_{n\times n}(\mathbb{R})$, then in this case $a_{ii} = 0$ $\forall\ 1 \le i \le n$.

③ If $A \in M_{n\times n}(\mathbb{C})$ is Hermitian, then $a_{ii} = \overline{a_{ii}}$ $\forall\ 1 \le i \le n$. So $a_{ii} \in \mathbb{R}$ $\forall\ 1 \le i \le n$.

④ For $A \in M_{n\times n}(\mathbb{C})$, we can write $A = \dfrac{A+A^*}{2} + \dfrac{A-A^*}{2}$. Note that, $\dfrac{A+A^*}{2}$ is Hermitian and $\dfrac{A-A^*}{2}$ is skew-Hermitian. If $A \in M_{n\times n}(\mathbb{R})$, then in this case $\dfrac{A+A^*}{2}$ is symmetric and $\dfrac{A-A^*}{2}$ is skew-symmetric.

<u>Examples</u>

$$\begin{pmatrix} -3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & 4 \end{pmatrix} \quad \text{symmetric}$$

$$\begin{pmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{pmatrix} \quad \text{skew-symmetric}$$

$$\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix}$$ Orthogonal

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 5 & i \\ 1 & -i & 3 \end{pmatrix}$$ Hermitian

$$\begin{pmatrix} i & 0 & 1 \\ 0 & 5i & i \\ 1 & i & 3i \end{pmatrix}$$ Skew-Hermitian

$$\begin{pmatrix} i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ unitary

Theorem Let $A \in M_{n\times n}(\mathbb{C})$.

1. If A is Hermitian, then all the eigen values of A are in $\mathbb{R}$.
2. If A is skew-Hermitian, then any eigen value of A is either 0 or purely imaginary.
3. If A is unitary, then the absolute value of each eigen value is 1.

__Proof__ Let λ be an eigen value of A.

So $\exists\ v(\neq 0) \in \mathbb{C}^n$ such that $Av = \lambda v$.

So, $v^* A v = v^* \lambda v \in \mathbb{C}$.

So, $\lambda = \dfrac{v^* A v}{v^* v}$. Note that, $v^* v \in \mathbb{R}$.

① If $A^* = A$ i.e. $\bar{A}^t = A$ i.e. $A^t = \bar{A}$.

$\overline{v^* A v} = v^t \bar{A} \bar{v} = v^t A^t \bar{v} = (\bar{v}^t A v)^t$
$= (v^* A v)^t$
$= v^* A v$

So, $v^* A v \in \mathbb{R}$.

So, $\lambda \in \mathbb{R}$.

② If $A^* = -A$ i.e. $A^t = -\bar{A}$.

$\overline{v^* A v} = -v^t A^t \bar{v} = -(v^* A v)^t$
$= -v^* A v$.

So, $v^* A v$ is either 0 or purely imaginary.

So, λ is either 0 or purely imaginary.

③ If A is unitary i.e. $A^* A = I = A A^*$.

$Av = \lambda v$, $v \neq 0$ So, $v^* \underbrace{A^* A}_{= I} v = v^* \lambda^* \lambda v$

$\Rightarrow (Av)^* = (\lambda v)^*$

$\Rightarrow v^* A^* = v^* \lambda^*$ So, $v^* I v = v^* \lambda^* \lambda v$.

So, $v^* v = \lambda^* \lambda v^* v$.

So, $|\lambda|^2 = 1$. So, $|\lambda| = 1$.

Corollary Let $A \in M_{n\times n}(\mathbb{R})$. Due to FTA, we have all the eigen values of A are in $\mathbb{C}$. If A is symmetric then all the eigen values of A are real. If A is skew-symmetric and $\lambda \in \mathbb{R}$ is an eigen value of A then $\lambda = 0$. If A is orthogonal and $\lambda \in \mathbb{R}$ is an eigen value of A, then $\lambda = 1$ or -1.

Orthogonal transformation

Let $A \in M_{n\times n}(\mathbb{R})$ be an orthogonal matrix. The linear map $T: F^n \longrightarrow F^n$ defined by $\boxed{T(v) = Av}$ is called an orthogonal transformation.

$v \in F^n$ is written as a column

Example $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ defined by

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

So, $T(x_1, x_2, x_3) = (x_3, -x_2, -x_1)$

Note that, $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ is an orthogonal matrix i.e. $AA^t = I = A^tA$.

we have, $T(1,0,0) = (0,0,-1)$,
$T(0,1,0) = (0,-1,0)$,
$T(0,0,1) = (1,0,0)$.

Note that, if we consider the standard inner product on $\mathbb{R}^3$, then in this case,

$$\langle (1,0,0), (0,1,0)\rangle = \langle T(1,0,0), T(0,1,0)\rangle,$$
$$\langle (1,0,0), (0,0,1)\rangle = \langle T(1,0,0), T(0,0,1)\rangle,$$
$$\langle (0,1,0), (0,0,1)\rangle = \langle T(0,1,0), T(0,0,1)\rangle.$$

Thm Let $A \in M_{n\times n}(\mathbb{R})$ be an orthogonal matrix. Then for $v, w \in \mathbb{R}^n$ we have,

$\langle v, w\rangle = \langle Av, Aw\rangle$ where $\langle\, ,\rangle$ denote the standard inner product on $\mathbb{R}^n$.

Pf Note that,

$$\langle v, w\rangle = w^t v.$$

So,
$$\begin{aligned} \langle Av, Aw\rangle &= (Aw)^t (Av) \\ &= w^t A^t A v \\ &= w^t I v \\ &= w^t v \\ &= \langle v, w\rangle. \end{aligned}$$

corollary Orthogonal transformation preserves the norm with respect to the standard inner product. To see this,

$$\|T(v)\|^2 = \langle T(v), T(v)\rangle = \langle Av, Av\rangle = \|v\|^2.$$

Thm Let $A \in M_{n\times n}(\mathbb{R})$. Then A is orthogonal if and only if rows and columns of A consist of orthonormal vectors in $\mathbb{R}^n$.

Pf We write $A = (c_1 \; \cdots \; c_n)$.

So, $A^t = \begin{pmatrix} c_1^t \\ \vdots \\ c_n^t \end{pmatrix}$.

$$A^t A = \begin{pmatrix} c_1^t \\ \vdots \\ c_n^t \end{pmatrix} (c_1 \cdots c_n)$$

$$= \begin{pmatrix} c_1^t c_1 & \cdots & c_1^t c_n \\ \vdots & \ddots & \vdots \\ c_n^t c_1 & \cdots & c_n^t c_n \end{pmatrix}$$

If $A^t A = I$, then

$$\underbrace{c_i^t c_j}_{= \langle c_j, c_i \rangle} = \delta_{ij} = \begin{cases} 1 & \text{for } i=j \\ 0 & \text{otherwise} \end{cases}$$

So the columns of A are orthonormal vectors.

Writing $A = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$ and using $AA^t = I$, we get the rows of A are orthonormal vectors.

Now if for any $A \in M_{n\times n}(\mathbb{R})$, if columns of A are orthonormal then we get $A^t A = I$ and if rows of A are orthonormal then $A^t A = I$ and hence A is orthogonal.

Example

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$ is orthogonal matrix.

Unitary transformation

Let $A \in M_{n\times n}(\mathbb{C})$ be a unitary matrix. Then the linear map $T: \mathbb{C}^n \longrightarrow \mathbb{C}^n$ defined by $\boxed{T(v) = Av}$ is called a unitary transformation.

v is written as a column vector

Thm Unitary transformation preserve the Standard inner product on $\mathbb{C}^n$.

Proof is analogous to the orthogonal transformation.

Corollary Unitary transformation preserves the norm with respect to the Standard inner product on $\mathbb{C}^n$.

<u>Thm</u> Let $A \in M_{n\times n}(\mathbb{C})$. Then A is unitary if and only if the rows and columns of A are orthonormal vectors in $\mathbb{C}^n$.

Proof is analogous to the orthogonal transformation.

<u>Important facts</u> (statements only)

① Every symmetric matrix in $M_{n\times n}(\mathbb{R})$ is diagonizable over $\mathbb{R}$.
$\exists$ a basis of $\mathbb{R}^n$ consisting of orthonormal eigen vectors of A.
with respect to standard inner product on $\mathbb{R}^n$

so one can find an orthogonal matrix $P \in M_{n\times n}(\mathbb{R})$ such that $P^tAP = \boxed{P^{-1}AP}$ is diagonal.

<u>Remark</u> This is special as Gram-Schmidt orthogonalization does not guarantee that from eigen vectors we get eigen vectors.

② Let $A \in M_{n\times n}(\mathbb{C})$. If A is Hermitian or skew-Hermitian or unitary, then A is diagonalizable over $\mathbb{C}$.
Also $\exists$ a basis of $\mathbb{C}^n$ consisting of orthonormal eigen vectors of $\mathbb{C}$ and

hence a unitary matrix P such that $P^{-1}AP$ is diagonal.

Remark: skew-symmetric matrices are diagonalizable over $\mathbb{C}$ but may not be over $\mathbb{R}$.

Orthogonal matrices are diagonalizable over $\mathbb{C}$ but may not be over $\mathbb{R}$.

For example,

$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is skew-symmetric, orthogonal

but not diagonalizable over $\mathbb{R}$.

Quadratic form over $\mathbb{R}$

A real quadratic form in the variables $x_1, \ldots, x_n$ is a $\mathbb{R}$-linear combination of degree 2 monomials in $x_1, \ldots, x_n$.

Examples

1. $Q(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$

Observe,

$$Q(x_1, \ldots, x_n) = (x_1 \cdots x_n) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

② $Q(x_1, x_2) = x_1 x_2 + x_1^2 + x_2^2$

Observe,

$$Q(x_1, x_2) = (x_1 \quad x_2) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\| $$

$$(x_1 \quad x_2) \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

③ $Q(x_1, x_2) = x_1 x_2$

Observe,

$$Q(x_1, x_2) = (x_1 \quad x_2) \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Every quadratic form $Q(x_1, \dots, x_n)$ can be written as $(x_1 \dots x_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ where $A \in M_{n \times n}(\mathbb{R})$ is symmetric.

<u>How to find this matrix $A \in M_{n\times n}(\mathbb{R})$ symmetric?</u>

Suppose $Q(x_1, \dots, x_n)$ is a real quadratic form. Then consider,

$A = (a_{ij})_{1 \leq i, j \leq n}$ where

a_{ii} = co-efficient of x_i^2 $\forall \, 1 \leq i \leq n$ and

$$a_{ij} = a_{ji} = \frac{1}{2}(\text{co-efficient of } x_{ij}) \quad \forall\ 1 \leq i \neq j \leq n.$$

Then $Q(x_1, \dots, x_n) = (x_1 \dots x_n)\, A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ where A is symmetric.

Now as $A \in M_{n\times n}(\mathbb{R})$ is symmetric, A is diagonalizable over $\mathbb{R}$. In fact $\exists$ an orthogonal matrix P such that

$D = P^{-1} A P$ is a Diagonal matrix.

So,
$$\begin{aligned} A &= P D P^{-1} \\ &= P D P^{t} \\ &= (P^t)^t D (P^t) \end{aligned}$$
Denote $P^t = R$

So,
$$\begin{aligned} Q(x_1, \dots, x_n) &= (x_1 \dots x_n)\, A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= (x_1 \dots x_n)\, R^t D R \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^t R^t D R \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= (Rx)^t D (Rx), \text{ denoting } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= y^t D y, \text{ denoting } y = Rx. \end{aligned}$$

If $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & & & \lambda_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, then

$$\boxed{Q(x_1, \cdots, x_n) = \lambda_1 y_1^2 + \cdots + \lambda_n^2 y_n^2.}$$

This is called the canonical form/ principal axes form of Q.

Note that $\lambda_1, \cdots, \lambda_n$ are eigen values of A.

Examples ① Let $Q(x_1, x_2) = 7x_1^2 + 6x_1x_2 + 7x_2^2$.

We want to find out the canonical form of this quadratic form Q and the required change of variables.

Firstly we write,

$$Q(x_1, x_2) = 7x_1^2 + 6x_1x_2 + 7x_2^2$$

$$= (x_1 \quad x_2) \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Denote $A = \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix}$.

Char. poly. of $A = \det(xI - A)$

$$= (x-7)^2 - 9$$

$$= x^2 - 14x + 40$$

$$= (x-4)(x-10)$$

So, A is similar to

$$\begin{pmatrix} 4 & 0 \\ 0 & 10 \end{pmatrix}.$$

$$E(A,4) = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \begin{array}{l} 7a + 3b = 4a, \\ 3a + 7b = 4b. \end{array} \right\}$$

$$= \left\{ \begin{pmatrix} a \\ -a \end{pmatrix} \in \mathbb{R}^2 \right\}$$

$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ is an orthonormal basis of $E(A,4)$.

$$E(A,10) = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \begin{array}{l} 7a + 3b = 10a, \\ 3a + 7b = 10b \end{array} \right\}.$$

$$= \left\{ \begin{pmatrix} a \\ a \end{pmatrix} \in \mathbb{R}^2 \right\}$$

$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ is an orthonormal basis of $E(A,10)$.

Now we note that,

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$ is orthogonal.

Also, $P^t A P = \begin{pmatrix} 4 & 0 \\ 0 & 10 \end{pmatrix}$

So, $\boxed{y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = P^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1 - x_2}{\sqrt{2}} \\ \frac{x_1 + x_2}{\sqrt{2}} \end{pmatrix}}$.

change of variable

$Q(x_1, x_2) = 4y_1^2 + 10y_2^2$ is the canonical form.

② $Q(x_1, x_2) = 3x_1^2 + 8x_1x_2 - 3x_2^2$.

$$= (x_1 \quad x_2) \underbrace{\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}}_{= A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\det(xI - A) = (x^2 - 9) - 16$$
$$= (x-5)(x+5).$$

So, A is similar to $D = \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}$.

$$E(A, 5) = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \begin{array}{l} 3a + 4b = 5a, \\ 4a - 3b = 5b \end{array} \right\}$$

$$= \left\{ \begin{pmatrix} 2b \\ b \end{pmatrix} \in \mathbb{R}^2 \right\}.$$

$\begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$ forms an orthonormal basis of $E(A,5)$

$$E(A,-5) = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \begin{matrix} 3a+4b=-5a, \\ 4a-3b=-5b \end{matrix} \right\}.$$

$$= \left\{ \begin{pmatrix} a \\ -2a \end{pmatrix} \in \mathbb{R}^2 \right\}$$

$\begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}$ forms an orthonormal basis of $E(A,-5)$

Take $P = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix}$.

Note that P is orthogonal.

Also, $P^t A P = D$.

$$\boxed{y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = P^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2x_1+x_2}{\sqrt{5}} \\ \frac{x_1-2x_2}{\sqrt{5}} \end{pmatrix}}.$$ Change of variable

$Q(x_1,x_2) = 5y_1^2 - 5y_2^2$ is the canonical form.

Definition A real quadratic form $Q(x_1, \dots, x_n)$ is called

① positive definite if
$$Q(x) = x^t A x > 0 \ \forall \ x = (x_1, \dots, x_n) \neq 0.$$

② positive semidefinite if
$$Q(x) = x^t A x \geq 0 \ \forall \ x = (x_1, \dots, x_n) \neq 0.$$

③ negative definite if
$$Q(x) = x^t A x < 0 \ \forall \ x = (x_1, \dots, x_n) \neq 0.$$

④ negative semidefinite if
$$Q(x) = x^t A x \leq 0 \ \forall \ x = (x_1, \dots, x_n) \neq 0.$$

Important fact Let $Q(\underline{x}) = \underline{x}^t A \underline{x}$ be a real quadratic form, where A is symmetric.

important for using this criterion

Then,

$Q(\underline{x})$ is positive definite $\Leftrightarrow$ all the eigen values of A are positive.

$Q(\underline{x})$ is positive semidefinite $\Leftrightarrow$ all the eigen values of A are non-negative.

$Q(\underline{x})$ is negative definite $\Leftrightarrow$ all the eigen values of A are negative.

$Q(\underline{x})$ is negative semidefinite $\Leftrightarrow$ all the eigen values of A are non-positive.

Example $Q(x_1, x_2) = x_1^2 + 2x_1 x_2 + x_2^2$

If we write $Q(x_1, x_2)$

$$= (x_1 \ x_2) \underbrace{\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}}_{\text{not symmetric}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

then note that although all the eigen values of $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ are positive but $Q(x_1, x_2)$ is not positive definite as $Q(1, -1) = 0$.

But if we write,

$$Q(x_1, x_2) = (x_1 \ x_2) \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}}_{\text{symmetric}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

then we can note that as eigen values of $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ are $0, 2$ we have $Q(x_1, x_2)$ is positive semidefinite.