+ ×

− ÷

Fields

Definition Field

A **field** is a set $\mathbb{F}$ together with binary operations

addition: $\mathbb{F}\times\mathbb{F}\to\mathbb{F}$, $(\alpha,\beta)\mapsto\alpha+\beta$

multiplication: $\mathbb{F}\times\mathbb{F}\to\mathbb{F}$, $(\alpha,\beta)\mapsto\alpha\beta$

satisfying the following axioms

Commutativity: $\forall\alpha,\beta\in\mathbb{F}$,

$$\alpha+\beta=\beta+\alpha \qquad \alpha\beta=\beta\alpha$$

Associativity: $\forall\alpha,\beta,\gamma\in\mathbb{F}$,

$$\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma \qquad \alpha(\beta\gamma)=(\alpha\beta)\gamma$$

Identity elements: $\exists\, 0,1\in\mathbb{F}$, $0\neq1$ such that for all $\mathbb{F}$,

$$\alpha+0=\alpha \qquad \alpha1=\alpha$$

Inverses: $\forall\alpha\in\mathbb{F}$, $\exists -\alpha\in\mathbb{F}$ such that

$$\alpha+(-\alpha)=0$$

$\forall\alpha\in\mathbb{F}$, $\exists\,\alpha^{-1}\in\mathbb{F}$ such that

$$\alpha\alpha^{-1}=1$$

Distributivity: $\forall\alpha,\beta,\gamma\in\mathbb{K}$, we have

$$\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$$

Example $\mathbb{R}$ and $\mathbb{C}$ are fields.

$x\in\mathbb{F} \Rightarrow x$ is a scalar.

$\boxed{\mathbb{F} \text{ is the field of scalars}}$

$\underline{x}\in\mathbb{F}^n \Rightarrow \underline{x}$ is a vector

$\boxed{\mathbb{F}^n \text{ is the field of vectors}}$

Linear Algebra

- $\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R}\}$
- $\mathbb{C}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{C}\}$
- $\mathbb{F}^n$ where $\mathbb{F}$ is a field (abstract algebra)

Example of solving linear system of equations

Consider the following system of equations

$$\boxed{\begin{aligned} &1)\ 3x_1 - x_2 + 2x_3 + x_4 = 1 \\ &2)\ -x_1 + x_2 + 0 + x_4 = 0 \end{aligned}}$$

We can write it in matrix form

$$\begin{pmatrix} -3 & -1 & 2 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

We now eliminate x_1 using second row : gaussian elimination

$$\begin{array}{l} r1 \\ r2 \end{array} \begin{cases} 3x_1 - x_2 + 2x_3 + x_4 = 1 \\ -x_1 + x_2 + 0 + x_4 = 0 \end{cases} \Longrightarrow \begin{array}{r} r_1 \\ r_1 + 3r_2 \end{array} \begin{cases} 3x_1 - x_2 + 2x_3 + x_4 = 1 & (*1) \\ 0 + 2x_2 + 2x_3 + 4x_4 = 1 & (*2) \end{cases}$$

$\Rightarrow$ From $(*2)$

$$x_2 = \frac{1}{2} - x_3 - 2x_4$$

$\Rightarrow$ Substituting into first $(*1)$

$$3x_1 - \left(\frac{1}{2} - x_3 - 2x_4\right) + 2x_3 + x_4 = 1$$

$$\Rightarrow x_1 = \frac{1}{3}\left(\frac{3}{2} - 3x_3 - 3x_4\right) = \frac{1}{2} - x_3 - x_4$$

Therefore the solution is

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

Writing in vector form:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} - x_3 - x_4 \\ \frac{1}{2} - x_3 - 2x_4 \\ x_3 \\ x_4 \end{pmatrix}$$

Note: Properties of vectors (can be extended to n-dimensions)

1) $\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+b \\ c+d \end{pmatrix}$

2) $\alpha \begin{pmatrix} a \\ b \end{pmatrix} + \beta \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \alpha a \\ \alpha b \end{pmatrix} + \begin{pmatrix} \beta c \\ \beta d \end{pmatrix} = \begin{pmatrix} \alpha a + \beta c \\ \alpha b + \beta d \end{pmatrix}$

Therefore another way of writing the solutions is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix}}_{(*)} \underbrace{+ x_3 \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}}_{(*)}$$

Remarks:

1) There are 2 free variables in the solution

$$\boxed{\text{number of free variables} = \text{number of variables} - \text{number of independant equations}}$$

2) $(*) = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix}$ is the particular solution to the problem

3) $(*) = x_3 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}$ is the homogeneous solution

The homogeneous system of equations is

$$\left.\begin{aligned} 3x_1 - x_2 + 2x_3 + x_4 &= 0 \\ -x_1 + x_2 + 0 + x_4 &= 0 \end{aligned}\right\}$$ 0 on RHS of r1, r2

or $\begin{pmatrix} -3 & -1 & 2 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

General system of linear equations

We can write a general system of p linear equations in n unknowns as

$$\boxed{\begin{matrix} A_{11}x_1 + \cdots + A_{1n}x_n = y_1 \\ \vdots \qquad\qquad \vdots \\ A_{p1}x_1 + \cdots + A_{pn}x_n = y_p \end{matrix}} \qquad (*1.1)$$

This can also be written as

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pn} \end{pmatrix}, \quad \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\boxed{A\underline{x} = \underline{y}} \qquad \underline{x} \in \mathbb{F}^n, \quad \underline{y} \in \mathbb{F}^p \qquad (*1.2)$$

Note: Some system of equations of the form (*1.1) may not have solutions.

For example
$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 0 \end{cases}$$

Solution is obviously empty

The solution set to (*1.2) is

$$\boxed{S = \{\underline{x} \in \mathbb{F}^n : A\underline{x} = \underline{y}\}}$$

LINEAR COMBINATION AND LINEAR SUBSPACES

The first and most crucial property is how solution set S behaves for homogeneous equations

Linear Combination

Definition Linear Combination

Given $\underline{v}_1, \dots, \underline{v}_q \in \mathbb{F}^n$ and $\alpha_1, \dots, \alpha_q \in \mathbb{F}$, then

$$\alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q = \sum_{j=1}^{q} a_j \underline{v}_j$$

is called a linear combination of $\underline{v}_1, \dots, \underline{v}_n$

Example: In $\mathbb{R}^3$, the vector $(0,1,0)$ is a linear combination of vectors

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + (-1) \begin{pmatrix} 2 \\ -1 \\ -4 \end{pmatrix}$$

Supspaces

Definition Subspaces

A subset $S \subseteq \mathbb{F}^n$ is called a subspace (or linear subspace) of $\mathbb{F}^n$ if

(S1) $S \neq \emptyset$

(S2) $\underline{0} \in S$

(S3) $\forall\, \underline{v}_1, \dots, \underline{v}_q \in S$, $a_1 \underline{v}_1 + \cdots + a_q \underline{v}_q \in S$ (closed under linear combination)

Here, $\boxed{\underline{0} = \underline{v} - \underline{v} \quad \forall \underline{v} \in S}$ $\boxed{S \neq \mathbb{F}^n \Rightarrow S \text{ is a proper subspace}}$

Examples: Examples of subspaces

1) $\mathbb{F}^n \leq \mathbb{F}^n$ and $\mathbb{F}^n$ is a subspace
2) $\{\underline{0}\} \leq \mathbb{F}^n$ and $\{\underline{0}\}$ is a subspace
3) $S_0 = \{(a,b) \in \mathbb{F}^2 \mid a+b=0\} \leq \mathbb{F}^2$ is a subspace as

$$\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \in S_0 \iff \begin{aligned} a+b&=0 \\ c+d&=0 \end{aligned}$$

$$\alpha\begin{pmatrix}a\\b\end{pmatrix} + \beta\begin{pmatrix}c\\d\end{pmatrix} = \begin{pmatrix}\alpha a + \beta c\\ \alpha b + \beta d\end{pmatrix} \Longrightarrow \alpha a + \beta c + \alpha b + \beta d = \alpha(a+b) + \beta(c+d)$$

$$= 0$$

$$\Longrightarrow \alpha\begin{pmatrix}a\\b\end{pmatrix} + \beta\begin{pmatrix}c\\d\end{pmatrix} \in S_0$$

4) $S_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 1\}$ is not a subspace

Since $\underline{0} = (0,0) \notin S_1$

5) $S_2 := \{(a,b) \in \mathbb{F}^2 \mid a^2 - b^2 = 0\}$

Here $\underline{0} = (0,0) \in S_2$

But not a subspace because it is not closed under linear combination

For example take

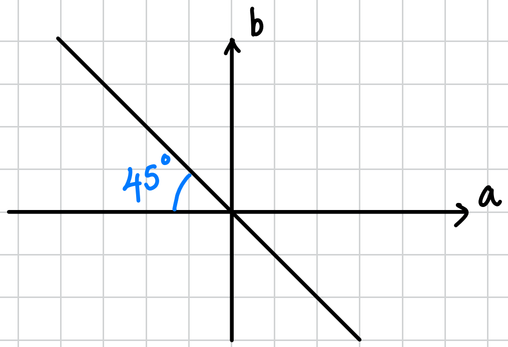
$\underline{u} = (1,-1) \in S_3$ and $\underline{v} = (1,1) \in S_3$ but

$\underline{u} + \underline{v} = (2,0) \notin S$

Geometry of subspace

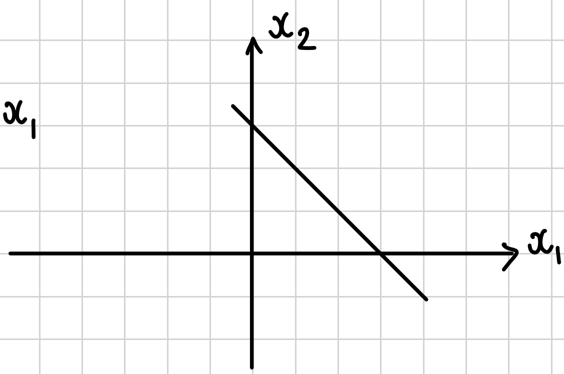
Take S_0:

$a + b = 0 \Longrightarrow b = -a$



Take S_1:

$x_1 + x_2 = 1 \Longrightarrow x_2 = 1 - x_1$

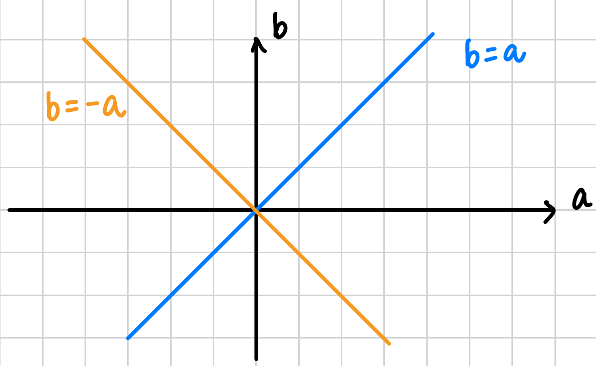


Take S_2:

$a^2 - b^2 = (a-b)(a+b) = 0$

union of curves

$b = a$, $b = -a$



Notation:

- $\underline{0}_p$ is the 0 vector with p dimensions
- $\underline{0}_n$ is the 0 vector with n dimensions

Consider the generalised linear system

$$\begin{aligned} A_{11}x_1 + \cdots + A_{1n}x_n &= y_1 \\ &\vdots \\ A_{p1}x_1 + \cdots + A_{pn}x_n &= y_p \end{aligned} \qquad \Bigg| \qquad A\underline{x} = \underline{y}, \quad A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pn} \end{pmatrix}, \quad \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$$

Theorem

The set of solutions

$$S = \{\underline{x} \in \mathbb{F}^n \mid A\underline{x} = \underline{y}\}$$

to any Linear system $A\underline{x} = \underline{y}$ of p equations in n variables is a linear subspace of $\mathbb{F}^n$

$$\Updownarrow$$

$\underline{y} = \underline{0}_p$ (linear system is homogeneous)

Proof:

$(\Leftarrow)$: Suppose $\underline{y} = \underline{0}_p$

Then indeed $\underline{0}_n$ is a solution to $A\underline{x} = \underline{0}_p$ as

$$A\underline{0}_n = \underline{0}_p$$

So $\underline{0}_n$ is in solution set $\underline{0} \in S$

Further let $\underline{v}_1, \dots, \underline{v}_q$ be solutions to $A\underline{x} = \underline{y}$, i.e. $A\underline{v}_1 = 0, \dots, A\underline{v}_q = \underline{0}$

Let $\alpha_1, \dots, \alpha_q \in \mathbb{F}$. Then we check that

$$A(\alpha_1\underline{v}_1 + \cdots + \alpha_q\underline{v}_q) = \underline{0}_p$$

$$A(\alpha_1\underline{v}_1 + \cdots + \alpha_q\underline{v}_q) = \alpha A\underline{v}_1 + \cdots + \alpha_q A\underline{v}_q = \underline{0}_p$$

$\Rightarrow \alpha_1\underline{v}_1 + \cdots + \alpha_q\underline{v}_q$ is a solution to $A\underline{x} = \underline{0}_p$

($\Longrightarrow$): Suppose that the set of solutions to $A\underline{x}=\underline{y}$ form a subspace

$\Longrightarrow$ $\underline{0}_n \in S$ is a solution

But then $A\underline{0}_n = \underline{0}_p = \underline{y} \Longrightarrow \underline{y} = \underline{0}$ ■

Example:

1) Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ $\quad \underline{y} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

The linear system $A\underline{x}=\underline{y}$ is then

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

$\Rightarrow \begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \end{cases}$ $\begin{array}{l} \rightarrow \text{defines plane in } x_1, x_2 \text{ as } x_3 = 0 \\ \rightarrow \text{defines plane in } x_2, x_3 \text{ as } x_1 = 0 \end{array}$

↳ intersection of planes $\Longrightarrow$ form a line

Solution:

$$\boxed{x_1 = -x_2 = x_3}$$

In vector form

$$S = \left\{ \begin{pmatrix} x_3 \\ -x_3 \\ x_3 \end{pmatrix} \middle| \, x_3 \in \mathbb{R} \right\} = x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

This is a vector equation of a line with x_3 as a parameter

2) Now consider same system with general $\underline{y}$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Longrightarrow \begin{cases} x_1 + x_2 = y_1 \\ x_2 + x_3 = y_2 \end{cases}$$

Solutions:

The general solution can be written as

$$\boxed{x_2 = y_2 - x_3}, \quad x_1 = y_1 - x_2 \Longrightarrow \boxed{x_1 = y_1 - y_2 + x_3}$$

$$\boxed{x_3: \text{ free parameter}}$$

Writing solution in vector form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 - y_2 + x_3 \\ y_2 - x_3 \\ 0 \end{pmatrix} = \underbrace{\begin{pmatrix} y_1 - y_2 \\ y_2 \\ 0 \end{pmatrix}}_{\text{shift from origin}} + x_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \rightarrow \text{solution to homogeneous part}$$

Vector equation of line with a shift from origin.

Solution of form : particular solution + homogeneous solution

Theorem

Let $A\underline{x} = y$ be a linear system of p equations, n variables

Then its solutions $S' = \{\underline{x} \in \mathbb{F}^n \mid A\underline{x} = \underline{y}\}$ are of form

$$S = \{ w + x_0 \mid \underbrace{A\underline{w} = \underline{0}_p}_{\text{homogeneous solution}},\ \underbrace{A\underline{x}_0 = y}_{\text{particular solution}} \}$$

In other terms, $\forall \underline{x}_0$ such that $A\underline{x}_0 = y$,

$$S = \{ w + x_0 \mid A\underline{w} = \underline{0} \}$$

i.e. $S = S'$

<u>Proof</u>: Using mutual containment, 1) $S \subseteq S'$, 2) $S' \subseteq S$

Let S' be the solution set

1) $S \subseteq S'$

$$A(\underline{w} + \underline{x}_0) = A\underline{w} + A\underline{x}_0 = \underline{0}_p + \underline{y} = y$$

2) $S' \subseteq S$

Let $\underline{v} \in S'$ be a solution. Then by definition

$$A\underline{v} = y \implies A(\underline{v} - \underline{x}_0) = A\underline{v} - A\underline{x}_0$$
$$= y - y$$
$$= \underline{0}$$

So $\underline{v} - \underline{x}_0$ is a solution to $A\underline{w} = \underline{0}$

So define $\underline{v} = x_0 + (\underline{v} - \underline{x}_0) = x_0 + w$ ∎

Example: Not every linear system has a solution for all $\underline{y}$. For example

$$A\underline{x} = \underline{y}$$

where

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The linear equations are then,

$$x_1 + x_2 + x_3 = 0 \qquad x_3 = 0, \qquad 0 = 1$$

This is NEVER true so solution set is empty, $S \neq \emptyset$

LINEAR INDEPENDENCE, BASES, DIMENSIONS

Linear Dependance/Independence

Definition Linear dependence

$\underline{v}_1, \ldots, \underline{v}_q \in \mathbb{F}^n$ is linearly dependant if $\exists (\alpha_1, \cdots, \alpha_q) \in \mathbb{F}^q \setminus \{(0, \cdots, 0)\}$ s.t

$$\alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q = \underline{0}$$

Otherwise, we say $\underline{v}_1, \ldots, v_q$ are linearly independant

Definition Linear independence

$\underline{v}_1, \cdots, \underline{v}_q$ are linearly independent if

$$\alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q = \underline{0} \implies \alpha_1 = 0, \cdots, \alpha_q = 0$$

The idea is linear dependence means one of these vectors can be written as a linear combination of others

For example since $\alpha_1 \neq 0$,

$$\underline{v}_1 = \frac{-1}{\alpha}(\alpha_2 \underline{v}_2 + \cdots + \alpha_q \underline{v}_q)$$

Remark: Any collection containing $\underline{0}$ is a linearly independent collection

$\underline{0}, \underline{v}_2, \cdots, \underline{v}_q$ is a linearly independant collection

$\underline{0} \cdot 1 + 0\,\underline{v}_2 + \cdots + 0\,\underline{v}_q = \underline{0}$, where

$$\alpha_1 = 1, \quad \alpha_2 = 0, \cdots, \alpha_q = 0$$

Example:

1) $S=\{\underline{v}\}\subseteq \mathbb{F}^n$ with $\underline{v}\neq \underline{0}$ is a linearly independent

$$\alpha\cdot \underline{v}=\underline{0} \iff \alpha=0$$

2) $\mathbb{F}^n$: $\underline{e_1},\dots,\underline{e_n}$ (standard basis)

$$\underline{e_1}=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad \underline{e_2}=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\quad \dots,\ \underline{e_n}=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

System $\underline{e_1},\dots,\underline{e_n}$ is linearly independant

$$\alpha_1\underline{e_1}+\dots+\alpha_n\underline{e_n}=\begin{pmatrix}\alpha_1\\\vdots\\\alpha_n\end{pmatrix}=0 \iff \alpha_i=0,\quad \forall\, 0\leq i\leq n$$

3) $\underline{u}=\begin{pmatrix}i\\1\end{pmatrix}$ $\quad \underline{v}=\begin{pmatrix}-1\\i\end{pmatrix}$ in $\mathbb{C}^2$

$$-i\,\underline{u}+\underline{v}=-i\begin{pmatrix}i\\1\end{pmatrix}+\begin{pmatrix}-1\\i\end{pmatrix}=\begin{pmatrix}1\\-i\end{pmatrix}+\begin{pmatrix}-1\\i\end{pmatrix}=\underline{0}$$

$\Longrightarrow$ linearly dependant

4) $\underline{u_2}=\begin{pmatrix}i\\1\end{pmatrix}$ $\quad \underline{v_2}=\begin{pmatrix}1\\i\end{pmatrix}$

$$\alpha,\beta\in\mathbb{C},\quad \alpha\underline{u_2}+\beta\underline{v_2}=\begin{pmatrix}0\\0\end{pmatrix}\Longrightarrow \begin{pmatrix}i\alpha\\\alpha\end{pmatrix}+\begin{pmatrix}\beta\\i\beta\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

$$\Longrightarrow \begin{pmatrix}i\alpha+\beta\\\alpha+i\beta\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

$$\Longrightarrow \begin{cases} i\alpha+\beta=0\\ \alpha+i\beta=0\end{cases}$$

gaussian elimination

$$\Longrightarrow \begin{cases} 2i\beta=0\\ \alpha+i\beta=0\end{cases}$$

$\Longrightarrow \beta=0,\ \alpha=0$ only solution

Therefore linearly independant

Spans

Definition Span

Let $\mathcal{C} \subset \mathbb{F}^n$ be a non-empty collection of vectors. $\mathcal{C} = \{\underline{v}_1, \cdots, \underline{v}_n\}$

The span of $\mathcal{C}$ denoted

$$Sp(\mathcal{C})$$

is the set of all linear combination of $\mathcal{C}$

$$Sp(\mathcal{C}) = \{ \underline{u} \in \mathbb{F}^n \mid \underline{u} = \alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n \text{ for some } \alpha_i \in \mathbb{F},\ \underline{v}_i \in S \}$$

By convention, $\boxed{Sp(\phi) = \{\underline{0}\}}$

Remark:

i) We always have $S \subseteq Sp(S) \subseteq \mathbb{F}^n$

ii) $\mathcal{C}$ may be infinite or finite but $Sp(\mathcal{C})$ consists of linear combination of a finitely many terms

Lemma

For any $\mathcal{C} \subset \mathbb{F}^n$, $\mathcal{C} \neq 0$,

$$Sp(\mathcal{C}) \text{ is a subspace of } \mathbb{F}^n$$

In fact, $Sp(\mathcal{C})$ is the smallest subspace of $\mathbb{F}^n$ containing $\mathcal{C}$, i.e. if $S \leq \mathbb{F}^n$ is any subspace with $\mathcal{C} \subseteq S$, then $Sp(\mathcal{C}) \subseteq S$

Proof: Take any collection of vectors

$$\underline{v}_1, \cdots, \underline{v}_n \quad \text{where} \quad \underline{v}_i \in \mathbb{F}^n, \quad i \in [\![1, n]\!]$$

Let $\mathcal{C} = \{\underline{v}_1, \cdots, \underline{v}_n\}$

Then the span is

$$Sp(\mathcal{C}) = \{ \alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n \mid \alpha_i \in \mathbb{F},\ \underline{v}_i \in \mathcal{C} \}$$

Clearly $\underline{0} \in Sp(\mathcal{C})$ when $\alpha_i = 0 \quad \forall i \in [\![1, n]\!]$

We need to show $Sp(\mathcal{C})$ is closed under linear combinations.

Suppose $\underline{a}, \underline{b} \in Sp(\mathcal{C})$. By definition of span,

$$\underline{a} = \alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n$$

$$\underline{b} = \beta_1 \underline{v}_1 + \cdots \beta_n \underline{v}_n$$

Then

$$\begin{aligned} \lambda \underline{a} + \mu \underline{b} &= \lambda(\alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n) + \mu(\beta_1 \underline{v}_1 + \cdots \beta_n \underline{v}_n) \\ &= (\lambda\alpha_1 + \mu\beta_1)\underline{v}_1 + \cdots + (\lambda\alpha_n + \mu\beta_n)\underline{v}_n \end{aligned}$$

$\Longrightarrow \lambda \underline{a} + \mu \underline{b} \in Sp(\mathcal{C})$ as it is a linear combination and by definition of span.

Finally we need to show that $Sp(\mathcal{C})$ is the smallest subspace

We have shown that $Sp(\mathcal{C})$ is a subspace and pretty clear that

$$\underline{v}_i \in Sp(\mathcal{C}) \quad \forall i \in [\![1, n]\!] \text{ since } \underline{v}_i = 0 . \underline{v}_1 + \cdots + 1 \underline{v}_i + 0 \underline{v}_{i+1} \cdots 0 \underline{v}_n$$

Suppose M is smallest subspace containing $\underline{v}_1, \cdots, \underline{v}_n$. We show that $Sp(\mathcal{C}) = M$

1) $\underline{v}_i \in Sp(\mathcal{C})$ but M is the smallest subspace containing $\underline{v}_1, \cdots, \underline{v}_n$

$$\Longrightarrow M \subseteq Sp(\mathcal{C})$$

2) Suppose $\underline{v}_i \in M$ for $1 \leq i \leq n \Longrightarrow \alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n \in M \quad \forall (\alpha_1, \cdots, \alpha_n) \in \mathbb{F}^n$ defn of subspace

$\Longrightarrow \alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n \in Sp(\mathcal{C})$ defn of span

$\Rightarrow Sp(\mathcal{C}) \subseteq M$

By mutual inclusion,

$$M = Sp(\mathcal{C})$$

■

For any subspace $S \subset \mathbb{F}^n$, we say $\mathcal{C}$ **spans** S if

$$Sp(\mathcal{C}) = S$$

and $\mathcal{C}$ is called the **spanning set** for S or S is **spanned by** $\mathcal{C}$

Example

1) $\mathbb{F}^3$; $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

$$sp(\underline{e}_1, \underline{e}_2) = \{\alpha \underline{e}_1 + \beta \underline{e}_2 \mid \alpha, \beta \in \mathbb{F}\}$$

$$= \left\{ \begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix} \;\middle|\; \alpha, \beta \in \mathbb{F} \right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{F}^3 \;\middle|\; x_3 = 0 \right\}$$

Now consider $sp(\underline{e}_1, \underline{e}_2, \underline{u})$, $\underline{u} = (1,1,0)$.

Since $\underline{u} = \underline{e}_1 + \underline{e}_2$, it is clear that

$$sp(\underline{e}_1, \underline{e}_2, \underline{u}) = sp(\underline{e}_1, \underline{e}_2)$$

Further, since $\underline{e}_1 = \underline{u} - \underline{e}_1$ and $\underline{e}_2 = u - \underline{e}_1$,

$$sp(\underline{e}_1, \underline{e}_2, \underline{u}) = sp(\underline{e}_1, \underline{u}) = sp(\underline{e}_2, \underline{u})$$

2) $\mathbb{F}^3$; define $\underline{v} = (1,1,1)$ then $\underline{e}_3 = \underline{v}_1 - \underline{e}_1 - \underline{e}_2$ so

$$sp(\underline{e}_1, \underline{e}_2, \underline{v}) = sp(\underline{e}_1, \underline{e}_2, \underline{e}_3) = \mathbb{F}^3$$

Example

1) $\mathbb{R}^2$; $\underline{v}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\underline{v}_2 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$, $\underline{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

<u>Questions</u>

a) Is it a linearly dependant system

b) Is $sp(\underline{v}_1, \underline{v}_2, \underline{v}_3) = \mathbb{R}^2$

<u>Answer</u>

$$1)\ \alpha \underline{v}_1 + \beta \underline{v}_2 + \gamma \underline{v}_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \alpha \begin{pmatrix} 3 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ 3 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 3\alpha - 2\beta + 0 \\ -\alpha + 3\beta + \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \boxed{\beta = \frac{3}{2}\alpha}$$

$$-\alpha + \frac{9}{2}\alpha + \gamma = 0 \Rightarrow \boxed{\gamma = -\frac{7}{2}\alpha}$$

Therefore we get

$$\boxed{\begin{aligned} \alpha &= -\frac{2}{7}\gamma \\ \beta &= \frac{-3}{7}\gamma \\ \gamma&: \text{free parameter} \end{aligned}}$$

Since $\gamma \neq 0, \Longrightarrow \alpha, \beta \neq 0 \Longrightarrow$ linearly dependent

2) $\alpha \underline{v}_1 + \beta \underline{v}_2 + \gamma \underline{v}_3 = \begin{pmatrix} x \\ y \end{pmatrix}$ where $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$

inhomogeneous system

We get

$$3\alpha - 2\beta + 0 = x$$
$$-\alpha + 3\beta + \gamma = y \quad \times 3$$

$$3\alpha - 2\beta = x$$
$$-7\beta + 3\gamma = x + 3y$$

gaussian elimination

$$\boxed{\begin{aligned} \alpha &= \frac{1}{3}x + \frac{2}{3}\beta \\ \beta &= \beta \ (\text{free parameter}) \\ \gamma &= \frac{1}{3}x + y + \frac{7}{3}\beta \end{aligned}} \Longrightarrow \text{spans } \mathbb{R}^2$$

Basis

Definition Basis

Let $S \subseteq \mathbb{F}^n$ be a non-trivial $S \neq \{0\}$ subspace of $\mathbb{F}^n$,

A collection $B = \{\underline{v}_1, \cdots, \underline{v}_q\} \subseteq S$ forms a basis if

i) $\underline{v}_1, \cdots, \underline{v}_q$ is linearly independent

ii) $sp(\underline{v}_1, \cdots, \underline{v}_q) = S$

By definition,

$$\boxed{\text{basis of } \{0\} \text{ is } \emptyset}$$

Lemma

Let $S \subseteq \mathbb{F}^n$ be a subspace with an ordered basis $(\underline{v}_1, \cdots, \underline{v}_q) = \underline{B}$

Then $\forall \underline{u} \in S$ can be uniquely written in form

$$\underline{u} = \alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q$$

Proof:

B is a basis so $sp(B) = S$.

Thus $\forall \underline{u} \in S$, $\exists \alpha_1, \cdots, \alpha_q \in \mathbb{F}$ such that

$$\underline{u} = \alpha_1 \underline{v}_1, + \cdots + \alpha_q \underline{v}_q$$

Let $\beta_1, \cdots, \beta_q \in \mathbb{F}$ such that

$$\underline{u} = \beta_1 \underline{v}_1 + \cdots + \beta_q \underline{v}_q$$

Then

$$\underline{u} = \alpha_1 \underline{v}_1, + \cdots + \alpha_q \underline{v}_q = \beta_1 \underline{v}_1 + \cdots + \beta_q \underline{v}_q$$

$$\Leftrightarrow \ (\alpha_1 - \beta_1)\underline{v}_1 + \cdots + (\alpha_q - \beta_q)\underline{v}_q = \underline{0}$$

But $(\underline{v}_1, \cdots, \underline{v}_q)$ is a basis so $\underline{v}_1, \cdots, \underline{v}_q$ is linearly independent

$$\Rightarrow \alpha_1 = \beta_1, \cdots, \alpha_q = \beta_q$$

$$\Rightarrow (\alpha_1, \cdots, \alpha_q) \text{ is unique}$$

■

Example:

i) $\mathbb{F}^n$; $\underline{e}_1, \dots, \underline{e}_n$ (standard basis)

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Checked lecture 3 that $\underline{e}_1, \dots, \underline{e}_n$ linearly independent

$$\forall \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n, \quad \alpha_1 \underline{e}_1 + \dots + \alpha_n \underline{e}_n = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \text{so} \quad sp(\underline{e}_1, \dots, \underline{e}_n) \in \mathbb{F}^n$$

Thus $\underline{e}_1, \dots, \underline{e}_n$ is a basis

ii) $$\underline{v}_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \underline{e}_1 + \dots + \underline{e}_n \quad \underline{v}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \underline{e}_2 + \dots + \underline{e}_n \quad \underline{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \underline{e}_3 + \dots + \underline{e}_n$$

$$\vdots$$

$$\dots \quad \underline{v}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \underline{e}_n$$

This forms a basis

Assume $\beta_1 \underline{v}_1 + \dots + \beta_n \underline{v}_n = \underline{0} \implies \beta_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \dots + \beta_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \underline{0}$

$$\implies \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Therefore we have system of equations

1) $\beta_1 = 0 \implies \beta_1 = 0$

2) $\beta_1 + \beta_2 = 0 \implies \beta_1 + \beta_2 = 0$ and $\beta_1 = 0 \implies \beta_2 = 0$

$\vdots$

n) $\beta_1 + \beta_2 + \dots + \beta_n = 0 \implies \beta_1 + \dots + \beta_n = 0$ and $\beta_1, \dots, \beta_{n-1} = 0 \implies \beta_n = 0$

Therefore $(\beta_1, \dots, \beta_n) = (0, \dots, 0) \implies$ linearly independent

Showing this spans $\mathbb{F}^n$, for any

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n$$

Lets try to find $\beta_1, \cdots, \beta_n$ such that

$$\beta_1 \underline{v}_1 + \cdots + \beta_n \underline{v}_n = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \Longrightarrow \beta_1 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \cdots + \beta_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Therefore we have system of equations:

$$\left.\begin{array}{ll} 1) & \beta_1 = \alpha_1 \\ 2) & \beta_1 + \beta_2 = \alpha_2 \end{array}\right\} \Longrightarrow \beta_1 = \alpha_1, \quad \Longrightarrow \beta_2 = \alpha_2 - \alpha_1$$

$$\vdots \qquad \Longrightarrow \beta_{n-1} = \alpha_{n-1} - \alpha_{n-2}$$

$$\left.\begin{array}{ll} n-1) & \beta_1 + \beta_2 + \cdots + \beta_{n-1} = \alpha_{n-1} \\ n) & \beta_1 + \beta_2 + \cdots + \beta_{n-1} + \beta_n = \alpha_n \end{array}\right\} \Longrightarrow \beta_n = \alpha_n - \alpha_{n-1}$$

gaussian elimination

The solution is therefore

$$\beta_1 = \alpha_1, \ \beta_2 = \alpha_2 - \alpha_1 \ \cdots\cdots \ \beta_n = \alpha_n - \alpha_{n-1}$$

Therefore system forms basis

Steinitz Exchange Lemma

Lemma Steinitz exchange lemma

Let $S \subseteq \mathbb{F}^n$ be a subspace. $S \neq \{\underline{0}\}$ non trivial and let $\{\underline{v}_1, \cdots, \underline{v}_q\}$ be a basis.

Then $\forall \underline{u} \in S \setminus \{\underline{0}\}$, $\exists\, j \in \{1, \cdots, q\}$ s.t swapping $\underline{u}$ and $\underline{v}_j$ forms a basis.

$\{\underline{v}_1, \cdots, \underline{v}_{j-1}, \underline{u}, \underline{v}_{j+1}, \cdots, \underline{v}_q\}$ is a basis of S as well.

Proof: $\underline{u} = \alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q$

As $\underline{u} \neq \underline{0}$, $\exists j \in [\![1, q]\!]$ such that $\boxed{\alpha_j \neq 0}$

Let's prove that $\underline{v}_1, \cdots, \underline{v}_{j-1}, \underline{u}, \underline{v}_{j+1}, \cdots, \underline{v}_q$ forms a basis

Since $\{\underline{v}_j,\ j = 1, \cdots, q\}$ forms a basis for S,

$$\underline{u} = \sum_{i=1}^{q} \alpha_i \underline{v}_i = \alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q \qquad (*)$$

Since $\alpha_j \neq 0$, we can write

$$\underline{v}_j = \frac{1}{\alpha_j}\left(\underline{u} - \sum_{\substack{i=1 \\ i \neq j}}^{q} \alpha_i \underline{v}_i\right) \implies \underline{v}_j = \alpha_j^{-1} \underline{u} + \sum_{\substack{i=1 \\ i \neq j}}^{q} (\alpha_j^{-1} \alpha_i)\, \underline{v}_i$$

Now $\mathcal{C} = \{\underline{v}_1, \cdots, \underline{v}_{j-1}, \underline{u}, \underline{v}_{j+1}, \cdots \underline{v}_q\}$ still spans S

To show linear independence, suppose

$$\underline{0} = \sum_{\substack{i=1 \\ i \neq j}}^{q} \beta_i \underline{v}_i + \gamma \underline{u}$$

for some $\beta_i \in \mathbb{F}$, $1 \leq i \leq q$, $\gamma \in \mathbb{F}$, $i \neq j$. Substituting $\underline{u}$ from $(*)$

$$\underline{0} = \sum_{\substack{i=1 \\ i \neq j}}^{q} \beta_K \underline{v}_K + \gamma \left(\sum_{i=1}^{q} \beta_K \underline{v}_K\right) = \sum_{\substack{i=1 \\ i \neq j}}^{q} (\beta_i + \gamma \alpha_i) v_i + \gamma \alpha_j \underline{v}_j$$

By linear independence of $\{\underline{v}_j,\ j = 1, \cdots, q\}$

$$\beta_K + \gamma \alpha_K = 0 \text{ for each } K \neq j$$

Since $\alpha_j \neq 0$ and $\gamma a_j = 0 \implies \gamma = 0$ and $\beta_k + \gamma\alpha_k = 0$ for each $k \neq j$

$$\implies \beta_k = 0 \quad \forall k \neq j$$

Thus $\mathcal{C}$ is linearly independant ■

> Moreover we can take/swap any index j where $\alpha_j \neq 0$ in
> $$\underline{u} = \sum_{j=1}^{q} \alpha_j \underline{v_j}$$

Example:

Consider a basis for $\mathbb{R}^2$, $\underline{v_1}$ and $\underline{v_2}$

$$\underline{v_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \underline{v_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \underline{u} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$\underline{u} = 2\underline{v_1} + 0.\underline{v_2}$

Swapping $\underline{u}$ and $\underline{v_1}$: $\{\underline{u}, \underline{v_2}\}$ forms a basis.

Since $\underline{u} = 2\underline{v_1} + \underline{0}\underline{v_2}$, swapping $\underline{u}$ and $\underline{v}$ does NOT form a basis

Dimensions

> **Theorem**
>
> Every subspace $S \subseteq \mathbb{F}^n$ has a basis and every basis of S has the same number of elements

Proof: Providing a method to construct a basis and show this method terminates after finitely many steps

1) CASE 1: $S = \{\underline{0}\}$ then S has basis $\emptyset$

2) CASE 2: If $S \neq \{\underline{0}\}$ then take $\underline{v_1} \neq \underline{0}$ and the following steps

Step 1: If $S = Sp(\{\underline{v_1}\})$, then we are done

Else if $S \neq Sp(\{\underline{v_1}\})$ then $S \supsetneq Sp(\{\underline{v_1}\})$ then take

$\underline{v_2} \in S \setminus Sp(\{\underline{v_1}\})$ (So $\underline{v_2}$ is independent of $\underline{v_1}$)

Step 2: If $S = Sp(\{\underline{v_1}, \underline{v_2}\})$ then $\underline{v_1}$ and $\underline{v_2}$ is a basis of S.

Else if $S \neq Sp(\{\underline{v_1}, \underline{v_2}\})$, then $S \supsetneq Sp(\{\underline{v_1}, \underline{v_2}\})$ then take

$\underline{v_3} \in S \setminus Sp(\{\underline{v_1}, \underline{v_2}\})$ (So $\underline{v_3}$ is independent of $\underline{v_1}, \underline{v_2}$)

<u>Step K</u>: If $S = Sp(\{\underline{v}_1, \dots, \underline{v}_k\})$ then $\underline{v}_1, \dots, \underline{v}_k$ is a basis of S

and $\underline{v}_1, \dots, \underline{v}_k$ are linearly independent

If not, then $S \neq Sp(\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\})$ then $S \supsetneq Sp(\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\})$ then take

$$\underline{v}_{k+1} \in S \setminus Sp(\{\underline{v}_1, \dots, \underline{v}_k\})$$

Then $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k, \underline{v}_{k+1}\}$ is linearly independent

<u>Claim</u>: This algorithm stops after $\leq n$ steps

<u>proof</u>: (via contradiction)

Suppose we have made n steps and we have n linearly independent vectors

$$\underline{v}_1, \dots, \underline{v}_n$$

and consider $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$.

Suppose the procedure does not stop

Apply iteratively Steinitz Exchange Lemma replacing $\underline{e}_j$ with $\underline{v}_i$ for some i, j

I) If we can apply Steinitz Exchange Lemma n times then we will get that

$$\underline{v}_1, \dots, \underline{v}_n \text{ is a basis of } \mathbb{F}^n$$

Indeed after the first application, we get

$$\underline{e}_1, \dots, \underline{e}_{j-1}, \underline{v}_i, \underline{e}_{j+1}, \dots, \underline{e}_n \text{ is a basis}$$

If after n steps we get

$$\underline{v}_1, \dots, \underline{v}_n, \text{ then } \mathbb{F}^n = Sp(\underline{v}_1, \dots, \underline{v}_n).$$

So in our procedure, necessarily $S = Sp(\{\underline{v}_1, \dots, \underline{v}_n\})$

II) Assume that after k steps we cannot swap $\underline{v}_{k+1}$ with $\underline{e}_{k+1}$

At this step, we have basis $\underline{v}_1, \dots, \underline{v}_k, \underline{e}_{k+1}, \dots, \underline{e}_n$

Consider $\underline{v}_{k+1} = \alpha_1 \underline{v}_1 + \dots + \alpha_k \underline{v}_k + \alpha_{k+1} \underline{e}_{k+1} + \alpha_n \underline{e}_n$

If $\underline{v}_{k+1}$ cannot be swapped with $\underline{e}_{k+1}, \dots, \underline{e}_n$ to get a basis, we necessarily have

$$\alpha_{k+1} = \dots = \alpha_n = 0$$

So $\underline{v}_{k+1} = \alpha_1 \underline{v}_1 + \cdots + \alpha_k \underline{v}_k \in Sp(\{\underline{v}_1 \ldots, \underline{v}_k\})$

But $\underline{v}_{k+1} \in S \setminus Sp(\{\underline{v}_1, \ldots, \underline{v}_k\}) \Longrightarrow \underline{v}_{k+1} \notin Sp(\{\underline{v}_1, \ldots, \underline{v}_k\})$

This is a contradiction.

◀

Claim 2: Every basis for S has the same numbes of elements

proof: (by contradiction)

Assume that $k < l$ and

$\underline{u}_1, \ldots, \underline{u}_k$ and $\underline{v}_1, \ldots, \underline{v}_l$ are a basis for S.

(linearly independent)

Apply iteratively Steinitz exchange lemma to insert $\underline{u}_1, \ldots, \underline{u}_k$ into $\underline{v}_1, \ldots, \underline{v}_l$

Assume that for $t < k$, we cannot swap $\underline{u}_{t+1}$ with either of $\underline{v}_j$ in

$$\underline{u}_1, \ldots, \underline{u}_t, \underline{v}_{t+1}, \ldots, \underline{v}_l$$

Then consider

$$\underline{u}_{t+1} = \alpha \underline{v}_1 + \cdots + \alpha_t \underline{u}_t + \alpha_{t+1} \underline{v}_{t+1} + \cdots + \alpha_l \underline{v}_l$$

Then by Steinitz Exchange Lemma, $\alpha_{t+1} = \cdots = \alpha_l = 0$ hence

$$\underline{u}_{t+1} = \alpha_1 \underline{u}_1 + \cdots + \alpha_t \underline{u}_t$$

This contradicts linear independance of $\underline{u}_1, \cdots, \underline{u}_k$

Thus w.l.o.g $\underline{u}_1, \ldots, \underline{u}_k, \underline{v}_{k+1}, \cdots, \underline{v}_l$ is a basis of S

But $S = Sp(\{\underline{u}_1, \cdots, \underline{u}_k\})$. In particular

$$\underline{v}_{k+1} = \alpha_1 \underline{u}_1 + \cdots + \alpha_k \underline{u}_k$$

which contradicts linear independence.

Thus $k = l$.

■

Definition

For any subspace $S \subseteq \mathbb{F}^n$, we define dimension of S by

$$\dim(S) = \#(\text{basis of } S) \quad \text{cardinality}$$

Example:

1) $\mathbb{F}^n$ has standard basis $\{\underline{e}_1, \dots, \underline{e}_n\}$, hence

$$\dim(\mathbb{F}^n) = |\{\underline{e}_1, \dots, \underline{e}_n\}| = n$$

2) For $\mathbb{C}$, $\dim(\mathbb{C})$ depends on the ground field.

3) Consider solution set to homogeneous linear system

$$x_1 + x_2 + x_3 = 0$$

Claim that $\underline{v}_1 = (1,0,-1)$ and $\underline{v}_2 = (0,1,-1)$ spans S.

Clearly $\underline{v}_1, \underline{v}_2 \in S$ and they are linearly independent since

$$\alpha_1(1,0,-1) + \alpha_2(0,1,-1) = (0,0,0) \iff (\alpha_1, \alpha_2, -\alpha_1 - \alpha_2) = (0,0,0)$$
$$\iff \alpha_1 = 0,\ \alpha_2 = 0$$

Further every solution has form

$$(x_1, x_2, -x_1 - x_2) = x_1(1,0,-1) + x_2(0,1,-1)$$

so every solution belongs to $Sp(\{\underline{v}_1, \underline{v}_2\})$

Properties of dimensions and basis

Lemma

Suppose $S \subseteq \mathbb{F}^n$ is a linear subspace of $\mathbb{F}^n$ of dimension q

(0) Every linear independent set of vectors $\{\underline{u}_1, \dots, \underline{u}_t\} \subset S$ can be extended to a basis of S

(i) Any linearly independent subset $\mathcal{G}$ has no more than q elements

(ii) Any linearly independent subset $\mathcal{G} \subseteq \mathbb{F}^n$ can be extended to a basis of $\mathbb{F}^n$

(iii) Any finite spanning set for S contains a basis of S

Hence no subset containing fewer than q elements span S

(iv) Any linearly independent subset of S containing q elements spans S so it is a basis of S

Similarly if a set of size q spans S then it is linearly independent and its a basis.

(v) If $q=0$, then $S = \{\underline{0}\}$. If $q=n$, then $S = \mathbb{F}^n$

__Proof:__

(0) $\dim(S)=q$. Let $\underline{v}_1, \dots, \underline{v}_q$ be a basis of S.

Apply Steinitz Exchange Lemma recursively to $\underline{u}_1, \dots, \underline{u}_t$ and basis $\underline{v}_1, \dots, \underline{v}_q$ (then to $\underline{u}_1, \underline{v}_2, \dots, \underline{v}_q$)

So at k^{th} step ($k<t$), you want to exchange $\underline{u}_{k+1}$ with $\underline{v}_{k+1}, \dots, \underline{v}_q$ in basis

$$\underline{u}_1, \underline{u}_2, \dots, \underline{v}_{k+1}, \dots, \underline{v}_q$$

Since its a basis

$$\underline{u}_{k+1} = \alpha_1 \underline{u}_1 + \dots + \alpha_k \underline{u}_k + \alpha_{k+1} \underline{v}_{k+1} + \dots + \alpha_q \underline{v}_q$$

Claim: Not all $\alpha_{k+1}, \dots, \alpha_q$ are 0's

proof: (via contradiction)

Indeed if $\alpha_{k+1} = \dots = \alpha_q = 0$, then in particular (due to being a basis)

$$\underline{u}_{k+1} = \alpha_1 \underline{u}_1 + \dots + \alpha_k \underline{u}_k$$

which contradicts linear independence of $\underline{u}_1, \dots, \underline{u}_k$ ◀

So $\exists\, j \in [\![k+1, q]\!]$ such that $\alpha_j \neq 0$ hence by Steinitz Exchange Lemma,

$$\underline{u}_1, \dots \underline{u}_k, \underline{v}_{k+1}, \dots, \underline{v}_{j-1}, \underline{u}_{k+1}, \underline{v}_{j+1}, \dots, \underline{v}_q$$

is a basis.

Upto renumbering v_i's, without loss of generality, assume $j=k+1$

$$\underline{u}_1, \dots, \underline{u}_k, \underline{u}_{k+1}, \underline{v}_{k+2}, \dots, \underline{v}_q \text{ is a basis.}$$

i) By $\underline{u}_1, \dots, \underline{u}_t$ be linearly independent subset of S. By (0), it could be extended to a basis of S

$$\underline{u}_1, \dots, \underline{u}_t, \underline{v}_{t+1}, \dots, \underline{v}_q$$

so $t \leq q$

ii) This is (0) with $S = \mathbb{F}^n$

iii) Let $\mathcal{G} = \{\underline{u}_1, \dots, \underline{u}_t\}$ verify $Sp(\mathcal{G}) = S$

(a) If $\mathcal{G}$ is linearly independent $\Longrightarrow$ it is a basis

(b) If not, $\exists (\alpha_1, \dots, \alpha_t) \neq (0, \dots, 0)$ such that $\alpha_1 \underline{u} + \dots + \alpha_t \underline{u}_t = \underline{0}$

Without loss of generality, assume $\alpha_t \neq \underline{0}$

Then $\underline{u}_t = -\frac{\alpha_1}{\alpha_t}\underline{u}_1 - \cdots - \frac{\alpha_{t-1}}{\alpha_t}\underline{u}_{t-1}$

substitute

Claim: $Sp(\{\underline{u}_1, \dots, \underline{u}_r\}) = S, \quad r \leq t$

proof:

Indeed $\underline{v} \in S$, $\underline{v} = \beta_1 \underline{u}_1 + \dots + \beta_{t-1}\underline{u}_{t-1} + \beta_t \underline{u}_t$

$$= \left(\beta_1 - \beta_t \frac{\alpha_1}{\alpha_t}\right)\underline{u}_1 + \cdots + \left(\beta_{t-1} - \beta_t \frac{\alpha_{t-1}}{\alpha_t}\right)\underline{u}_{t-1} \Longrightarrow Sp(\{\underline{u}_1, \dots, \underline{u}_{t-1}\}) = S$$

If $\{\underline{u}_1, \dots, \underline{u}_{t-1}\}$ is linearly independent, done

If not repeat steps. Iterating this procedure, we arrive after $\leq t$ steps, we arrive to the linearly independent set

(basis) $\underline{u}_1, \dots, \underline{u}_r$ such that $Sp(\{\underline{u}_1, \dots, \underline{u}_r\}) = S$

iv) $\dim(S) = q \Longrightarrow \exists$ a basis $\underline{u}_1, \dots, \underline{u}_q$

a) If $\underline{u}_1, \dots, \underline{u}_q$ is linearly independent

Incase if $Sp(\{\underline{u}_1, \dots, \underline{u}_q\}) \neq S$ then complete this set to a basis of S

$$\underline{u}_1, \dots, \underline{u}_q, \underline{u}_{q+1} \cdots \underline{u}_{q+s}$$

But then we have a basis with $q+s > q \Longrightarrow$ contradicts theorem that all basis have same number elements

b) Assume $Sp(\{\underline{u}_1, \dots, \underline{u}_q\}) = S$. By (iii), a subset of $\underline{u}_1, \dots, \underline{u}_q$ is a basis of S

But by Thm above this basis has q elements so $\underline{u}_1, \dots, \underline{u}_q$ is a basis

(v) a) $q = 0 \Rightarrow$ basis $\emptyset \Rightarrow S = Sp(\emptyset) = \{\underline{0}\}$

b) $q = n$, let $\underline{v}_1, \dots, \underline{v}_n$ be a basis of S. By (iii), this set can be extended to a basis

$$B = \{\underline{v}_1, \dots, \underline{v}_n, \underline{v}_{n+1}, \dots\} \text{ of } \mathbb{F}^n$$

But if B is strictly larger than $\underline{v}_1, \dots, \underline{v}_n$, then we have a basis of $\mathbb{F}^n$ with $> n$ elements ↯

Indeed $\mathbb{F}^n$ has basis

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \Longrightarrow \dim \mathbb{F}^n = n$$

$\blacksquare$

Example:

1) Consider the 3 vectors

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

As shown before these are not linearly independent.

Theorem

Let $A\underline{x}=\underline{y}$ be a linear system of p equations in n variables.

If its set of solutions is not empty then every solution has form

$$\underline{x} = \alpha_1 \underline{v}_1 + \cdots + \alpha_q \underline{v}_q + \underline{x}_0, \quad \alpha_1, \cdots, \alpha_q \in \mathbb{F} \qquad (*)$$

where $\{\underline{v}_1, \cdots, \underline{v}_q\}$ is a basis for the solution set of $A\underline{x} = \underline{0}$ and

$\underline{x}_0$ is the particular solution to $A\underline{x}=\underline{y}$

The expression $(*)$ is known as $(*)$ is called the general solution to the system of equations

Example:

Consider

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_2 - x_3 = 0 \end{cases} \qquad A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \underline{y} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Particular solution

$$\underline{x}_0 = (1,0,0)$$

Via gaussian elimination

$$\underline{w} = \alpha_1 (2,-1,1)$$

So $\underline{x} = \alpha_1(2,-1,1) + (1,0,0) = (2\alpha_1 + 1, -\alpha_1, -\alpha_1)$, one degree of freedom, choice of α_1

Another particular is $\underline{x}_0 = (-1,1,1)$ and if $\alpha = \alpha_1/2$,

$$\underline{w} = \alpha(-2,1,1)$$

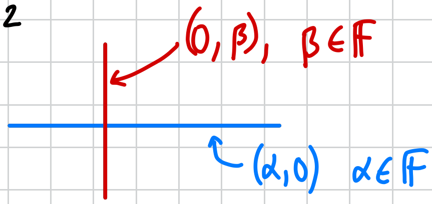
So $\underline{x} = \alpha(-2,1,1) + (-1,1,1) = (-1-2\alpha, \alpha+1, \alpha+1)$

SUM AND DIRECT SUM OF SUBSPACES

We want to construct a subspace from subspaces.

Let S_1 and S_2 be 2 subspaces of $\mathbb{F}^n$

<u>Problem</u> $S_1 \cup S_2$



$(\alpha, 0) + (0, \beta) = (\alpha, \beta)$

$S_1 = \{(\alpha, 0) : \alpha \in \mathbb{R}\}$ $\quad$ $S_2 = \{(0, \beta) : \beta \in \mathbb{R}\}$

$S_1 \cap S_2 = \{(0,0)\}$ is a subspace

$S_1 \cup S_2$ is NOT a subspace

Direct Sums

Definition Sum of Subspaces

Let $S_1, \dots, S_q \subset \mathbb{F}^n$ be subspaces. Then, sum

$$S_1 + S_2 + \dots + S_q = Sp(S_1 \cup \dots \cup S_q) = \{\alpha_1 \underline{v}_1 + \dots + \alpha_q \underline{v}_q \mid \alpha_j \in \mathbb{F},\ \underline{v}_j \in S_j\}$$

When

$$S_j \cap \left(\sum_{k \neq j} S_k \right) = \{\underline{0}\} \qquad \forall\ 1 \leq j \leq q$$

we call this the direct sum denoted

$$S_1 \oplus S_2 \oplus \dots \oplus S_q = \bigoplus_{j=1}^{q} S_j$$

Theorem

For any subspaces $S_1, \dots, S_q \in \mathbb{F}^n$

i) $S_1 \cap \dots \cap S_q$ is a subspace

(ii) $S_1 + \dots + S_q$ is a subspace

<u>Proof</u>

ii) Span(anything) is always a subspace $\Longrightarrow$ $S_1 + \dots + S_q = Sp(S_1 \cup \dots S_q)$ is a subspace

i) $\{\underline{0}\} \in S_k \ \forall k=1,\dots,q \Longrightarrow \{\underline{0}\} \in S_1 \cap \dots \cap S_q$

$\underline{v}_1,\dots,\underline{v}_t \in S_1 \cap \dots \cap S_q \iff \forall 1,\dots,t \ \forall j=1,\dots,q \ , \ \underline{v}_i \in S_j$

So $\underline{v}_1,\dots,\underline{v}_q \in S_j \ , \ j=1,\dots,q$

S_j is a subspace, $\forall \alpha_1,\dots,\alpha_t, \ \alpha_1\underline{v}_1+\dots+\alpha_t\underline{v}_t \in S_j \quad \forall j=1,\dots,q$

$\Longrightarrow \alpha_1\underline{v}_1+\dots+\alpha_q\underline{v}_q \in S_1 \cap \dots \cap S_q \Longrightarrow S_1 \cap \dots \cap S_q$ is a subspace. ■

Example:

1) Let $S_1 \subseteq \mathbb{F}^n$ be defined by $\sum_{j=1}^{n} \alpha_j \underline{x}_j = \underline{0}$

$S_2 \subseteq \mathbb{F}^n$ be defined by $\sum_{j=1}^{n} \beta_j \underline{x}_j = \underline{0}$

Then $S_1 \cap S_2$ is defined by $\begin{cases} \sum_{j=1}^{q} \alpha_j \underline{x}_j = 0 \\ \sum_{j=1}^{n} \beta_j \underline{x}_j = \underline{0} \end{cases}$

2) $\mathbb{F}^3$ and standard basis vectors

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \underline{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Let $S_j = Sp(\underline{e}_j)$: each of these is a line we call an axis when $\mathbb{F}=\mathbb{R}$

Notice that $S_j \cap S_k = \{\underline{0}\}$ for $j \neq k$

$S_1 \oplus S_2 = Sp(\underline{e}_1, \underline{e}_2)$ is the plane defined by $x_3 = 0$

3) $\mathbb{F}^3$, $V_1 = Sp(\underline{e}_1, \underline{e}_2) = \{(\alpha,\beta,0) \mid \alpha,\beta \in \mathbb{F}\}$

$V_2 = Sp(\underline{e}_2, \underline{e}_3)$

$V_1 + V_2 = Sp(V_1 \cup V_2) = Sp(\underline{e}_1, \underline{e}_2, \underline{e}_3) = \mathbb{F}^3$

$V_1 \cap V_2 = Sp(\underline{e}_2)$

Lemma

Let S_1, S_2 be subspaces of $\mathbb{F}^n$. Then

$$\dim(S_1+S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2)$$

In particular for direct sum

$$\dim(S_1 \oplus S_2) = \dim(S_1) + \dim(S_2)$$

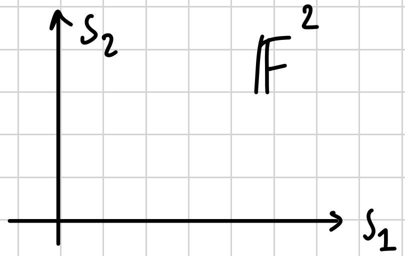
Example

$\dim V_1 = 2$

$\dim V_2 = 2$

$\dim(V_1 + V_2) = 3$

$\dim(V_1 \cap V_2) = 1$



$\dim(S_1) = 1$

$\dim(S_2) = 1$

$\dim(S_1 \cap S_1) = 0$

$\dim(S_1 \cup S_2) = 2$

Lemma

Let $S_1 \oplus S_2 \oplus \cdots \oplus S_q$ be a direct sum of subspaces and

$$\underline{v}_j \in S_j \setminus \{\underline{0}\} \text{ (non-zero) for } j = 1, \dots, q$$

Then $\underline{v}_1, \dots, \underline{v}_q$ are linearly independent

Proof:

Assume $\sum_{j=1}^{q} \alpha_j \underline{v}_j = \underline{0} \quad \forall j = 1, \dots, q$

$$\alpha_j \underline{v}_j = -\sum_{\substack{k=1 \\ k \neq j}}^{q} \alpha_k \underline{v}_k \in (S_1 \oplus \cdots S_{j-1} \oplus S_{j+1} \oplus \cdots \oplus S_q) \cap S_j = \{\underline{0}\}$$

$$\Longrightarrow \alpha_j \underline{v}_j = \underline{0}$$

$$\Longrightarrow \alpha_j = 0 \quad \forall j = 1, \dots, q$$

∎

2. Matrices and Linear Maps

LINEAR MAPS

Definition Linear Maps

A map $L:\mathbb{F}^n \longrightarrow \mathbb{F}^p$ is called linear map if

$$L(\alpha \underline{u} + \beta \underline{v}) = \alpha L(\underline{u}) + \beta L(\underline{v}) \qquad \forall \alpha, \beta \in \mathbb{F}, \ \forall \underline{u}, \underline{v} \in \mathbb{F}$$

Example:

Let A be a $n \times p$ matrix. Then

$$A(\alpha \underline{u} + \beta \underline{v}) = \alpha A \underline{u} + \beta A \underline{v}$$

$\therefore$ $\underline{u} \longmapsto A\underline{u}$ is a linear map

Lemma

A map $L:\mathbb{F}^n \longrightarrow \mathbb{F}^p$ is a linear map if and only if

$$\exists A \in Mat_{p \times n} \text{ s.t } L(\underline{u}) = A\underline{u}$$

Proof:

$(\Longrightarrow)$: Consider

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \ \cdots\cdots \ \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

and define $A = (L(\underline{e}_1) \cdots L(\underline{e}_n)) \in Mat_{p \times n}$ that is columns of A are vectors $L(\underline{e}_i), i = 1, \cdots, n$

Lets verify that $\forall \underline{u} \in \mathbb{F}^n$, we have $L(\underline{u}) = A\underline{u}$

We have $\underline{u} = \alpha_1 \underline{e}_1 + \cdots + \alpha_n \underline{e}_n = \sum_{i=1}^{n} \alpha_j \underline{e}_j$

$$L(\underline{u}) = L(\alpha_1 \underline{e}_1 + \cdots + \alpha_n \underline{e}_n) = \alpha_1 L(\underline{e}_1) + \cdots + \alpha_n L(\underline{e}_n) = (L(\underline{e}_1), \cdots, L(\underline{e}_n)) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = A\underline{u}$$

Hence $\boxed{A = (L(\underline{e}_1) \cdots L(\underline{e}_n))}$ To find A ■

Remark:

We distinguish between matrices and linear maps

For example, the linear map $L:\mathbb{R}^2 \longrightarrow \mathbb{R}^3$; $L(x_1, x_2) = (x_1, x_2, 0)$ is represented by matrix

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

Remark If M is a matrix with real co-efficients, then, it defines a linear map

$$\mathbb{R}^n \to \mathbb{R}^p$$

but also a linear map

$$\mathbb{C}^n \to \mathbb{C}^p$$

The converse NOT true

Further any linear map $L: \mathbb{R}^n \to \mathbb{R}^p$ has a natural extension to a linear map $L: \mathbb{C}^n \to \mathbb{C}^p$
The matrix just not change, just using the fact $\mathbb{R} \subseteq \mathbb{C}$.

Converse NOT true, for example

$L: \mathbb{C}^2 \to \mathbb{C}$; $L(x_1, x_2) = ix_1$ is represented by

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

clearly does not map $\mathbb{R}^2$ into $\mathbb{R}$

Remark

Let M and N be linear maps with the corresponding matrices A and B. Then

$\alpha L + \beta M$: $(\alpha L + \beta M)(\underline{u}) = \alpha L(\underline{u}) + \beta M(\underline{u})$ is a linear map with

$$\alpha A + \beta B$$

Recap: Multiplication of matrices

$$A(A_{ij}) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{np} \end{pmatrix} \qquad B(B_{ij}) = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1q} \\ B_{21} & B_{22} & \cdots & B_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ B_{1p} & B_{2p} & \cdots & B_{pq} \end{pmatrix}$$

$$\boxed{AB = C_{ij} \quad \text{where} \quad C_{ij} = \sum_{k=1}^{p} A_{ik} B_{kj}}$$

$A \in \mathbb{R}^{np}$ $\quad B \in \mathbb{R}^{pq}$

$AB \in \mathbb{R}^{n \times q}$

Proposition

Matrix multiplication satisfies the following properties

if $A \in \mathbb{R}^{m \times n}$ and $B, C \in \mathbb{R}^{n \times p}$ then

$$A(B+C) = AB + BC$$

and if $A, B \in \mathbb{R}^{m \times n}$ and $C \in \mathbb{R}^{n \times p}$ then

$$(A+B)C = AC + BC$$

Proposition

i) If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ and $r, s \in \mathbb{R}$ then

$$(rA)(sB) = rs(AB)$$

ii) If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$ and $C \in \mathbb{R}^{p \times q}$ then

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

The definition above is compatible with matrix multiplication

$$A\underline{x}$$

of a matrix $n \times p$ by a vector $\underline{x} \in \mathbb{F}^p$ if you consider a vector as $p \times 1$ matrix

In particular for matrices A and B and a vector $\underline{x}$ of appropriate size, we have

$$(AB)\underline{x} = A(B\underline{x})$$

$$\mathbb{F}^n \xrightarrow{L} \mathbb{F}^p \xrightarrow{M} \mathbb{F}^q$$

$$\forall \underline{x} \in \mathbb{F}^n, \quad L\underline{x} = B\underline{x} \implies M(L(\underline{x})) = A(B\underline{x})$$

$$\implies (M \circ L)(\underline{x}) = A(B\underline{x}) = (AB)\underline{x}$$

composition associativity

Lemma

Let $L : \mathbb{F}^n \longrightarrow \mathbb{F}^p$ and $M : \mathbb{F}^p \longrightarrow \mathbb{F}^q$ be 2 linear maps represented by

$A \in \mathrm{Mat}_{p \times n}(\mathbb{F})$ and $B \in \mathrm{Mat}_{q \times p}(\mathbb{F})$ respectively.

Then

$$M \circ L \text{ is linear}$$

represented by

$$BA$$

IMAGES AND KERNEL; RANK AND NULLITY

Images

We can use linear maps to rephrase the problem of existence and uniqueness for linear system of equations

To a system of p linear equations in n unknowns, $A\underline{x}=\underline{y}$, we assign the linear map

$$L: \mathbb{F}^n \to \mathbb{F}^p; \quad L(\underline{x}) = A\underline{x}$$

Definition Image and Kernel

Let L be a Linear map from $\mathbb{F}^n$ to $\mathbb{F}^p$; $L: \mathbb{F}^n \longrightarrow \mathbb{F}^p$

Image of L: $\operatorname{Im}(L) = \{\underline{y} \in \mathbb{F}^p \mid \underline{y} = L(\underline{x}) \text{ for some } \underline{x} \in \mathbb{F}^n\}$

Kernel of L: $\operatorname{Ker}(L) = \{\underline{x} \in \mathbb{F}^n \mid L(\underline{x}) = \underline{0}\}$ also called null-space

Lemma

Suppose L is a linear map $L: \mathbb{F}^n \longrightarrow \mathbb{F}^p$

- $\operatorname{Im}(L)$ is a subspace of $\mathbb{F}^p$
- $\operatorname{Ker}(L)$ is a subspace of $\mathbb{F}^n$

Further when L is represented by a matrix $A \in \operatorname{Mat}_{p\times n}(\mathbb{F})$

$\operatorname{Im}(L)$ is the subspace spanned by the columns of A

Proof:

- $\underline{0} \in L(\underline{0}) \implies L(\underline{0}) = \underline{0} \in \operatorname{Im}(L)$

For any $\underline{u}, \underline{v} \in \operatorname{Im}(L)$, then $\exists\ \underline{x}, \underline{y} \in \mathbb{F}^n$ s.t $\underline{u} = L(\underline{x})$ and $\underline{v} = L(\underline{y})$

$$\alpha\underline{u} + \beta\underline{v} = \alpha L(\underline{x}) + \beta L(\underline{y}) = L(\alpha\underline{x} + \beta\underline{y}) \in \operatorname{Im}(L)$$

$$\implies \text{subspace}$$

- $\underline{0} \in \operatorname{Ker}(L)$ as $L(\underline{0}) = \underline{0}$

$\forall\ \underline{u}, \underline{v} \in \operatorname{Ker}(L)$, $\forall \alpha, \beta \in \mathbb{F}$

$$L(\alpha\underline{u} + \beta\underline{v}) = \alpha L(\underline{u}) + \beta L(\underline{v}) = 0 \implies \alpha\underline{u} + \beta\underline{v} \in \operatorname{Ker}(L)$$

$$\implies \text{subspace}$$

Now recall from above that columns of A are $L(\underline{e}_1), \cdots, L(\underline{e}_n)$

$$A = (L(\underline{e}_1) \cdots L(\underline{e}_n))$$

Now if $v \in \mathrm{Im}(L)$ if and only if $\underline{v} = L(\underline{u})$ for some $\underline{u} \in \mathbb{F}^n$ and we can write

$$\underline{u} = \sum_{j=1}^{n} \alpha_j \underline{e}_j \quad \text{for some } \alpha_j \in \mathbb{F}$$

Linearity gives

$$L\left(\sum \alpha_j \underline{e}_j\right) = \sum \alpha_j L(\underline{e}_j) \in Sp(L(\underline{e}_1), \cdots, L(\underline{e}_n))$$

and we conclude

$$\mathrm{Im}(L) = Sp(L(\underline{e}_1), \ldots, L(\underline{e}_n))$$

■

Notation: Some additional notation for image and kernel

* $\mathrm{Im}(L) = \{L(\underline{a}) : a \in \mathbb{F}^n\} \subseteq \mathbb{F}^p$
* $\mathrm{Ker}(L) = \{\underline{a} = \mathbb{F}^n \mid L(\underline{a}) = \underline{0}\}$

Definition Rank/Nullity

Let L be a linear map.

Rank of L, $rk(L)$ is the dimension of $\mathrm{Im}(L)$

Nullity of L, $null(L)$ is the dimension of $\ker(L)$

Remark:

Let A be the matrix that represents L

By Lemma 2.4, $\mathrm{Im}(L) = Sp(\text{columns of } A)$

So $\dim L$ = maximal number of linearly independent columns of A.

Fact: $rkA = rkA^T$. That is maximal number of linearly independent columns

= maximal number of linearly independent rows

Theorem Rank-Nullity Theorem

For a linear map $L: \mathbb{F}^n \to \mathbb{F}^p$

$$n = rk(L) + null(L)$$

Proof: Consider $Ker L$ and $Im L$

Since $Ker L$ is a subspace, let

$$B = \{\underline{u}_1, \cdots, \underline{u}_q\} \text{ be a basis of } Ker L \Longrightarrow \dim ker(L) = null(L) = q$$

By Lemma 1.12 (ii) we can extend basis B to a basis of $\mathbb{F}^n$:

$$\{\underline{u}_1, \cdots, \underline{u}_q, \underline{v}_1, \cdots, \underline{v}_r\} \qquad (n = q + r)$$

We are going to show that $L(\underline{v}_1), \ldots, L(\underline{v}_r)$ is a basis of $Im(L)$

(i) Linear Independence: Let $\alpha_1 L(\underline{v}_1) + \cdots + \alpha_r L(\underline{v}_r) = \underline{0}$

$$\Longleftrightarrow L(\alpha_1 \underline{v}_1 + \cdots + \alpha_r \underline{v}_r) = \underline{0} \Longleftrightarrow \alpha_1 \underline{v}_1 + \cdots + \alpha_r \underline{v}_r \in ker(L).$$

Since B is a basis for $Ker L$

$$\alpha_1 \underline{v}_1 + \cdots + \alpha_r \underline{v}_r = \beta_1 \underline{u}_1 + \cdots + \beta_q \underline{u}_q \Longleftrightarrow -\beta_1 \underline{u}_1 - \cdots - \beta_q \underline{u}_q + \alpha_1 \underline{v}_1 + \cdots + \alpha_r \underline{v}_r = \underline{0}$$

But $\{\underline{u}_1, \cdots, \underline{u}_q, \underline{v}_1, \cdots, \underline{v}_r\}$ is a basis so these vectors are linearly independent

$$\Longrightarrow \alpha_1 = \cdots = \alpha_r = 0$$

$\Longrightarrow L(\underline{v}_1) \cdots L(\underline{v}_r)$ are linearly independent

(ii) $Im L = Sp(L(\underline{v}_1), \cdots L(\underline{v}_r))$: Let $\underline{u} \in Im(L)$. Then $\exists\, \underline{v} \in \mathbb{F}^n$ such that

$$\underline{u} = L(\underline{v})$$

Since B is a basis for $\mathbb{F}^n$

$$\underline{v} = \beta_1 \underline{u}_1 + \cdots + \beta_q \underline{u}_q + \alpha_1 \underline{v}_1 + \cdots + \alpha_r \underline{v}_r$$

Therefore we have

$$\begin{aligned}
\underline{u} = L(\underline{v}) &= L(\beta_1 \underline{u}_1 + \cdots + \beta_q \underline{u}_q + \alpha_1 \underline{v}_1 + \cdots + \alpha_r \underline{v}_r) \\
&= \beta_1 \cancel{L(\underline{u}_1)} + \cdots + \beta_q \cancel{L(\underline{u}_q)} + \alpha_1 L(\underline{v}_1) + \cdots + \alpha_r L(\underline{v}_r) \\
&= \alpha_1 L(\underline{v}_1) + \cdots + \alpha_r L(\underline{v}_r) \in Sp(L(\underline{v}_1), \ldots, L(\underline{v}_r))
\end{aligned}$$

$$\Longrightarrow Im(L) = Sp(L(\underline{v}_1), \cdots, L(\underline{v}_r))$$

By (i) and (ii),

$$L(\underline{v}_1), \cdots, L(\underline{v}_r) \text{ is a basis for } \mathrm{Im}\, L \implies \dim \mathrm{Im}(L) = \mathrm{rk}\, L = r$$

Therefore $n = q + r = \mathrm{null}(L) + \mathrm{rk}(L)$ ∎

Reminder: Let $L: \mathbb{F}^n \longrightarrow \mathbb{F}^p$ be a linear map

- L is one-to-one (injective) if $L(\underline{u}_1) = L(\underline{u}_2) \implies \underline{u}_1 = \underline{u}_2$
- L is onto (surjective) if $\forall\, \underline{a} \in \mathbb{F}^p\ \exists\, \underline{b} \in \mathbb{F}^n$ s.t $L(\underline{b}) = \underline{a}$ $(\mathrm{Im}(L) = \mathbb{F}^p)$
- L is bijective if f is both one to one and onto

Lemma

A linear map $L: \mathbb{F}^n \to \mathbb{F}^p$ is

i) one to one $\iff \mathrm{Ker}(L) = \{\underline{0}\} \iff \mathrm{null}(L) = 0$

ii) onto $\iff \mathrm{rk}(L) = p$

iii) bijective $\iff \mathrm{null}(L) = 0$ and $n = p$

Proof:

i) $(\Leftarrow)$: $\ker L = \{\underline{0}\}$. Then if $L(\underline{u}) = L(\underline{v}) \iff L(\underline{u} - \underline{v}) = \underline{0}$

$$\iff \underline{u} - \underline{v} \in \ker L$$

$$\iff \underline{u} - \underline{v} = \underline{0}$$

$$\iff \underline{u} = \underline{v}$$

$(\Rightarrow)$: L is 1 to 1. Since $L(\underline{0}) = \underline{0}$

$\underline{0} \in \ker L$. Then $\forall\, \underline{u} \in \ker L$, $L\underline{u} = \underline{0}_p = L\underline{0}_n \implies \underline{u} = \underline{0}_n$

ii) $\mathrm{rk}(L) = p = \dim \mathrm{Im}(L) \iff \dim \mathrm{Im}(L) = p$. But the only subspace of $\mathbb{F}^p$ of dimension p is $\mathbb{F}^p$. So

$$\dim \mathrm{Im}(L) = p \iff \mathrm{Im}\, L = \mathbb{F}^p$$

iii) By rank-nullity theorem,

$$(\text{onto})\ p = n - \mathrm{null}\, L = n\ (\text{one-to-one})$$

∎

Corollary

A system of $L(\underline{x}) = \underline{y}$ has a solution $\iff$ $\underline{y} \in \operatorname{Im}(L)$

When it has a solution, it is unique $\iff$ $\operatorname{Ker} L = 0$ (one-to-one)

Proof:

$\exists$ a solution $\underline{x} \iff L(\underline{x}) = \underline{y} \iff \underline{y} \in \operatorname{Im} L$

A solution is unique $\iff$ We can have $\begin{cases} L(\underline{x}_1) = \underline{y} \\ L(\underline{x}_2) = \underline{y} \end{cases} \iff \underline{x}_1 = \underline{x}_2$

$(\Rightarrow)$: Contrapositive: We are going to prove that $\operatorname{Ker} L \neq \{\underline{0}\}$ then there is more than one solution

Let $\underline{u} \in \operatorname{Ker} L \setminus \{\underline{0}\}$, that is $L(\underline{u}) = \underline{0}$ and $\underline{u} \neq \underline{0}$

Let $\underline{x}_0$ be a solution $\Longrightarrow L(\underline{x}_0) = \underline{y}$

Then $\underline{x}_0 + \underline{u} \neq \underline{x}_0$ and $L(\underline{x}_0 + \underline{u}) = L(\underline{x}_0) + L(\underline{u}) = L(\underline{x}_0) = \underline{y}$

$(\Leftarrow)$: Contropositive: Lets show that if $\underline{x}_1 \neq \underline{x}_2$ are both solutions to $L(\underline{x}) = \underline{y}$ then $\operatorname{Ker} L \neq \{\underline{0}\}$

Indeed $\underline{x}_1 - \underline{x}_2 \neq \underline{0}$ and $L(\underline{x}_1 - \underline{x}_2) = L(\underline{x}_1) - L(\underline{x}_2) = \underline{y} - \underline{y} = \underline{0}$

$$\Longrightarrow \underline{x}_1 - \underline{x}_2 \in \operatorname{Ker} L \setminus \{\underline{0}\}$$

$\blacksquare$

Remark: By the corollary above, the uniqueness the uniqueness of solution to $L\underline{x} = \underline{y}$ depends on L only (not on $\underline{y}$)

This is not a case for general non-linear system

Counterexamples: 1) $F: \mathbb{R}^2 \to \mathbb{R}^2$

$$(x_1, x_2) \mapsto (x_1^2 + x_2^2, x_2)$$

Consider equation $F(x_1, x_2) = (1, a) \iff \begin{cases} x_1^2 + x_2^2 = 1 \\ x_2 = 0 \end{cases}$

If $|a| > 1$ then there is no solution

If $|a| = 1$, then a unique solution $(0, a)$

If $|a| < 1$, then $\exists$ 2 solutions $(\pm\sqrt{1-a^2}, a)$

2) Consider $g(x_1, x_2) = (x_1 x_2 + 1, x_2)$ and $g(x_1, x_2) = (1, a)$

If $a \neq 0$, a unique solution $\left(\frac{a-1}{a}, a\right)$

If $a = 0$, infinitely many solutions $(b, 0)$, $b \in \mathbb{R}$

Corollary

For a homogeneous linear system of p equations in n unknowns,

$$A\underline{x} = \underline{0}$$

the number of linearly independent solutions equal $n - \text{rk} A$

Proof:

The number of linearly independent solutions $= \dim \ker A = \text{null} A$, $\text{rk} A = \dim A$.

By rank nullity theorem

$$\text{rk} A + \text{null} A = n \implies \text{null} A = n - \text{rk} A$$

$\blacksquare$

Fact: By using the fact that the # of linearly independent columns of A = # of linearly indp rows of A,

The number of linearly independent solutions $= n - r$

r is the number of linearly independent equations

INVERTIBLE LINEAR MAPS; CHANGE OF BASIS

Lemma

If $L: \mathbb{F}^n \to \mathbb{F}^n$ is an invertible linear map, represented by an $n \times n$ matrix A, then

$$L^{-1} \text{ is linear}$$

and is represented by A^{-1}

Recall: Invertible map $L: A \to B \iff \exists L^{-1}: B \to A$ s.t

$$L \circ L^{-1} = id_A \quad \text{and} \quad L^{-1} \circ L = id_A$$

Inverse matrix: Inverse to a matrix A is a matrix A^{-1} such that

$$A \cdot A^{-1} = A^{-1} A = I_n$$

Proof: We want to show that $\forall \underline{u}, \underline{v} \in \mathbb{F}^n$ (codomain) and $\forall \alpha, \beta \in \mathbb{F}$,

$$L^{-1}(\alpha \underline{u} + \beta \underline{v}) = \alpha L^{-1}(\underline{u}) + \beta L^{-1}(\underline{v})$$

L is invertible, in particular, 1-to-one, enough to show

$$L(lhs) = L(rhs)$$

$$L(L^{-1}(\alpha \underline{u} + \beta \underline{v})) = \alpha \underline{u} + \beta \underline{v}$$

$$L(\alpha L^{-1}(\underline{u}) + \beta L^{-1}(\underline{v})) = \alpha L(L^{-1}(\underline{u})) + \beta L(L^{-1}(\underline{v})) = \alpha \underline{u} + \beta \underline{v} \implies L(lhs) = L(rhs)$$

Let L^{-1} be represented by a matrix B

$L \cdot L^{-1}$ is represented by BA and $L \circ L^{-1} = I$, so

$I = BA$ and similarly $I = BA \implies B = A^{-1}$ ■

Lemma

A basis $\{\underline{v}_1, \cdots, \underline{v}_n\}$ of $\mathbb{F}^n$. A linear map

$$L: \mathbb{F}^n \to \mathbb{F}^n$$

is invertible $\iff L(\underline{v}_1), \cdots, L(\underline{v}_n)$ is a basis

In particular an $n \times n$ matrix is invertible $\iff$ its columns provide a basis of $\mathbb{F}^n$

proof: Suppose L is invertible.

Lets check $L(\underline{v}_1), \cdots, L(\underline{v}_n)$ is linearly independent, then they span $\mathbb{F}^n$ by lemma 1.12 hence they form a basis

Assume $\alpha L(\underline{v}_1) + \cdots + \alpha_n L(\underline{v}_n) = \underline{0}$

We know by Lemma 2.8, L^{-1} is linear. Then

$$\underline{0} = L^{-1}(\underline{0}) = L^{-1}(\alpha_1 L(\underline{v}_1) + \cdots + \alpha_n L(\underline{v}_n))$$

$$= \alpha \underline{v}_1 + \cdots + \alpha_n \underline{v}_n$$

But $\{\underline{v}_1, \cdots, \underline{v}_n\}$ is a basis $\implies \alpha_1 = \alpha_2 = \cdots \alpha_n = 0$

$\implies L(\underline{v}_1), \cdots, L(\underline{v}_n)$ are linearly independent

$\implies$ span $\mathbb{F}^n$ (lemma 1.12)

$\implies$ form a basis

($\Leftarrow$): Now assume $L(\underline{v}_1), \cdots, L(\underline{v}_n)$ form a basis of $\mathbb{F}^n$

By the rank-nullity theorem, it is enough to show that $\ker L = \{\underline{0}\}$

(because then $\operatorname{null} L = 0 \Rightarrow \operatorname{rk} L = n - 0 = n \Rightarrow \operatorname{Im} L = \mathbb{F}^n$

$\Rightarrow \{L(\underline{v}_1), \cdots, L(\underline{v}_n)\}$ spans $\mathbb{F}^n$

$\Rightarrow$ basis of $\mathbb{F}^n$ by lemma 1.12)

Let $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \ker L$. Then $L\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \underline{0} \iff \left(L(\underline{v}_1) \cdots L(\underline{v}_n)\right)\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \underline{0}$

$\overset{?}{\iff} \alpha_1 L(\underline{v}_1) + \cdots + \alpha_n L(\underline{v}_n) = \underline{0}$

$\Rightarrow \alpha_1 = \cdots = \alpha_n = 0 \Rightarrow \ker L = \{\underline{0}\}$

By rank nullity theorem $\operatorname{rk} L = n \iff \operatorname{Im} L = \mathbb{F}^n$

$\iff L$ is bijection (invertible) $\blacksquare$

Example:

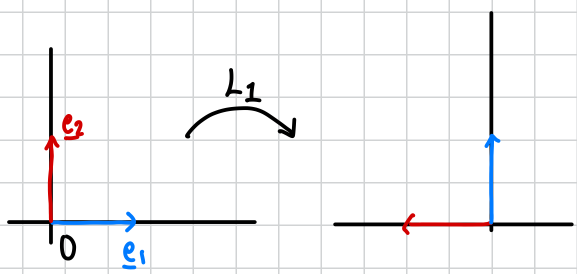
1) $A_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = (\underline{e}_2 \ \ -\underline{e}_1)$

It is the linear map represented by

$$L_1(x_1, x_2) = (-x_2, x_1)$$

$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\underline{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\underline{e}_1 \mapsto \underline{e}_2$

$\underline{e}_2 \mapsto -\underline{e}_1$



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

$\underline{e}_1 \mapsto \underline{e}_2$

$\underline{e}_2 \mapsto \underline{e}_1$

EIGENVECTORS AND EIGENVALUES

Notation:

$L: \mathbb{F}^n \to \mathbb{F}^n \quad (L: \mathbb{F}^n \circlearrowleft)$

Definition

A linear map $L: \mathbb{F}^n \circlearrowleft$

An eigenvector of L is a non-zero vector $\underline{v} \in \mathbb{F}^n$ such that

$$L\underline{v} = \lambda \underline{v} \quad \text{where} \quad \lambda \in \mathbb{F}$$

In this case λ is an eigenvalue of L

The same definition applicable to matrices

$$A\underline{v} = \lambda \underline{v}$$

The set of all eigenvalues of L is called the spectrum of L : Spec L

$$\text{Spec } L = \{\lambda \in \mathbb{F} \mid L - \lambda I_n \text{ is not invertible}\}$$

Indeed

$$L\underline{v} = \lambda \underline{v} \iff (L - \lambda Id_n)\underline{v} = 0$$

Example:

i) $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \lambda_1 \neq \lambda_2$

Then λ_1 and λ_2 are eigenvalues, the corresponding eigenvectors are $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

ii) If $\lambda_1 = \lambda_2$ in i) then we have the matrix $\lambda_1 I_2$ which has precisely one eigenvalue λ_1

iii) In $\mathbb{R}$ it is possible to have matrix with no eigenvalues

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

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix} \qquad \begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} -b = \lambda a \\ a = \lambda b \end{cases} \implies -b = \lambda^2 b \implies -1 = \lambda^2$$

iv) Matrix with one eigenvalue and one linearly independent eigenvector

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \longrightarrow \lambda = 2 \qquad \underline{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Remark:

Eigenvectors are never unique.

If $\underline{v}$ is an eigenvector, so is $\lambda \underline{v}$, $\lambda \in \mathbb{F}$

$$L\alpha\underline{v} = \lambda\alpha\underline{v}, \quad \underline{v} \neq 0, \lambda \in \mathbb{F}, \lambda \in \mathbb{F}$$

Remark: If $L\underline{v}_1 = \lambda \underline{v}_1$

$L\underline{v}_2 = \lambda \underline{v}_2$

$$\Rightarrow L(\underline{v}_1 + \underline{v}_2) = \lambda(\underline{v}_1 + \underline{v}_2)$$

$$\Rightarrow L\underline{v}_1 + L\underline{v}_2 = \lambda\underline{v}_1 + \lambda\underline{v}_2$$

Definition, Eigenspace

Given an eigenvalue λ of a linear map $L: \mathbb{F}^n \to \mathbb{F}^n$, we call

$$\ker(L - \lambda I_n) = \{\underline{v} \in \mathbb{F}^n \mid (L - \lambda I_n)\underline{v} = \underline{0}\}$$

the eigenspace for eigenvalue λ of L

$\dim \ker(L - \lambda I_n)$ is called the geometric multiplicity of λ

Note: $\underline{0}$ is not an eigenvector, eventhough it belongs to the eigenspace

Example:

$L: \mathbb{C}^2 \to \mathbb{C}^2$, $L_1(x_1, x_2) = (-x_2, x_1)$

L_1 represented by $A_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{cases} y = \lambda x \\ -x = \lambda y \end{cases} \Rightarrow -x = \lambda^2 x \Rightarrow \lambda^2 = -1$$

$$\Rightarrow \lambda = \pm i$$

1) $\lambda = i$ $\begin{cases} y = ix \\ -x = iy \end{cases} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ iy \end{pmatrix}$

2) $\lambda = -i$ $\begin{cases} y = -ix \\ -x = -iy \end{cases} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -ix \end{pmatrix}, \quad x \neq 0$

$$\ker(L - iI_2) = \ker\left(\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix}\right) = Sp\begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$\ker(L - (i)I_2) = \ker\left(\begin{pmatrix} +i & -1 \\ +1 & +i \end{pmatrix}\right) = Sp\begin{pmatrix} i \\ -1 \end{pmatrix}$$

as all linear combinations are eigenvectors

$\Rightarrow$ belong to $\ker(L + \lambda I_n)$

$\Rightarrow$ span.

Lemma

Let $\lambda_1, \dots, \lambda_q$ be distinct eigenvalues of $L: \mathbb{F}^n \circlearrowleft$

Then the corresponding eigenspaces form a direct sum

$$S_1 \oplus S_2 \oplus \cdots \oplus S_q$$

In particular eigenvectors $\underline{v}_1, \dots, \underline{v}_q$ for distinct eigenvalues are linearly independent

Proof: Reminder: $S_1 \oplus \cdots \oplus S_q$ mean that $\forall\, \ell = 1, \dots, q$, $S_\ell \cap \left(\sum_{k \neq \ell} S_k\right) = \{\underline{0}\}$

For example if $q=2$, we need to check $S_1 \cap S_2 = \{\underline{0}\}$

$\underline{v} \in S_1 \cap S_2$: $L\underline{v} = \lambda_1 \underline{v} = \lambda_2 \underline{v} \Longrightarrow (\lambda_1 - \lambda_2)\underline{v} = 0 \Longrightarrow \underline{v} = \underline{0}$

proof using induction

Inductive step

Assume that $S_1, \dots, S_{l-1}$ form a direct sum.

Consider $S_\ell \cap \left(\bigoplus_{k=1}^{\ell-1} S_k \right) \ni \underline{v}$

$\underline{v} \in S_\ell \Rightarrow L\underline{v} = \lambda_\ell \underline{v}$. Furthermore $v \in \bigoplus_{k=1}^{\ell-1} S_k$, hence

$$\underline{v} = \underline{v}_1 + \cdots + \underline{v}_{\ell-1} \quad (\underline{v}_k \in S_k,\ k = 1, \dots, \ell-1)$$

Therefore we have

$$L\underline{v} = \lambda_1 \underline{v}_1 + \cdots + \lambda_{\ell-1} \underline{v}_{\ell-1}$$

$\underline{v} \in S_\ell \Rightarrow L\underline{v} = \lambda_\ell \underline{v}$. Furthermore $v \in \bigoplus_{k=1}^{\ell-1} S_k$, hence

$$\underline{v} = \underline{v}_1 + \cdots + \underline{v}_{\ell-1} \quad (\underline{v}_k \in S_k,\ k = 1, \dots, \ell-1)$$

Therefore we have

$$L\underline{v} = \lambda_1 \underline{v}_1 + \cdots + \lambda_{\ell-1} \underline{v}_{\ell-1}$$

$$\lambda_\ell(\underline{v}_1 + \cdots + \underline{v}_{\ell-1}) = \lambda_\ell \underline{v} = \lambda_1 \underline{v}_1 + \cdots + \lambda_{\ell-1} \underline{v}_{\ell-1}$$

$$\Longrightarrow (\lambda_1 - \lambda_\ell)\underline{v}_1 + \cdots + (\lambda_{l-1} - \lambda_l)\underline{v}_{l-1} = \underline{0}$$

$$\Longrightarrow (\lambda_1 - \lambda_l)\underline{v}_1 = \cdots = (\lambda_{l-1} - \lambda_l)\underline{v}_{l-1} = \underline{0}$$

$$\Longrightarrow \underline{v}_1 = \cdots = v_{l-1} = 0 \Longrightarrow v = \underline{0}$$

Hence $S_1, \dots, S_l$ form a direct sum $\Longrightarrow$ follows by induction ∎

DIAGONALIZABILITY

Definition Diagonalizable

A linear map $L: \mathbb{F}^n \to \mathbb{F}^n$ is diagonalizable over $\mathbb{F}$ when $\exists$ an invertible $n \times n$ matrix $P \in M_{n\times n}(\mathbb{F})$ for which

$P^{-1}AP$ is a diagonal matrix

Notation

$$A\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = A(\alpha_1 \underline{e}_1 + \cdots + \alpha_n \underline{e}_n)$$

Let $\underline{v}_1, \cdots, \underline{v}_n$ be another basis of $\mathbb{F}^n$

$$\underline{e}_1 = \beta_{11}\underline{v}_1 + \cdots + \beta_{1n}\underline{v}_n$$
$$\vdots$$
$$\underline{e}_n = \beta_{n1}\underline{v}_1 + \cdots + \beta_{nn}\underline{v}_n$$

$$P = \begin{pmatrix}\beta_{11} & \cdots & \beta_{n1}\\ \vdots & \ddots & \\ \beta_{1n} & \cdots & \beta_{nn}\end{pmatrix}$$

Hence
$$A\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = A(\alpha_1 \underline{e}_1 + \cdots + \alpha_n \underline{e}_n)$$
$$= A(\alpha_1(\beta_{11}\underline{v}_1 + \cdots + \beta_{1n}\underline{v}_n) + \cdots + \alpha_n(\beta_{n1}\underline{v}_1 + \cdots + \beta_{nn}\underline{v}_n))$$
$$= AP\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}$$

Notation Diagonal Matrix

$$P^{-1}AP \text{ is diagonal} \iff P^{-1}AP = \begin{pmatrix}\lambda_1 & \cdots & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

Remark:

$\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$: not diagonalizable as an element of $M_{2\times 2}(\mathbb{R})$

: diagonalizable as an element of $M_{2\times 2}(\mathbb{C})$

Theorem

(1) A linear map $L: \mathbb{F}^n \to \mathbb{F}^n$ is diagonalizable over $\mathbb{F}$

$\Longleftrightarrow$

(2) $\mathbb{F}^n$ has a basis consisting of eigenvectors of L

(3) This is equivalent to saying that $\mathbb{F}^n$ has a direct sum of eigenspaces of L

$\Longleftrightarrow$

(4) sum of all dimensions of the eigenspaces of L equal to n

Proof:

$1 \Rightarrow 2$: Let A represent L. Diagonalizable $\Longrightarrow \exists\, P \in M_{n\times n}(\mathbb{R})$ such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = D \iff AP = PD$$

Remark: $\forall\, \underline{e}_j$, $D\underline{e}_j = \lambda_j \underline{e}_j \quad \forall j = 1, \ldots, n$

Therefore

$$AP\underline{e}_j = A(P\underline{e}_j) = PD\underline{e}_j$$
$$= P(\lambda_j \underline{e}_j) = \lambda_j P\underline{e}_j$$

$\Rightarrow A(P\underline{e}_j) = \lambda_j (P\underline{e}_j) \Longrightarrow P\underline{e}_j$ is an eigenvector of A

So $P\underline{e}_1, \ldots, P\underline{e}_n$ are eigenvectors of L.

But $P\underline{e}_j$ is the j^{th} column of P. By Lemma pg 44, $P\underline{e}_1, \ldots, P\underline{e}_n$ is a basis. (as P is invertible)

$2 \Rightarrow 1$: Let $\underline{v}_1, \ldots, \underline{v}_n$ be a basis of $\mathbb{F}^n$ s.t $L\underline{v}_j = \lambda_j \underline{v}_j$

Consider the matrix $P = (\underline{v}_1, \ldots, \underline{v}_n)$. Then $P\underline{e}_j = \underline{v}_j$

Correspondingly $\underline{e}_j = P^{-1}\underline{v}_j$. Then

$$P^{-1}AP\underline{e}_j = P^{-1}A\underline{v}_j = \lambda_j P^{-1}\underline{v}_j = \lambda_j \underline{e}_j \Longrightarrow P^{-1}AP = \begin{pmatrix} \lambda_1 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Proving the other equivalences

Let $\mu_1, \ldots, \mu_q$ be all the eigenvalues of L and let $S_1, \ldots, S_q$ be the corresponding eigenspaces

$$n_j = \dim S_j = \# B_j \quad , \quad j = 1, \dots, q$$

Choose a basis B_j for S_j, $j = 1, \dots, q$. Now, if

$$\mathbb{F}^n = \bigoplus_{j=1}^{q} S_j \text{ then } \bigcup_{j=1}^{q} B_j \text{ is a basis of } \mathbb{F}^n$$

Then $n = \sum_{j=1}^{q} \dim(S_j) = \sum_{j=1}^{q} n_j = \# \bigcup_{j=1} B_j$

i.e. this union consists of n linearly independent vectors $\Longrightarrow$ by Lemma pg 44, $\mathbb{F}^n$ has a basis of eigenvectors

Conversely if L has a basis of eigenvectors, then we can group this basis by corresponding eigenvalues to get the basis of each S_j

i.e. if $\bigcup_{j=1}^{q} B_j$ has n elements, then this is a basis of eigenvectors for $\mathbb{F}^n$, therefore

$$\mathbb{F}^n = \bigoplus_{j=1}^{q} S_j$$

$\blacksquare$

Example:

$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ does not have any real eigenvalue

$\Longrightarrow$ A is not diagonalizable

However, over $\mathbb{C}$, the matrix A has eigenvalues $\pm i$. So as $1+1 \geq 2$, we have a basis of eigenvectors of A

$\Longrightarrow$ A is diagonalizable over $\mathbb{C}$

CHARACTERISTIC POLYMOMIAL

Definition Characteristic Polynomial

$A \in M_{n\times n}(\mathbb{F})$

$$c_A(\lambda) = \det(\lambda I_n - A) = \lambda^n + c_1(A)\lambda^{n-1} + \cdots + c_n(A) = \lambda^n + \sum_{j=1}^{n} c_j(A)\lambda^{n-j}$$

If a linear map L is represented by L, then we also call c_A the characteristic polynomial of L

Properties of Determinants

Theorem Determinants

$A, B \in Mat_{n\times n}(\mathbb{F})$

i) $\det A = 0$ iff $rk A < n$ (equally A is not invertible)

ii) $\det(AB) = \det(A)\det(B)$

In particular, $\det A^{-1} = \frac{1}{\det A}$ if A is invertible

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

iii) If A is upper or lower triangular,

$\det A =$ product of diagonal elements $a_{11}, a_{22}, \dots, a_{nn}$

In particular, $\det(\alpha \cdot I_n) = \alpha^n$. Hence $\det(\alpha A) = \alpha^n \det A$

iv) $\det A^T = \det A$

Example:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$ch_A(\lambda) = \det\left(\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right)$$

$$= \det\begin{pmatrix} \lambda - a_{11} & \lambda - a_{12} \\ \lambda - a_{21} & \lambda - a_{22} \end{pmatrix}$$

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

$$= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + \det A = \lambda^2 - trace(A)\lambda + \det A$$

Lemma

Let $A \in Mat_{n\times n}(\mathbb{F})$. Then each $c_j(A)$ is a polynomial function of degree j in entries of A

Furthermore $\quad c_1(A) = -\sum_{j=1}^{n} a_{jj} = tr(A)$

$$c_n(A) = (-1)^n \det A$$

For every j $\quad c_j(B^{-1}AB) = c_jA \quad$ for every invertible matrix B

i.e. similar matrices have same characteristic polynomial

Proof:
$$ch_A(\lambda) = \det(\lambda I_n - A)$$
$$ch_A(\lambda) = \det B^{-1} \det(\lambda I_n - A) \det B$$
$$\begin{aligned} ch_A(\lambda) &= \det(B^{-1}(\lambda I_n - A)B) \\ &= \det(B^{-1}(\lambda I_n)B - B^{-1}AB) \\ &= \det(\lambda I_n - B^{-1}AB) \\ &= ch_{B^{-1}AB}(\lambda) \end{aligned}$$

■

Geometric Multiplicity

Let $L: \mathbb{F}^n \to \mathbb{F}^n$ be a linear map, λ is an eigenvalue of L $\left(\text{i.e } \exists\, \underline{v} \in \mathbb{F}^n \setminus \{\underline{0}\} \text{ such that } L\underline{v} = \lambda \underline{v}\right)$

$$L\underline{v} = \lambda\underline{v} \implies (L - \lambda I_n)\underline{v} = \underline{0}$$
$$\implies \ker(L - \lambda I_n) \neq \{\underline{0}\}$$

Definition Geometric multiplicity

Let $\ker(L - \lambda I_n)$ be the eigenspace

The geometric multiplicity is $\dim \ker(L - \lambda I_n)$

Algebraic Multiplicity

Let A be the matrix representing L.

$$c_A(\lambda) = \det(\lambda I_n - A) \quad \text{is the characteristic polynomial.}$$

So $c_A(\lambda_0) = 0 \implies c_A(\lambda) = (\lambda - \lambda_0)^k p(\lambda)$

$p(\lambda_0) \neq k$ is called algebraic multiplicity

Example:

$A_1 = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, $c_A(\lambda) = \det(\lambda I_2 - A_1)$

$= \det\begin{pmatrix} \lambda-3 & -1 \\ 0 & \lambda-3 \end{pmatrix} = (\lambda-3)^2 \Rightarrow$ Algebraic multiplicity w.r.t A_1 is 2

However $\dim \ker(A_1 - 3I_2) = \dim \ker\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$

geometric multiplicity by rank nullity thm

as $\text{rank}(A_1 - 3I_2) = \text{num of lin ind columns} = 1$

$A_2 = \begin{pmatrix} 3 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 3 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 3 & 1 & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & & 3 \end{pmatrix}$ $c_{A_1}(\lambda) = (\lambda-3)^n \Longrightarrow$ algebraic multiplicity is n

$\dim \ker(\lambda I_n - A_2) = 1$

Theorem

Let $L: \mathbb{F}^n \to \mathbb{F}^n$ be a linear map represented by $A \in \text{Mat}_{n\times n}(\mathbb{F})$

i) The eigenvalues of L are the roots in $\mathbb{F}$ of its characteristic polynomial $c_A(\lambda)$

The multiplicity of the root is called the algebraic multiplicity of the eigenvalue

ii) For a given eigenvalue λ, its algebraic multiplicity is bigger than its geometric multiplicity

(iii) If A is upper or lower triangular, then

$$c_A(\lambda) = (\lambda - a_{11}) \times \cdots \times (\lambda - a_{nn})$$

(iv) There are at most n different eigenvalues of L

(v) If L has n distinct eigenvalues (i.e. $c_A(\lambda)$ has distinct roots in $\mathbb{F}$,) then L is diagonalizable

Proof:

ii) Let geometric multiplicity of λ be q

$$\dim \ker(\lambda I_n - A) = q$$

and this, we can find a basis $S_1 = \{\underline{v}_1, \cdots\cdots, \underline{v}_q\}$ of $\ker(\lambda I_n - A)$

Complete S_1 to be a basis of $S = \{\underline{v}_1, \cdots\cdots, \underline{v}_q, \underline{u}_{q+1}, \cdots\cdots, \underline{u}_n\}$ of $\mathbb{F}^n$.

Let M_s be the matrix with columns $\underline{v}_1, \dots, \underline{u}_n$

$$M_s = (\underline{v}_1, \dots, \underline{v}_q, \underline{u}_{q+1}, \dots \underline{u}_n)$$

Note: $M_s \underline{e}_j = \underline{v}_j \quad \forall j = 1, \dots, q$

$$\Longrightarrow \underline{e}_j = M_s^{-1} \underline{v}_j$$

Claim: First q columns of $M_s^{-1} A M_s$ are $\lambda e_j \quad j \in \{1, \dots, q\}$

$$\forall j \in [\![1, q]\!] \quad M_s^{-1} A M_s \underline{e}_j = M_s^{-1} A \underline{v}_j$$

$$= M_s^{-1}(\lambda \underline{v}_j) \qquad \underline{v}_j \text{ belongs to eigenspace}$$

$$= \lambda M_s^{-1} \underline{v}_j$$

$$= \lambda \underline{e}_j$$

$$\text{So } M_s^{-1} A M_s = \begin{array}{c} \begin{array}{ccc} & & \downarrow q \end{array} \\ \begin{pmatrix} \lambda & 0 & 0 & \\ 0 & \lambda & 0 & \\ 0 & 0 & \lambda & \\ \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \end{pmatrix} \end{array}$$

The characteristic polynomial of $M_s^{-1} A M$ is

$$c_{M_s^{-1} A M_s}(x) = \det(x I_n - M_s^{-1} A M_s) = \det \begin{pmatrix} x-\lambda & 0 & 0 & \\ 0 & x-\lambda & 0 & \\ 0 & 0 & x-\lambda & \\ \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \end{pmatrix}$$

$$= (x-\lambda)^q p(x)$$

$$\Longrightarrow \text{algebraic multiplicity of } \lambda \text{ w.r.t } M_s^{-1} A M_s \text{ is } \geq q$$

But $c_A(x) = c_{M^{-1}AM}(x) \Longrightarrow c_A(x) = (x-\lambda)^q p(x) \Longrightarrow$ algebraic multiplicity of λ w.r.t $A \geq q$

Example:

i) $\begin{pmatrix} 2 & \pi & 0 & e \\ 0 & \pi^2 & 3/4 & 27 \\ 0 & 0 & \log 2 & 8 \\ 0 & 0 & 0 & 117 \end{pmatrix}$ upper triangular, has eigenvalues

$$-2, \pi^2, \log(2), 117 \quad \text{diagonal entries}$$

distinct $\Rightarrow$ diagonalizable

ii) $L(x_1, x_2) = (-x_2, x_1)$ has characteristic polynomial

$$c_{A_1}(\lambda) = \lambda^2 + 1$$

Since $tr(A_1) = 0$, $\det(A_1) = 1$. So as a linear map from $\mathbb{R}^2$ to $\mathbb{R}^2$, this has no eigenvalues in $\mathbb{R}$

$\Rightarrow$ not diagonalizable in $\mathbb{R}$

if $L: \mathbb{C}^2 \to \mathbb{C}^2$, eigenvalues are

$$(\lambda^2 + 1) = (\lambda - i)(\lambda + i) \Rightarrow \lambda = i \text{ or } \lambda = -i$$

$$\Rightarrow \text{diagonalizable}$$

Follows from section on diagonalizability

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

iii) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ upper triangular $\Rightarrow$ eigenvalues are $1, 1, 0$

Algebraic multiplicity of 1 is 2.

Algebraic multiplicity of 0 is 1

$$\ker(A - I_3) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = Sp(\underline{e}_1, \underline{e}_2)$$

and $\ker(A) = Sp(\underline{e}_2 - \underline{e}_3)$

Basis of eigenvectors $\underline{e}_1, \underline{e}_2, \underline{e}_2 - \underline{e}_3$

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

DETERMINANTS: a reminder

Definition Determinant

Let $A=(a_{ij})$ be a matrix $A \in Mat_{n\times n}(\mathbb{F})$.

The **determinant** is defined to be

$$\det A = \begin{vmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{vmatrix} = \sum_{\sigma \in S_n} (-1)^{N(\sigma)} a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

Here

$$\boxed{S_n = \{\sigma : \{1,\dots,n\} \longrightarrow \{1,\dots,n\} \mid \sigma \text{ is invertible}\}}$$

S_n is the set of all permutations (Symmetric group)

$$\boxed{N(\sigma) = \#\{(j,k) \mid 1 \le j \le k \le n \text{ and } \sigma(j) > \sigma(k)\}}$$ number of inversions of σ

Theorem

Let $A \in Mat_{n\times n}(\mathbb{F})$ with columns $A_1, \dots, A_n$ with $\det A$

i) Swapping columns changes sign of $\det A$

ii) If 2 columns of A are scalar multiples of each other, then $\det(A) = 0$

iii) If j^{th} column of A is replaced by $\alpha A_j + \beta A_k$, then the new matrix has determinant $\alpha \det(A)$

Properties are equally true if "columns" are replaced by "rows"

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{23}a_{31}$$

Cofactors

Definition Minor and Cofactor

Let $A = (A_{jk}) \in M_{n \times n}(\mathbb{F})$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

For each $1 \le j, k \le n$, we can get an $(n-1) \times (n-1)$ matrix by deleting the j^{th} row and k^{th} column

The determinant of such a matrix is denoted by $M_{jk} \in \mathbb{F}$ is called the (j,k)-minor of A

The (j,k)- cofactor of A is defined to be

$$C_{jk} := (-1)^{j+k} M_{jk}$$

Theorem Laplace Formula for Determinant

Let $A = (A_{jk}) \in M_{n \times n}(\mathbb{F})$.

1) For each fixed all $1 \le j \le n$, expansion along j-th row

$$\det(A) = \sum_{k=1}^{n} A_{jk} C_{jk}$$

2) For each fixed all $1 \le k \le n$, expansion along k-th column

$$\det(A) = \sum_{j=1}^{n} A_{jk} C_{jk}$$

Definition Cofactor Matrix

Let $A \in Mat_{n \times n}(\mathbb{F})$

The cofactor matrix of A is

$$\mathrm{cof}(A) = (C_{jk}) = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{2n} & \cdots & C_{nn} \end{pmatrix}$$

((j,k) entry is the $(j-k)$- cofactor of A)

The classical adjoint (or adjugate matrix) of A is the transpose of cofactor matrix

$$\mathrm{adj}(A) = \mathrm{cof}(A)^T = (C_{jk})$$

Theorem Adjugate, determinant and inverse

Let $A = (A_{ij}) \in M_{n \times n}(\mathbb{F})$. Then

$$A \cdot \operatorname{adj} A = \det(A) I_n = \operatorname{adj}(A) A$$

In particular, if $\det(A) \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

CAYLEY HAMILTON THEOREM

Suppose

$$p(x) = \alpha_m x^m + \alpha_{m-1} x^{m-1} + \cdots + \alpha_1 x^1 + \alpha_0$$

is a polynomial with coefficients in $\mathbb{F}$

Given $A \in Mat_{n \times n}(\mathbb{F})$, define $P(A)$

$$P(A) = \alpha_m A^m + \alpha_{m-1} A^{m-1} + \cdots + \alpha_1 A + \alpha_0 I$$

Theorem The Cayley Hamilton Theorem

Let $A \in M_{n \times n}(\mathbb{F})$. Let c_A be the characteristic polynomial of A. Then

$$c_A(A) = 0 \in Mat_{n \times n}(\mathbb{F})$$

Proof:

$B = A - xI$

$c_A(x) = (-1)^n \det B$

By thm above, $B\operatorname{adj}(A) = \det(B) I_n$

$\operatorname{adj}(B) = (B_{ij})$. Every B_{ij} is a polynomial of degree $n-1$ by defn of cofactor

$$B_{ij} = \sum_{k=0}^{n-1} b_{ijk} x^k = p_{ij}(x) \qquad \operatorname{adj} B = \begin{pmatrix} p_{11}(x) & \cdots & p_{1n}(x) \\ \vdots & \ddots & \\ p_{n1}(x) & \cdots & p_{nn}(x) \end{pmatrix}$$

$\operatorname{adj}(B) = B_0 + B_1 x + B_2 x^2 + \cdots + B_{n-1} x^{n-1}$ where

$$B_0 = (b_{ij0}), \ \cdots \ B_\ell = (b_{ij\ell})$$

$$\det(B) I_n = B\operatorname{adj}(B) = B(B_0 + B_1 x + \cdots + B_{n-1} x^{n-1})$$

$$= (A - xI)(B_0 + Bx + \cdots + B_{n-1} x^{n-1})$$

$$= AB_0 + AB_1x + \cdots + AB_{n-1}x^{n-1} - B_0x - B_1x^2 - \cdots - B_{n-1}x^n$$

$$= AB_0 + (AB_1 - B_0)x + \cdots - B_{n-1}x^n \quad \text{(eq 1)}$$

$$c_A(x) = c_0 + c_1x + \cdots + c_nx^n$$

$$\det BI_n = c_A(x) \implies \det(B)I_n = c_0I + c_1Ix + \cdots + c_{n-1}Ix^{n-1} + c_nIx^n \quad \text{(eq 2)}$$

Comparing coefficients of (eq 1) and (eq 2)

$$\begin{aligned}
c_0I &= AB_0 && \times I \\
c_1I &= AB_1 - B_0 && \times A \\
c_2I &= AB_2 - B_1 && \times A^2 \\
&\vdots \\
c_{n-1}I &= AB_{n-1} - B_{n-2} && \times A^{n-1} \\
c_nI &= -B_{n-1} && \times A^n
\end{aligned}$$

$$\implies \quad \begin{aligned}
c_0I &= AB_0 \\
c_1A &= A^2B_1 - AB_0 \\
c_2A^2 &= A^3B_2 - A^2B_1 \\
&\vdots \\
c_{n-1}A^{n-1} &= AB_{n-1} - B_{n-2} \\
c_nA^n &= -A^nB_{n-1}
\end{aligned}$$

Adding these up, all terms on RHS cancel

$$\implies c_A(A) = c_0I + c_1A + \cdots + c_nA^n = 0 \qquad \blacksquare$$

Example: Using Cayley-Hamilton Theorem to find inverse of

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

$\det A = 5 \implies A^{-1}$ exists

$$c_A(x) = \begin{vmatrix} 3-x & 0 & 4 \\ 1 & 1-x & 2 \\ 1 & 0 & 3-x \end{vmatrix} = (x-1)^2(x-5) = -x^3 + 7x^2 - 11x + 5$$

According to Cayley-Hamilton theorem

$$-A^3+7A^2-11A+5I=0_{n\times n}$$

Multiplying this by A^{-1} on the left, we find that

$$-A^2+7A-11I+5A^{-1}=0_{n\times n} \quad \text{or} \quad A^{-1}=\frac{1}{5}(A^2-7A+11I)$$

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

Linking some ideas

$A\in M_{n\times n}(\mathbb{F})$, $n,K\in\mathbb{F}$

$$A^K=A\cdots\cdots A \quad K \text{ times} \qquad A^0=I$$

$$\sum_{K=0}^{N}\alpha_K A^K=p(A), \qquad p(x)=\sum_{K=0}^{N}\alpha_K x^K$$

$$c_A(A)=0_{n\times n}$$

$\displaystyle\sum_{K=0}^{\infty}\alpha_K A_K$: power series : used to denote analytic functions like e^A

$$\operatorname{adj}(A)\,A=\det(A)\,I_{n\times n}=A\operatorname{adj}(A)$$

We have characteristic monic polynomial

$$c_A(x)=x^n+a_1x^{n-1}+\cdots\cdots+a_{n-1}x+a_n$$

$$a_n=(-1)^n\det A\, I_{n\times n}$$

Substituting A by Cayley-Hamilton Theorem

$$0_{n\times n}=c_A(A)=A^n+a_1A^{n-1}+\cdots\cdots+a_{n-1}A+(-1)^n\det(A)\,I_{n\times n}$$

$$\Longrightarrow (-1)^{n+1}\det A\, I_{n\times n}=A^n+a_1A^{n-1}+\cdots\cdots+A) \qquad \text{by subtracting } (-1)^n\det A$$

$$=(A^{n-1}+a_1A^{n-2}+\cdots\cdots+a_{n-1})A$$

$$=A\underbrace{(A^{n-1}+a_1A^{n-2}+\cdots\cdots+a_{n-1})}_{=M}$$

$$\Rightarrow (-1)^{n+1}\det(A)I_{n\times n} = A\cdot M = M\cdot A$$

$$\boxed{\det A \neq 0 \Rightarrow M = \mathrm{adj}(A)(-1)^{n+1} = (-1)^{n+1}(A^{n-1} + a_1 A^{n-2} + \cdots + a_{n-1})}$$

MINIMAL POLYNOMIAL

Let $L : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ be a linear map (represented by $A \in \mathrm{Mat}_{n\times n}(\mathbb{F})$)

Define $L^2, L^3, \ldots\ldots$

$$L^2 = L\circ L,\quad L^3 = L^2\circ L = L\circ L\circ L,\quad \ldots\ldots,\quad L^k = L^{k-1}\circ L = L\circ L\circ\cdots\cdots\circ L \quad k \text{ times}$$

represented by $A^2 = AA$ $\quad A^3 \quad A^k$

Let $p(x)$ be the polynomial

$$p(x) = \alpha_m x^m + \alpha_{m-1}x^{m-1} + \cdots\cdots + a_1 x + a_0 x^0$$

with coefficients $\alpha_i \in \mathbb{F}$

Define linear transformation by formula: $p(L) : \mathbb{F}^n \to \mathbb{F}^n$

$$\boxed{p(L) = \alpha_m L^m + \alpha_{m-1}L^{m-1} + \cdots\cdots + \alpha_1 L + \alpha_0 I}$$

Definition Minimal Polynomial of A

Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{F})$.

The minimal polynomial of A $d_A(x)$ is the monic polynomial $p(x)$ of least degree vanishing at A:

$$p(A) = 0$$

Example

$A = I_n$. $\quad c_{I_n} = (x-1)^n = \det(xI_n - I_n)$

At the same time $\quad d_{I_n} = (x-1)$

Indeed $d_{I_n}(I_n) = I_n - I_n = 0_{I_n}$

Definition Minimal Polynomial of L

Let $L : \mathbb{F}^n \to \mathbb{F}^n$ be a linear map

The minimal polynomial of L $d_L(x)$ is the monic polynomial $p(x)$ of least degree vanishing at L: $\quad p(L) = 0$

From now on, let $A \in Mat_{n\times n}(\mathbb{F})$ represent linear map $L: \mathbb{F}^n \longrightarrow \mathbb{F}^n$.

I use L and A interchangeably

Lemma

The minimal polynomial of A is unique and fixed and divides the characteristic polynomial

Proof:

$\deg d_A(x) = m$. Let $p(x)$ be the polynomial s.t $p(A) = 0$

$$\Longrightarrow \deg p(x) \geq m$$

By Euclidean division,

$$p(x) = q(x)\, d_A(x) + r(x) \qquad \deg r \leq m-1$$

Substituting $x = A$

$$p(A) = q(A)\, d_A(A) + r(A) \Longrightarrow 0_{n\times n} = q(A)\, 0_{n\times n} + r(A)$$

$$\Longrightarrow r(A) = 0_{n\times n}$$

This contradicts assertion that $d_A(x)$ has the smallest degree and $\deg(r) < \deg(d_A)$ unless

$$r(x) \equiv 0$$

$$\Longrightarrow p(x) = q(x)\, d_A(x)$$

$$\Longrightarrow d_A(x) \mid p(x)$$

Uniqueness:

Let $d_A(x)$ and $\delta_a(x)$ both verify defn of minimial polynomial.

$$\Longrightarrow \begin{cases} d_A(x) \mid \delta_a(x) \\ \delta_A(x) \mid d_A(x) \end{cases} \Longrightarrow d_A(x) = \beta \delta_A(x) \qquad \beta \in \mathbb{F} \setminus \{0\}$$

As d_A and δ_a both monic, $\beta = 1$ by comparing co-efficients with x^n

$$\Longrightarrow d_A(x) = \delta_a(x)$$

$\blacksquare$

Let $L: \mathbb{F}^n \longrightarrow \mathbb{F}^n$ be a linear map. (represented by $A \in Mat_{n\times n}(\mathbb{F})$)

Define $L^2, L^3, \dots\dots$

$$\underbrace{L^2}_{\substack{\text{represented}\\ \text{by } A^2 = AA}} = L \circ L, \quad \underbrace{L^3}_{A^3} = L^2 \circ L = L \circ L \circ L, \quad \dots\dots, \underbrace{L^K}_{A^K} = L^{K-1} \circ L = L \circ L \circ \dots\dots \circ L \quad K \text{ times}$$

Let $p(x)$ be the polynomial

$$p(x) = \alpha_m x^m + \alpha_{m-1} x^{m-1} + \dots\dots + a_1 x + a_0 x^0$$

with coefficients $\alpha_i \in \mathbb{F}$

Define linear transformation by formula: $p(L): \mathbb{F}^n \to \mathbb{F}^n$

$$\boxed{p(L) = \alpha_m L^m + \alpha_{m-1} L^{m-1} + \dots\dots + \alpha_1 L + \alpha_0 I}$$

By property of linear maps, $p(L)$ is represented by $P(A)$

So by cayley-hamilton, $c_L(L) = 0_{n\times n}$

We also have minimial polynomial of L, $d_L(x)$

Theorem

Let $L: \mathbb{F}^n \to \mathbb{F}^n$ a linear map

A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of $L \iff d_L(\lambda) = 0$

Proof:

$(\Longleftarrow)$: Assume λ be a root of $d_L(x)$

$$d_L(\lambda) = 0 \implies d_L(x) = (x-\lambda)\, p(x) \qquad \deg d_L = \deg(x-\lambda) + \deg p(x)$$

$$\deg p(x) < \deg d_L(x) \implies p(L) \neq 0$$

So $\exists\, \underline{v} \in \mathbb{F}^n \setminus \{\underline{0}\}$ such that

$$\underline{w} := p(L)\underline{v} \neq \underline{0}$$

$$\begin{aligned} d_L(L)(\underline{v}) = \underline{0} &= [(L - \lambda I)\, p(\lambda)]\, \underline{v} \\ &= (L - I\lambda)[P(L)\underline{v}] \quad \text{associativity} \\ &= (L - I\underline{v})\, \underline{w} = L\underline{w} - \lambda \underline{w} = \underline{0} \implies L\underline{w} = \lambda \underline{w} \implies \lambda \text{ is an eigenvalue of } L \end{aligned}$$

$(\Longrightarrow)$: Assume $\underline{v} \in \mathbb{F}^n \setminus \{\underline{0}\}$ is an eigenvector

$\exists \underline{v} \in \mathbb{F}^n \setminus \{\underline{0}\}$ s.t $L\underline{v} = \lambda \underline{v}$

By definition of minimal polynomial, $d_L(L) = 0_{n \times n}$

$$0_{n\times n} = d_L(L)(\underline{v}) = (L^n + \alpha_{n-1}L^{n-1} + \cdots + \alpha_1 L + \alpha_0 \cdot id)(\underline{v})$$
$$= L^n\underline{v} + \alpha_{n-1}L^{n-1}\underline{v} + \cdots + \alpha_1 L\underline{v} + \alpha_0\underline{v}$$

Note: $L^2\underline{v} = L(\lambda\underline{v}) = \lambda L\underline{v} = \lambda^2\underline{v} \Longrightarrow \forall k \in \mathbb{N},\ L^k\underline{v} = \lambda^k\underline{v}$

Therefore

$$0_{n\times n} = d_L(\lambda) = \lambda^n\underline{v} + \alpha_{n-1}\lambda^{n-1}\underline{v} + \cdots + \alpha_1\lambda\underline{v} + \alpha_0\underline{v}$$
$$= (\lambda^n \cdot id + \alpha_{n-1}\lambda^{n-1} \cdot id + \cdots + \alpha_1\lambda id + \alpha_0 id)$$
$$= d_L(\lambda) \cdot id\, \underline{v} = d_L(\lambda)\underline{v}$$

$\Rightarrow d_L(\lambda)\underline{v} = 0_{n\times n}$ and $\underline{v} \neq \underline{0}$

$\Rightarrow d_L(\lambda) = 0$

Example

$$A = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Characteristic polynomial

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x-1 & -1 & -4 \\ 0 & x-2 & -1 \\ 0 & 0 & x-2 \end{vmatrix} = (x-1)^2(x-2)^2$$

$$c_B(x) = (x-1)(x-2)^2$$

$\Longrightarrow$ has eigenvalues 1 and 2

For A: Try $p(x) = (x-1)(x-2)$ both have deg 1 least power such that $d_A | c_A$

$$p(A) = (A-I)(A-2I) = \begin{pmatrix} 0 & 1 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \neq 0 \Longrightarrow p(x) \neq d_A(x)$$

$\Longrightarrow d_A(x) = c_A(x)$

For B: testing $p(x) = (x-1)(x-2)$

$$p(B) = (B-I)(B-2I) = \begin{pmatrix} 0 & 1 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies p(x) = d_B(x)$$

Definition Minimal Multiplicity

The **minimal multiplicity** $m_\lambda \in \mathbb{N}$ of an eigenvalue λ of a linear map $L: \mathbb{F}^n \to \mathbb{F}^n$ or of $A \in \mathrm{Mat}_{n\times n}(\mathbb{F})$ representing L is the multiplicity of λ as a root of $d_L(x)$ or $d_A(x)$

Note:

$d_L(x) \mid c_L(x)$

multiplicity of λ as a root of $c_L(x)$ is algebraic multiplicity

$\boxed{\text{minimal multiplicity} \leq \text{algebraic multiplicity}}$

3. Jordan's Theorem

Definition Elementary Jordan block

For $\lambda \in \mathbb{C}$, the **elementary Jordan block** (of size ℓ with eigenvalue λ) is the $\ell \times \ell$ matrix

$$J_{\lambda,\ell} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

$$\boxed{J_{\lambda,\ell} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} + \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}}$$

Definition Jordan Normal Form

The $n \times n$ matrix is said to have the **Jordan normal form** if

$$J = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_k} \end{pmatrix}$$

where for each $i = 1, \dots, k$

$$J_i = J_{\lambda_i, \ell_i}$$

for some complex numbers $\lambda_1, \dots, \lambda_k \in \mathbb{C}$, integers $\ell_1, \dots, \ell_k \in \mathbb{N}$

Example:

$$\begin{pmatrix} \boxed{\begin{matrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix}} & & \\ & & \boxed{5} & \\ 0 & & & \boxed{\begin{matrix} 7 & 1 \\ 0 & 7 \end{matrix}} \end{pmatrix}$$

Theorem Jordan's Theorem

i) $\forall A \in Mat_{n\times n}(\mathbb{C})$, $\exists P, J \in Mat_{n\times n}(\mathbb{C})$ where P is invertible, J has a Jordan normal form and

$$A = PJP^{-1}$$

ii) The collection of pairs $(\ell_1, \lambda_1), \ldots\ldots, (\lambda_k, \ell_k)$ is determined uniquely by the given matrix A upto reordering these pairs

(iii) The matrix P can chosen so that the diagonal blocks of J with same eigenvalue λ appear consecutively (one after another), and the sizes of these consecutive blocks with the same λ do not increase as one goes the diagonal

MULTIPLICITIES AND EIGENVALUES

Example:

$$J = \begin{pmatrix} \boxed{\begin{matrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{matrix}} & & 0 \\ & \boxed{5} & \\ 0 & & \boxed{\begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix}} \end{pmatrix}$$

$\lambda_1 = 5 \quad \lambda_2 = 3$

	alg mult	geo mult	min mult
λ_1:	4	2	3
λ_2:	2	1	2

Algebraic multiplicity:

1) $\lambda_1 = 5$: 5 appears 4 times in diagonal $\Rightarrow$ alg mult $= 4$

2) $\lambda_2 = 3$: 3 appears 2 times in diagonal $\Rightarrow$ alg mult $= 2$

Geometric Multiplicity

1) $\lambda_1 = 5$: There are 2 blocks with diagonal $\lambda = 5 \Rightarrow$ geo mult $= 2$

2) $\lambda_2 = 3$: There are 1 blocks with diagonal $\lambda = 3 \Rightarrow$ geo mult $= 1$

Note:

- For any upper triangular matrix, its eigenvalues are its diagonal elenents

$$\mathcal{D} = \{a_{11}, \ldots, a_{nn}\}$$

alg multiplicity is the number of times it appears in $\mathcal{D}$

- $\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$ geo multiplicity of λ is 1

General Case:

Let J be a matrix with Jordan Normal Form

$$J = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_K} \end{pmatrix}$$

- **Algebraic multiplicity**: The number of times any given $\lambda \in \mathbb{C}$ appears on the diagonal of J is algebraic multiplicity a_λ

To see this, det of any upper triangular matrix is product of diagonal

$$c_j(x) = \prod_{i=1}^{K} (\lambda_i - x)^{\ell_i}$$

$\Longrightarrow$ a_λ is the total size of all Jordan blocks of J with the given eigenvalue λ

- **Geometric multiplicity**: The number of elementary Jordan blocks with same eigenvalue λ = geo mult g_λ

By defn, g_λ = maximal number of linearly independent eigenvectors associated to λ.

Each elementary Jordan block has only one eigenvector associated to it

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \quad \text{geo mult} = 1, \text{ eigenvector } \vec{v} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

To generalize to arbitrary J, let

$X \in Mat_{p \times p}$ $\qquad$ $Y = Mat_{q \times q}$

$\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_p \end{pmatrix}$ $p \times 1$ $\qquad$ $\underline{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_q \end{pmatrix}$ $q \times 1$

$$A := \begin{array}{cc} & \\ p & \\ & q \end{array} \begin{pmatrix} \boxed{X} & 0 \\ 0 & \boxed{Y} \end{pmatrix} \qquad \underline{w} = \begin{pmatrix} \underline{u} \\ \underline{v} \end{pmatrix} \begin{matrix} \}\, p \\ \}\, q \end{matrix}$$

$$\mu := A\underline{w} = \begin{pmatrix} \boxed{x} & 0 \\ 0 & \boxed{y} \end{pmatrix}\begin{pmatrix} \underline{u} \\ \underline{v} \end{pmatrix} = \begin{pmatrix} x \cdot \underline{u} \\ y \cdot \underline{v} \end{pmatrix} = \lambda \begin{pmatrix} \underline{u} \\ \underline{v} \end{pmatrix} = \begin{pmatrix} \lambda \underline{u} \\ \lambda \underline{v} \end{pmatrix}$$

if this is an eigenvector

For any $\lambda \in \mathbb{C}$, $A\underline{w} = \lambda \underline{w} \implies \begin{pmatrix} \underline{u} \\ \underline{v} \end{pmatrix}$ is an eigenvector corresponding to eigenvalue λ of μ

$$\Updownarrow$$

a) $\underline{u}$ is an eigenvector of x corresponding to λ

$\underline{v}$ is an eigenvector of y corresponding to λ

b) $\underline{u} = \underline{0}$

$\underline{v}$ is an eigenvector of y corresponding to λ

c) $\underline{v} = \underline{0}$

$\underline{u}$ is an eigenvector of x corresponding to λ

Example: Continuing example above

Assume

$\begin{pmatrix} \underline{u} \\ \underline{w} \\ \underline{v} \end{pmatrix}$ is an eigenvector of J corresponding to eigenvalue λ

$\underline{u}$ is in eigenspace $ES_5\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix} = Sp(e_1)$

$\underline{v} \in Es_5(5) = Sp(e_2)$

$\underline{w} \in Es_5\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ $\quad \underline{w} = \underline{0}$

So $\quad Es_5(J) = Sp\left(\begin{pmatrix} e_1 \\ \underline{0}_1 \\ \underline{0}_2 \end{pmatrix} + \begin{pmatrix} \underline{0}_3 \\ e_2 \\ \underline{0}_2 \end{pmatrix}\right)$

▶ Minimal Multiplicity: The maximal size of any elementary Jordan block with eigenvalue λ is the minimal multiplicity m_λ

To see this observe

$$\begin{pmatrix} \underbrace{\boxed{x}}_{a} & 0 \\ 0 & \underbrace{\boxed{y}}_{b} \end{pmatrix}\begin{pmatrix} \underbrace{\boxed{A}}_{a} & 0 \\ 0 & \underbrace{\boxed{B}}_{b} \end{pmatrix} = \begin{pmatrix} \boxed{xA} & 0 \\ 0 & \boxed{yB} \end{pmatrix}$$

Definition Nilpotent

If a matrix A satisfies $A^m = 0$ but $A^{m-1} \neq 0$ then

A is **nilpotent** of degree m

Example:

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

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

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

$$A^4 = A^3 A = 0_{n \times n} \implies A^4 = 0_{n \times n}$$

$$\implies \text{nilpotent degree} = 4$$

Therefore if $J_{\lambda,\ell}$ is an elementary Jordan block of size ℓ with eigenvalue λ, that is

$$J_{\lambda,\ell} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{pmatrix}$$

then

$$J_{\lambda,\ell} - \lambda I_\ell = J_{0,\ell} = \left.\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{pmatrix}\right\} \ell \text{ rows}$$

and this is nilpotent of degree ℓ

$$(J_{\lambda,\ell} - \lambda I_\ell)^\ell = 0_{\ell \times \ell} = J_{0,\ell}^\ell$$

Now consider the matrix J having Jordan normal form

Notation: $\sigma(J) = \{\text{set of all distinct eigenvalues of } J\}$
$= \{\lambda_1, \cdots\cdots, \lambda_k : \lambda_i \neq \lambda_j\}$

For each $\lambda \in \sigma(J)$, denote by ℓ_λ, the largest ℓ_i associated to eigenvalue $\lambda = \lambda_i$

Define
$$p(x) = \prod_{\lambda \in \sigma(J)} (x-\lambda)^{\ell_\lambda}$$

Theorem

$$d_J(x) = p(x) = \prod_{\lambda \in \sigma(J)} (x-\lambda)^{\ell_\lambda}$$

proof:

Fix any $\lambda \in \sigma(J)$ and consider any Jordan block $J_i = J_{\lambda_i \ell_i}$ such that $\lambda_i = \lambda$

Then
$$p(x) = (x-\lambda)^{\ell_\lambda} \cdot \prod_{\mu \in \sigma(J)} (x-\mu)^{\ell_\mu}$$

Then
$$p(J_{\lambda_i,\ell_i}) = (J_{\lambda_i,\ell_i} - \lambda I_{\ell_i})^{\ell_\lambda} \cdot \prod_{\mu \in \sigma(J) \setminus \{\lambda\}} (J_{\lambda_i,\ell_i} - \mu I_{\ell_i})^{\ell_\mu}$$

$$= J_{0,\ell_i}^{\ell_\lambda} \cdot \prod_{\mu \in \sigma(J) \setminus \{\lambda\}} J_{\lambda_i - \mu_i \ell_i}^{\ell_\mu}$$

$$= 0$$

as $\ell_i < \ell_\lambda$ and $J_{0,\ell_i}^{\ell_i} = 0 \implies J_{0,\ell_i}^{\ell_\lambda} = 0$

$\implies p(J_{\lambda_i \ell_i}) = 0$ $\forall$ Jordan blocks of J

$\implies P(J) = 0$

So $p(x)$ annhilates J.

Showing $p(x)$ is the minimal polynomial

Consider any other polynomial $q(x)$ that divides $p(x)$. Then
$$q(x) = \prod_{\lambda \in \sigma(J)} (x-\lambda)^{\ell_\lambda'}$$

for some $\ell_\lambda' \leq \ell_\lambda$ where atleast one of the inequalities is strict

Fix any λ with $\ell_\lambda' < \ell_\lambda$. Take any Jordan block $J_i = J_{\lambda_i, \ell_i}$ such that $\lambda_i = \ell_i$ and $\ell_i = \ell_\lambda$
That is for our fixed λ, we take Jordan block of maximal size. Then

$$q(J_{\lambda_i, \ell_i}) = J_{0,\ell_i}^{\ell_\lambda'} \cdot \prod_{\mu \in \sigma(J) \setminus \{\lambda\}} J_{\lambda_i - \mu, \ell_i}^{\ell_\mu'} \neq 0$$

because $\ell_\lambda' < \ell_\lambda = \ell_i$ and $J_{0,\ell_i}^{\ell_\lambda'} \neq 0$ by above argument.

Also each matrix $J_{\lambda_i - \mu, \ell_i}$ is non-singular as $\lambda_i \neq \mu$. Thus $q(\lambda_i, \ell_i) \neq 0$

$$\Rightarrow q(J) = 0$$

■

Example: What are the algebraic, geometric and minimal multiplicities of

$$A = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 15 \end{pmatrix}$$

Solution:

	Alg mult	Geo Mult	Min Mult
$\lambda_1 \cdot 5$	2	1	2
λ_2: 15	1	1	1

Eigenvectors:

1) $\lambda_1 : 5$ $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{0} \Longrightarrow \begin{cases} y = 0 \\ 10z = 0 \Rightarrow z = 0 \end{cases}$

Hence $\vec{v} = \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}$ $\alpha = 1$

2) $\lambda_2 = 15$: $\begin{pmatrix} -10 & 1 & 0 \\ 0 & -10 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{0} \Longrightarrow \begin{cases} -10x + y = 0 \Rightarrow x = 0 \\ -10y = 0 \Rightarrow y = 0 \end{cases}$

$\vec{v} = \begin{pmatrix} 0 \\ 0 \\ \alpha \end{pmatrix}$ $\alpha = 1$

Minimal polynomial: $(x-5)^2(x-15)$

Notice

$$A-5I=\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 10 \end{array}\right)$$

$$(A-5I)^2=\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 10^2 \end{array}\right)$$

while

$$A-15I=\left(\begin{array}{cc|c} -10 & 1 & 0 \\ 0 & -10 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

and so

$$(A-5I)^2(A-15I)=\left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 10^2 \end{array}\right)\left(\begin{array}{cc|c} -10 & 1 & 0 \\ 0 & -10 & 0 \\ \hline 0 & 0 & 0 \end{array}\right) = \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

whereas

$$(A-5I)(A-15I)=\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 10 \end{array}\right)\left(\begin{array}{cc|c} -10 & 1 & 0 \\ 0 & -10 & 0 \\ \hline 0 & 0 & 0 \end{array}\right) = \left(\begin{array}{cc|c} 0 & -10 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

Definition Spectrum

For any matrix A with complex entries, the set of all pairwise distinct eigenvalues of A is called the *spectrum* of A and denoted by $\sigma(A)$

$$\sigma(A)=\{\lambda_i \mid \lambda_i \neq \lambda_j \quad \forall i \neq j\}$$

4. Constructing Jordan Normal Form

<u>Observe</u>: Consider elementary Jordan block

$$J_{\lambda,\ell} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{pmatrix}_{\ell \times \ell} \quad \text{with any eigenvalue } \lambda \in \mathbb{C}$$

This $\ell \times \ell$ matrix can be regarded as a linear transformation of co-ordinate vector space $\mathbb{C}^\ell$

Let $\underline{e}_1, \cdots \underline{e}_\ell$ be standard basis of $\mathbb{C}^\ell$

$$J - \lambda I = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

Then

$$\left.\begin{array}{l} \blacktriangleright\ (J-\lambda I)\underline{e}_1 = \underline{0} \\ \blacktriangleright\ (J-\lambda I)\underline{e}_2 = \underline{e}_1 \\ \blacktriangleright\ (J-\lambda I)\underline{e}_3 = \underline{e}_2 \\ \quad\vdots \\ \blacktriangleright\ (J-\lambda I)\underline{e}_\ell = \underline{e}_{\ell-1} \end{array}\right\} \Longrightarrow J_{\lambda,\ell} - \lambda I_\ell : \underline{e}_\ell \mapsto \underline{e}_{\ell-1} \mapsto \cdots \mapsto \underline{e}_2 \mapsto \underline{e}_1 \mapsto 0$$

Arrange the basis vectors into a column array called a tower:

$$\begin{array}{c} \underline{e}_\ell \\ \downarrow \\ \underline{e}_{\ell-1} \\ \vdots \\ \downarrow \\ \underline{e}_2 \\ \downarrow \\ \underline{e}_1 \\ \downarrow \\ \underline{0} \end{array}$$

<u>Example</u>:

$$A = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{matrix}} & & & \\ & \boxed{\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{matrix}} & & \\ & & \boxed{\begin{matrix} 2 & 0 \\ 1 & 2 \end{matrix}} & \\ & & & \boxed{2} \end{pmatrix}_{9 \times 9}$$

$9\times 9 \Rightarrow 9$ basis vectors: $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_9$

$$A-2I=\begin{pmatrix} \boxed{\begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix}} & & & 0 \\ & \boxed{\begin{matrix}0&1&0\\0&0&1\\0&0&0\end{matrix}} & & \\ 0 & & \boxed{\begin{matrix}0&1\\0&0\end{matrix}} & \\ & & & \boxed{0} \end{pmatrix}$$

$(A-2I)\underline{e}_1 = \underline{0}$

$(A-2I)\underline{e}_2 = \underline{e}_1$

$(A-2I)\underline{e}_3 = \underline{e}_2$

$(A-2I)\underline{e}_4 = \underline{0}$

$(A-2I)\underline{e}_5 = \underline{e}_4$

$(A-2I)\underline{e}_6 = \underline{e}_5$

$(A-2I)\underline{e}_7 = \underline{0}$

$(A-2I)\underline{e}_8 = \underline{e}_7$

$(A-2I)\underline{e}_9 = \underline{0}$

So we get tower

$$\begin{array}{cccc} \underline{e}_3 & \underline{e}_6 & & \\ \downarrow & \downarrow & & \\ \underline{e}_2 & \underline{e}_5 & \underline{e}_8 & \\ \downarrow & \downarrow & \downarrow & \\ \underline{e}_1 & \underline{e}_4 & \underline{e}_7 & \underline{e}_9 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{array}$$

pyramid of basis of vectors

$$\begin{array}{llll} \underline{e}_3 & \underline{e}_6 & & \\ \underline{e}_2 & \underline{e}_5 & \underline{e}_8 & \\ \underline{e}_1 & \underline{e}_4 & \underline{e}_7 & \underline{e}_9 \end{array}$$

More generally consider Jordan normal form A of size $n\times n$ with blocks of same eigenvalue (only one eigenvalue)

$$\lambda = \lambda_1 = \dots = \lambda_k$$

$$A=\begin{pmatrix} \boxed{J_{\lambda,\ell_1}} & & 0 \\ & \boxed{J_{\lambda,\ell_2}} & \\ & & \ddots & \\ 0 & & & \boxed{J_{\lambda,\ell_k}} \end{pmatrix}$$

with $\ell_1 \geq \ell_2 \geq \ell_3 \geq \dots \geq \ell_k$

Each block has its own tower. Put k towers next to each other: pyramid:

$$\begin{array}{cccccc}
l & e_{l_1} & & & & \\
 & \vdots & e_{l_1+l_2} & & & \\
m & \vdots & \vdots & & & \\
 & \vdots & \vdots & \cdots & e_n & \\
 & \vdots & \vdots & \cdots & \vdots & \\
2 & e_2 & e_{l_1+2} & \cdots & e_{n-l_k+2} & \\
1 & e_1 & e_{l_1+1} & \cdots & e_{n-l_k+1} &
\end{array}$$

where $l_1 + \cdots + l_{k-1} = n - l_k$

Note: All basis vectors at level 1 is mapped to 0, i.e.

$$\ker(J-\lambda I_n) \leq \mathbb{C}^n = Sp(\{e_1, e_{l+1}, \cdots e_{n-l_k+1}\}) \quad ; \text{ spanned by } 1 \text{ vectors}$$

More generally for any $m = 1, 2, \cdots$

$$\ker(J-\lambda I_n)^m \leq \mathbb{C}^n = Sp(\{\text{first level } m \text{ vectors from } 1, \cdots, m\})$$

Example: In A above

$$\ker(A-2I) = Sp(\{e_1, e_4, e_7, e_9\})$$

$$\ker(A-2I)^2 = Sp(\{e_1, e_4, e_7, e_9, e_2, e_5, e_8\})$$

If we are given an arbitrary Jordan form matrix J with several distinct eigenvalues,

For each $\lambda \in \sigma(J)$, we can seperately consider the pyramid of standard basis of $\mathbb{C}^n$ which correspond to the Jordan blocks with same eigenvalue λ.

For each λ, subspace

$$\ker(J-\lambda I_n)^m \leq \mathbb{C}^n = Sp(\{\text{first level } m \text{ vectors from } 1, \cdots, m \text{ of that pyramid of } \lambda\})$$

General Procedure:

1) Evaluate characteristic polynomial $c_A(x)$ and all eigenvalues for given matrix $A \in Mat_{n\times n}(\mathbb{C})$
2) Determine spectrum $\sigma(A)$
3) Seperately for each eigenvalue $\lambda \in \sigma(A)$, compute

$$\ker(A-\lambda I_n)^m \leq \mathbb{C}^n \qquad \forall\, m = 1, 2, \cdots$$

Note:

$$\ker(A-\lambda I_n) \leq \ker(A-\lambda I_n)^2 \leq \cdots$$

4) Choose certain basis vectors and collect them all together for all $\lambda \in \sigma(A)$ will finally construct a basis $\underline{v}_1, \dots, \underline{v}_n$ of $\mathbb{C}^n$ such that

$$\forall \text{ index } j, \qquad A v_j = \sum_{i=1}^{n} J_{ij} v_i$$

where $J = (J_{ij})_{i,j=1}^{n}$ is a matrix of Jordan normal Form

In particular, if $\sigma(A) = \{\lambda\}$, one eigenvalue has basis $\underline{v}_1, \dots, \underline{v}_n$ with pyramid

$$\begin{array}{cccc}
\underline{v}_{l_1} & & & \\
\vdots & \underline{v}_{l_1+l_2} & & \\
\vdots & \vdots & \cdots & \underline{v}_n \\
\vdots & \vdots & \cdots & \vdots \\
\underline{v}_2 & \underline{v}_{l_1+2} & \cdots & \underline{v}_{n-l_k+2} \\
\underline{v}_1 & \underline{v}_{l_1+1} & \cdots & \underline{v}_{n-l_k+1}
\end{array}$$

Each application of $A - \lambda I_n$ will map a basis vector a level down

Hence $\forall m = 1, 2, \dots,$ $\ker(A - \lambda I_n)^m \subseteq \mathbb{C}^n = Sp(\{\text{first level } m \text{ vectors from } 1, \dots, m\})$

Note: In case of $\sigma(A) = \{\lambda\}$, suffices to choose

$$\underline{v}_1, \underline{v}_{l_1+l_2}, \dots, \underline{v}_n \quad \text{top row}$$

Observe $\forall m = 1, \dots$

$$\dim \ker(A - \lambda I) = \# \text{ of towers}$$

$$\dim \ker(A - \lambda I)^m - \dim \ker(A - \lambda I_n)^{m-1} = \# \text{ of basis vectors at } m\text{-level.}$$

For matrix A with several distinct eigenvalues,

1) For each $\lambda \in \sigma(A)$, find pyramid of linearly independent vectors of n

2) For this λ, $\ker(A - \lambda I)^m = Sp(\{\text{first level } m \text{ vectors from } 1, \dots, m\})$

3) Collect all vectors in pyramid corresponding to λ

4) $\underline{v}_1, \dots, \underline{v}_n$ forms Jordan basis

Let

$$\underline{v}_1 = \begin{pmatrix} p_{11} \\ \vdots \\ p_{n1} \end{pmatrix}, \dots, \underline{v}_n = \begin{pmatrix} p_{1n} \\ \vdots \\ p_{nn} \end{pmatrix}$$

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

Claim: $A = PJP^{-1}$

Since $Av_j = \sum_{i=1}^{n} J_{ij} v_i$; J is the matrix of linear transformation of A of $\mathbb{C}^n$ relative to Jordan basis $v_1, \dots, v_n$

Let $\mathcal{V} = \{\underline{v}_1, \dots, \underline{v}_n\}$

A is itself, the matrix of linear transformation of A of $\mathbb{C}^n$ relative to standard basis

$$\underline{e}_1 = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \underline{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$$

Let $\mathcal{E} = \{\underline{e}_1, \dots, \underline{e}_n\}$

Hence P is co-ordinate change matrix, by definition

$$P = C_{\mathcal{V}}^{\mathcal{E}}$$

Hence

$$A = M_{\mathcal{E}}(A) = C_{\mathcal{V}}^{\mathcal{E}} M_{\mathcal{V}}(A) (C_{\mathcal{V}}^{\mathcal{E}})^{-1} = PJP^{-1}$$

Observe: For an eigenvalue λ and its corresponding

a_λ = total number of vectors in pyramid

g_λ = number of towers of pyramid = number of elementary Jordan blocks of eigenvalue λ

m_λ = size of largest tower = maximal size of elementary Jordan block with eigenvalue λ

So

m_λ = least η such that

$$\dim(\ker(A - \lambda I)^{\eta}) = \text{algebraic multiplicity}$$

5. Four Examples

Example A

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

Finding eigenvalues, calculating characteristic polynomial

$$c_A(x) = \det(A - \lambda I) = \begin{vmatrix} -\lambda & 0 & 0 & 0 & 0 \\ 0 & -\lambda & 0 & 1 & 0 \\ 0 & 0 & -\lambda & 1 & 0 \\ -1 & -1 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & 0 & -\lambda \end{vmatrix} = 0$$

$$\Longrightarrow (-\lambda)^2 \begin{vmatrix} -\lambda & 0 & 1 \\ 0 & -\lambda & 1 \\ -1 & 1 & -\lambda \end{vmatrix} = 0$$

$$\Rightarrow (-\lambda)^2 \left(-\lambda \begin{vmatrix} -x & 1 \\ 1 & -x \end{vmatrix} + 1 \begin{vmatrix} 0 & -x \\ -1 & 1 \end{vmatrix} \right) = 0 \Longrightarrow -\lambda^5 = 0$$

So $c_A(\lambda) = -\lambda^5 = 0 \Longrightarrow \lambda = 0$

The spectrum is $\sigma(A) = \{0\}$ and 0 has algebraic multiplicity $a_0 = 5$

<u>Computing Kernels</u>:

For $\lambda = 0$: Define

$$T_\lambda = (A - \lambda I) = A$$

$$\underbrace{\ker T_\lambda}_{T_1} \subseteq \underbrace{\ker T_\lambda^2}_{T_2} \subseteq \underbrace{\ker T_\lambda^3}_{T_3} \subseteq \cdots \subseteq \mathbb{C}^5$$

$T_0 = A \Longrightarrow T_1 = \ker A = \{\vec{x} \in \mathbb{C}^5 : A\vec{x} = 0\}$.

Let $\vec{x} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix}$. Then $A\vec{x} = A\begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = 0 \Longleftrightarrow \begin{cases} 0 = 0 \\ d = 0 \\ d = 0 \\ -a + b + c = 0 \\ 0 = 0 \end{cases}$

Therefore

$$\operatorname{Ker} A = \left\{ \begin{pmatrix} a \\ b \\ a+b \\ 0 \\ e \end{pmatrix} : a, b, e \in \mathbb{C} \right\} ; \quad \dim \ker A = \text{number of free variables} = 3$$

$\Longrightarrow$ 1st row: number of elements $= \dim \ker A = 3$

$1 \quad \square \; \square \; \square$

and $T_1 = Sp\{\underline{v}_1, \underline{v}_4, \underline{v}_5\}$

Next

$$T_2 = A^2 = AA = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\operatorname{Ker} A^2 = \left\{ \begin{pmatrix} a \\ b \\ a+b \\ d \\ e \end{pmatrix} : a, b, e, d \in \mathbb{C} \right\} ; \quad \dim \ker A^2 = \text{number of free variables} = 4$$

$\Longrightarrow$ 2nd row: number of elements $= \dim \ker A^2 - \dim \ker A = 1$

$2 \quad \square$

$1 \quad \square \; \square \; \square$

and $T_2 = Sp\{\underline{v}_1, \underline{v}_4, \underline{v}_5, \underline{v}_2\}$

Next

$$T_3 = A^3 = 0 \Longrightarrow \ker A^3 = \mathbb{C}^5 \Longrightarrow \dim \ker A^3 = 5$$

$\Longrightarrow$ 3rd row: number of elements $= \dim \ker A^3 - \dim \ker A^2 = 1$

Therefore we get pyramid

$3 \quad \square$

$2 \quad \square$

$1 \quad \square \; \square \; \square$

and $T_3 = Sp\{\underline{v}_1, \underline{v}_4, \underline{v}_5, \underline{v}_2, \underline{v}_3\}$

Constructing Jordan normal form

Pyramid: for $\lambda = 0$

$$\begin{array}{llll} 3 & \boxed{v_3} & & \\ 2 & \boxed{v_2} & & \\ 1 & \boxed{v_1} & \boxed{v_4} & \boxed{v_5} \end{array}$$

$$J_{0,3} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad J_{0,1} = (0) \qquad J_{0,1} = (0)$$

Each Tower: one elementary Jordan block

size $= \ell =$ height of tower

Therefore

$$J_A = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix}} & & \\ & \boxed{0} & \\ & & \boxed{0} \end{pmatrix} = \begin{pmatrix} J_{0,3} & & \\ & J_{0,1} & \\ & & J_{0,1} \end{pmatrix}$$

Observe: For an eigenvalue λ and its corresponding

$a_\lambda =$ total number of vectors in pyramid

$g_\lambda =$ number of towers of pyramid $=$ number of elementary Jordan blocks of eigenvalue λ

$m_\lambda =$ size of largest tower $=$ maximal size of elementary Jordan block with eigenvalue λ

Finding matrix P, starting from top of pyramid, in our case $\underline{v}_3$

Start from, choosing any $\underline{v}_3 \in \operatorname{Ker} A^3$ and $\underline{v}_3 \notin \operatorname{Ker} A^2$

$$\underline{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Note: $\underline{v}_3 \in \operatorname{Ker} A^3 \Longrightarrow \underline{v}_2 = A\underline{v}_3 \Longrightarrow A\underline{v}_2 = \underline{v}_1$

$\underline{e}_2$ is the image of $\underline{e}_3$ etc

$$\begin{array}{c} \underline{v}_3 \\ \downarrow \\ \underline{v}_2 \\ \downarrow \\ \underline{v}_1 \end{array} \qquad \begin{array}{l} \underline{v}_2 = (A - \lambda I)\underline{v}_3 \\ \underline{v}_1 = (A - \lambda I)\underline{v}_2 \end{array}$$

Hence

$$\underline{v}_2 = A\underline{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \implies \underline{v}_1 = A\underline{v}_2 = \begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

To construct $\underline{v}_4$ and $\underline{v}_5$, observe

$$T_1 = \ker A = sp\{\underline{v}_1, \underline{v}_4, \underline{v}_5\} \implies \underline{v}_1, \underline{v}_4, \underline{v}_5 \text{ linearly independent and}$$

$$\underline{v}_4, \underline{v}_5 \in \ker A$$

$$\implies \underline{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad \underline{v}_5 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Therefore

$$P = (\underline{v}_1 \ \ \underline{v}_2 \ \ \underline{v}_3 \ \ \underline{v}_4 \ \ \underline{v}_5) = \begin{pmatrix} 0 & 0 & 1 & 0 & 2 \\ -1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$A = PJ_AP^{-1}$$

Example B

$$B = \begin{pmatrix} 10 & -4 & 0 \\ 1 & 5 & 9 \\ -1 & 1 & 9 \end{pmatrix}$$

<u>1st step:</u> Calculating eigenvalues

$$c_B(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 10-\lambda & -4 & 0 \\ 1 & 5-\lambda & 9 \\ -1 & 1 & 9-\lambda \end{vmatrix}$$

$$= (10-\lambda)\begin{vmatrix} 5-\lambda & 9 \\ 1 & 9-\lambda \end{vmatrix} - (-4)\begin{vmatrix} 1 & 9 \\ -1 & 9-\lambda \end{vmatrix}$$

$$= (10-\lambda)\left((5-\lambda)(9-\lambda) - 9\right) + 4(9-\lambda+9)$$

$$= (10-\lambda)(45 - 5\lambda - 9\lambda + \lambda^2 - 9) + 4(18-\lambda)$$

$$= 450 - 50\lambda - 90\lambda + 10\lambda^2 - 90 - 45\lambda + 5\lambda^2 + 9\lambda^2 - \lambda^3 + 9\lambda + 72 - 4\lambda$$

$$= -\lambda^3 + 24\lambda^2 - 180\lambda + 432$$

$$= -(\lambda^3 - 24\lambda^2 + 180\lambda - 432)$$

$$= -(\lambda-6)^2(\lambda-12)$$

$$c_B(\lambda) = 0 \implies \lambda = 6, \quad \lambda = 12$$

Algebraic multiplicities are

$$\left.\begin{aligned} &1)\ \lambda = 6: \quad a_6 = 2 \\ &2)\ \lambda = 12: \quad a_{12} = 1 \end{aligned}\right\} \implies \sigma(\lambda) = \{6, 12\}$$

<u>2nd step:</u> Finding Kernels and constructing pyramids for each λ

I) $\lambda = 6$:

$$T = B - 6I = \begin{pmatrix} 4 & -4 & 0 \\ 1 & -1 & 9 \\ -1 & 1 & 3 \end{pmatrix}$$

Finding $\operatorname{Ker}(B - 6I) = \left\{ \underline{x} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} : (B - 6I)\underline{x} = \underline{0} \right\}$

We can use row echelon method to solve for a, b, c

$$\left(\begin{array}{ccc|c} 4 & -4 & 0 & 0 \\ 1 & -1 & 9 & 0 \\ -1 & 1 & 3 & 0 \end{array}\right) \overset{r_1/4}{\sim} \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 1 & -1 & 9 & 0 \\ -1 & 1 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\Longrightarrow \begin{cases} a-b=0 \\ c=0 \\ 0+0+0=0 \end{cases} \Longrightarrow \begin{cases} a=b \\ c=0 \\ 0 \end{cases}$$

Hence

$$\operatorname{Ker}(B-I6) = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} : a=b,\ c=0 \right\} = \left\{ \begin{pmatrix} b \\ b \\ 0 \end{pmatrix} : b \in \mathbb{C} \right\}$$

$g_6 = \dim(\ker T) =$ number of free variables $= 1 \Longrightarrow 1$ tower

$\Longrightarrow$ 1st row: number of elements $= \dim \ker T = 1$

$1 \quad \square$

Observe:

$a_6 = 2 \Longrightarrow$ number of basis vectors $= 2$ and hence pyramid has form

pyramid:

$2 \quad \boxed{\underline{v}_2}$

$1 \quad \boxed{\underline{v}_1}$

$$J_{1,6} = \begin{pmatrix} 6 & 1 \\ 0 & 6 \end{pmatrix}$$

Finding $\underline{v}_1$ and $\underline{v}_2$, computing $\operatorname{Ker}(B-6I)^2$

$$T^2 = (B-6I)^2 = \begin{pmatrix} 12 & -12 & -36 \\ -6 & 6 & 18 \\ -6 & 6 & 18 \end{pmatrix}$$

Reducing to row-echelon form