06.10.2015
Tuesday

DEFINITION: A field is a set F together with two mappings called addition and multiplication shown as:

$$\oplus : F \times F \longrightarrow F$$
$$\odot : F \times F \longrightarrow F$$

which have properties

(A1) $a \oplus b = b \oplus a \quad \forall a,b \in F$ (commutativity)

(A2) $a \oplus (b \oplus c) = (a \oplus b) \oplus c \quad \forall a,b,c \in F$ (associativity)

(A3) There is an element in F, denoted by 0_F such that

$$a \oplus 0_F = a \quad \forall a \in F \quad \text{(additive identity)}$$

(A4) For each $a \in F$, there is an element in F denoted as $-a$ such that

$$a \oplus -a = 0_F \quad \forall a \in F \quad \text{(additive inverse)}$$

(M1) $a \odot b = b \odot a \quad \forall a,b \in F$ (commutativity)

(M2) $a \odot (b \odot c) = (a \odot b) \odot c \quad \forall a,b,c \in F$ (associativity)

(M3) $\exists\ 1_F$ s.t

$$a \odot 1_F = a \quad \forall a \in F \quad \text{(multiplicative identity)}$$

(M4) For each $a \neq 0_F$, $\exists$ an element in F denoted by a^{-1} s.t

$$a \odot a^{-1} = 1_F \quad \text{(multiplicative inverse)}$$

(D) The operations $\oplus$ and $\odot$ satisfy

$$a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c) \quad \forall a,b,c \quad \text{(Distributive law)}$$

EXAMPLE: Set of real numbers $\mathbb{R}$ with standard addition and multiplication operations.

$$F = \mathbb{R}$$

EXAMPLE: Let $F = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ The operation $\oplus$ and $\odot$ are defined as

$$x = (x_1, x_2) \in F$$
$$y = (y_1, y_2) \in F$$

$$x \oplus y = (x_1 + y_1 \, , \, x_2 + y_2)$$

$$x \odot y = (x_1 y_1 - x_2 y_2 \, , \, x_2 y_1 + x_1 y_2)$$

$$(x_1 + j x_2) \odot (y_1 + j y_2) = x_1 y_1 - x_2 y_2 + j(x_2 y_1 + x_1 y_2)$$

$$\left.\right\} \quad F = \mathbb{C} \quad 0_F = (0,0) \quad 1_F = (1,0)$$

EXAMPLE: Let $F = \mathbb{R}$

$$\left.\begin{array}{l} x \oplus y = xy \\ x \odot y = x + y \end{array}\right\} \quad \text{Is } F \text{ with } \oplus, \odot \text{ a field?}$$

* $x \oplus y = y \oplus x$
* $0_F = 1$
* $-x = 1/x$
* $1_F = 0$
* $x^{-1} = -x$
* $x \odot (y \oplus z) \overset{?}{=} (x \odot y) \oplus (x \odot z)$

$$x + yz \overset{?}{\neq} (x+y)(x+z)$$

No, F with $\oplus, \odot$ is not a field because $\oplus$ & $\odot$ do not satisfy distributive law.

$\left(\frac{1}{x^2+1}\right)$

- Are polynomials a field? No $\rightarrow$ you cannot set up field out of poly.
- Are matrices a field? No $\rightarrow$ "

09. 10. 2015
Friday

Review

$F, \oplus, \odot$

$x, y, z \in F$

$\oplus$

$x \oplus y = y \oplus x$

$x \oplus (y \oplus z) = (x \oplus y) \oplus z$

$\exists 0_F \quad x \oplus 0_F = x$

$\exists -x \in F \quad x \oplus -x = 0_F$

$\odot$

$x \odot y = y \odot x$

$x \odot (y \odot z) = (x \odot y) \odot z$

$\exists 1_F \in F \quad x \odot 1_F = x$

$\exists x^{-1} \in F \quad x \odot x^{-1} = 1_F$

$\oplus, \odot$

$x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$

called as a FIELD

$F = \mathbb{R}$

$F = \mathbb{C}$ → complex number

F = rational number

$F \neq$ integer

Polynomial? → not a field $\left(\frac{1}{x+1}\right)$

Matrices? → not a field → we don't always have inverse of Matrices

LINEAR (VECTOR) SPACES

DEFINITION: A linear space V is a set whose elements are called vectors, associated with a field F, whose elements are called scalars, and two operations: addition $\oplus_V$ and scalar multiplication $\odot_V$.

(V, F)

V → vectors, F → scalars

The following axioms should hold for the addition and scalar multiplication.

A1) $x \oplus y = y \oplus x \quad \forall x, y \in V$ (commutativity)

A2) $(x \oplus y) \oplus z = x \oplus (y \oplus z) \quad \forall x, y, z \in V$ (associativity)

A3) $\exists$ a vector 0_v s.t

$$x \oplus 0_v = x \quad \forall x \in F \quad \text{(additive identity)}$$

A4) For every element $x \in V$ $\exists$ a unique element $-x$ st

$$x \oplus -x = 0_v \quad \text{(additive inverse)}$$

SM1) $a \odot_v (b \odot_v x) = (a \odot_F b) \odot_v x \quad \forall x \in V \quad \forall a, b \in F$

SM2) $a \odot (x \oplus y) = (a \odot x) \oplus (a \odot y) \quad \forall x, y \in V \quad \forall a \in F$

SM3) $(a \oplus_F b) \odot_v x = (a \odot_v x) \oplus_v (b \odot_v x) \quad x \in V \quad a, b \in F$

SM4) With 1_F being the multiplicative identity of the field F, we have

$$1_F \odot x = x \quad \forall x \in V$$

EXAMPLE: Show that

$$0_F \odot x = 0_V \quad \forall x \in V$$

$\oplus_v : V \times V \to V$

$\odot_v : F \times V \to V$

$$0_F \odot x = (0_F \oplus 0_F) \odot x$$

$$\underbrace{0_F \odot x}_{\in V} = \underbrace{(0_F \odot x)}_{\in V} \oplus \underbrace{(0_F \odot x)}_{\in V}$$

$\exists -(0_F \odot x)$

Add both sides $-(0_F \odot x)$

$$(0_F \odot x) \oplus -(0_F \odot x) = (0_F \odot x) \oplus (0_F \odot x) \oplus -(0_F \odot x)$$

$$0_V = (0_F \odot x) \oplus 0_V = 0_F \odot x$$

$$\boxed{0_F \odot x = 0_V} \quad \text{QED}$$

EXAMPLE: Linear space $(\underset{\substack{\downarrow \\ V}}{F^n}, \underset{\downarrow}{F})$

Set of vectors

i.e, the linear space of n-tuple in F over the field F.

$x = (x_1, x_2, \ldots\ldots, x_n)$ where $x_i \in F \quad \forall i$

$y = (y_1, y_2, \ldots\ldots, y_n)$ " y_i " "

Addition $\oplus$

$$x \underset{V}{\oplus} y = (x_1 \underset{F}{\oplus} y_1, x_2 \underset{F}{\oplus} y_2, \ldots\ldots, x_n \underset{F}{\oplus} y_n)$$

Scalar Multiplication $\odot$

$$a \underset{V}{\odot} x = (a \underset{F}{\odot} x_1, a \underset{F}{\odot} x_2, \ldots\ldots, a \underset{F}{\odot} x_n)$$

The most common examples are $\mathbb{R}^n$ & $\mathbb{R}^n$

EXAMPLE: The set of all real-valued functions $t \to f(t)$ defined over the real time.

$F : \mathbb{R}$

V: real valued functions

addition:

$$(f \oplus g)(t) = f(t) + g(t) \quad \forall t \in \mathbb{R}$$

scalar mult:

$$(af)(t) = a\, f(t) \quad \forall t \in \mathbb{R}$$

EXAMPLE: Set of all polynomials of degree $n=2$

$$V = \{ a_0 t^2 + a_1 t + a_2 \;/\; \underset{a_0 \neq 0}{a_0, a_1, a_2 \in \mathbb{R}} \}$$

$F = \mathbb{R}$

$$(2t^2 + t + 3) \oplus (-2t^2 + 2t + 4) = 3t + 7 \notin V$$

NOT A VECTOR SPACE

EXAMPLE: Set of all polynomials with degree $n \leq 2$

THIS IS VECTOR SPACE

DEFINITION: Let V be a vector space over the field F. A subset W of V is a subspace of V $\underbrace{\text{if and only if}}_{\text{iff}}$ (W, F) is itself a vector space.

$V = \mathbb{R}^2$

$$W = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \;/\; a \in \mathbb{R} \right\}$$

$W \in V$

We can check whether a subset W is a subspace or not by checking only the following:

(S1) $w_1 + w_2 \in W \quad \forall\, w_1 \,\&\, w_2 \in W$

(S2) $a\, w_1 \in W \quad \forall\, w_1 \in W \quad \forall\, a \in F$

Subset has to be closed under addition and scalar multiplication.

All other properties are inherited from the original linear space

Prove that $W = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \middle/ a \in \mathbb{R} \right\}$ is a subspace of $V = \mathbb{R}^2$

$x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \in W$

$y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \in W$

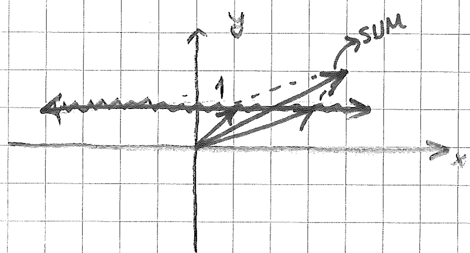
* $x + y = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ 0 \end{bmatrix} \in W$ ✓

* $ax = \begin{bmatrix} ax_1 \\ 0 \end{bmatrix} \in W$ ✓

$\Rightarrow$ W is a subspace

EXAMPLE: $V = \mathbb{R}^2$

$W = \left\{ \begin{bmatrix} a \\ 1 \end{bmatrix} \middle/ a \in \mathbb{R} \right\}$



$x = \begin{bmatrix} x_1 \\ 1 \end{bmatrix}$ & $y = \begin{bmatrix} y_1 \\ 1 \end{bmatrix}$

$x + y = \begin{bmatrix} x_1 + y_1 \\ 2 \end{bmatrix} \notin W$

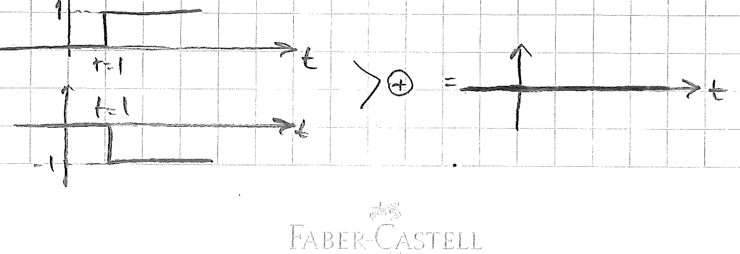
$\therefore$ W is NOT a subspace

EXAMPLE: V: set of real valued functions.

W_1: set of all continuous functions.

W_1 is a subspace

W_2: set of all real valued functions which are discontinuous at $t=1$.



$\Rightarrow$ W is NOT subspace

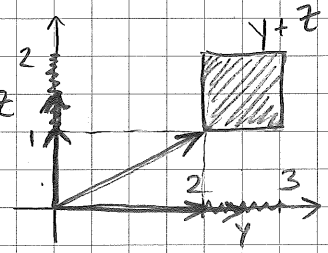
DEFINITION: The sum $Y+Z$ of two subsets Y & Z of a linear space X is the set of all vectors $y+z$ where $y \in Y$ & $z \in Z$.

$$Y+Z \triangleq \{ y+z \mid y \in Y,\ z \in Z \}$$

EXAMPLE: $X = \mathbb{R}^2$

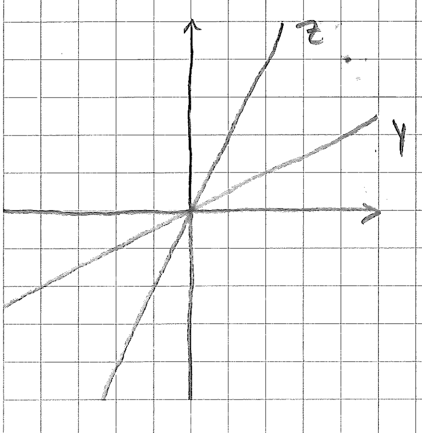
$$Y = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \Big/ \ 2 \leq a \leq 3 \right\}$$

$$Z = \left\{ \begin{bmatrix} 0 \\ b \end{bmatrix} \Big/ \ 1 \leq b \leq 2 \right\}$$



$$Y+Z = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \Big/ \ \begin{matrix} 2 \leq a \leq 3 \\ 1 \leq b \leq 2 \end{matrix} \right\}$$

$V = \mathbb{R}^2$



$Y+Z = \mathbb{R}^2$

* 0 vector is itself subspace & it is the smallest subspace.

EXAMPLE: Show that $Y+Z$ is a subspace if both Y & Z are subspaces.

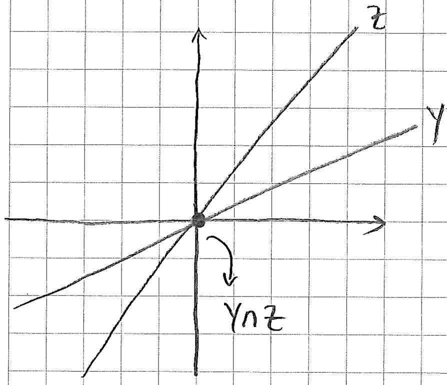
$$x_1 \in Y+Z \Rightarrow x_1 = y_1 + z_1 \text{ where } y_1 \in Y \quad z_1 \in Z$$

$$x_2 \in Y+Z \Rightarrow x_2 = y_2 + z_2 \text{ where } y_2 \in Y \quad z_2 \in Z$$

$$x_1 + x_2 = \underbrace{y_1 + y_2}_{\in Y} + \underbrace{z_1 + z_2}_{\in Z} \Rightarrow x_1 + x_2 \in Y+Z$$

$$a x_1 = \underbrace{a y_1}_{\in Y} + \underbrace{a z_1}_{\in Z} \Rightarrow a x_1 \in Y+Z$$

EXAMPLE:



Show that if Y & Z are subspaces then so is their intersection $Y \cap Z$.

$x_1, x_2 \in Y \cap Z$

$x_1 \in Y \qquad x_1 \in Z$

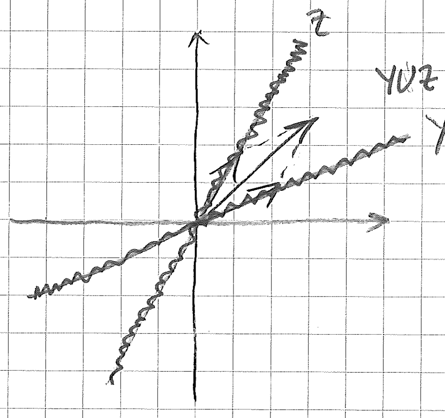
$x_2 \in Y \qquad x_2 \in Z$

$x_1 + x_2 \in Y \qquad x_1 + x_2 \in Z \implies x_1 + x_2 \in Y \cap Z$

$a x_1 \in Y \qquad a x_1 \in Z \implies a x_1 \in Y \cap Z$

$\implies$ $Y \cap Z$ is a SUBSPACE

EXAMPLE: If Y & Z are subspaces, Is $Y \cup Z$ a subspace?



$Y \cup Z$ does NOT have to be a subspace in general.

DEFINITION: (Product Space)

Let (V, F) and (W, F) be two linear space defined over the same field F. The product space of (V, F) & (W, F) is defined as

$$V \times W \triangleq \{ (v, w) \mid v \in V, w \in W \}$$

$(v_1, w_1), (v_2, w_2) \in V \times W$

vector addition

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$$

scalar multiplication

$$a \in F \quad a(v_1, w_1) = (av_1, aw_1)$$

EXAMPLE:

$$\underbrace{(\mathbb{R}^2}_{V \times W}, \underset{F}{\mathbb{R}}) = (\underset{V}{\mathbb{R}}, \underset{F}{\mathbb{R}}) \times (\underset{W}{\mathbb{R}}, \underset{F}{\mathbb{R}})$$

$x_1 + x_2$

$a x_1$

13. 10. 2015
Tuesday

Review: $F \oplus \odot$

(V, F)

$\oplus \quad V \times V \to V$

$\odot \quad F \times V \to V$

$W \subset V$

$w_1, w_2 \in W$

$w_1 + w_2 \in W$

$\underset{\in F}{a} w_1 \in W$

W_1, W_2 are subspaces

$W_1 + W_2$ is a subspace

$W_1 \cap W_2$ " "

$W_1 \cup W_2$ is general no a subspace

$W_1 \times W_2$

$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$

DEFINITION: A linear combination of n-vectors $x_1, x_2, \dots, x_n$ of a linear space V is a vector of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = \sum_{i=1}^{n} a_i x_i$$

where $a_i \in F$ for $i = 1, 2, \dots, n$

The set of all linear combinations of $x_1, x_2, \dots, x_n$

i.e $\left\{ \sum_{i=1}^{n} a_i x_i \;/\; a_i \in F \;,\; i = 1, 2, \dots, n \right\}$

is called the span of $\{x_1, x_2, \dots, x_n\}$

We denote span of $x_1, \dots, x_n$ as $\text{Span}\{x_1, \dots, x_n\}$

* $x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$\text{Span}\{x_1, x_2\} = \left\{ a_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} \;/\; a_1, a_2 \in \mathbb{R} \right\}$$

$$= \left\{ \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \;/\; a_1, a_2 \in \mathbb{R} \right\} = \mathbb{R}^2$$

* $x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\text{Span}\{x_1, x_2, x_3\} = \mathbb{R}^2 \rightarrow$ in this case, x_3 is linearly dependent with x_1 & x_2.

* Every set is subspace of itself !

EXAMPLE: Show that $\text{span}\{x_1, \ldots, x_n\}$ is a subspace.

$$w_1, w_2 \in \text{span}\{x_1, \ldots, x_n\}$$

$$w_1 = \sum_{i=1}^{n} a_i x_i \qquad a_i \in F$$

$$w_2 = \sum_{i=1}^{n} b_i x_i \qquad b_i \in F$$

$$w_1 + w_2 = \sum_{i=1}^{n} \underbrace{(a_i + b_i)}_{\in F} x_i \quad \in \text{span}\{x_1, \ldots, x_n\}$$

$c \in F$

$$c\, w_1 = \sum_{i=1}^{n} \underbrace{(c a_i)}_{\in F} x_i \quad \in \text{span}\{x_1, \ldots, x_n\}$$

Definition: Vectors $x_1, \ldots, x_n \in V$ are said to be linearly independent if

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = 0$$

implies

$$a_1 = a_2 = \ldots = a_n = 0$$

Otherwise, they are called linearly dependent.

ie if $\exists\ a_1, a_2, \ldots, a_n$ not all zero

for which $a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = 0$

then $x_1, \ldots, x_n$ are linearly dependent.

EXAMPLES:

$$X_1 = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

$$a_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow a_1 = a_2 = 0$$

linear independence

$$X_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\} \qquad a_1 = 1 \quad a_2 = -2 \quad a_3 = 1$$

$$\begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = 2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad a_1 \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} + a_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + a_3 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 0$$

X_2 is linearly dependent

EXAMPLE: Consider linear space of polynomials with degree ≤ 2.

Let $P = \{ p_1(t), p_2(t), p_3(t) \}$ be a set of polynomials

$$p_1(t) = 1$$
$$p_2(t) = 1 + t$$
$$p_3(t) = 1 + t + t^2$$

Is P linearly independent?

$$a_1 \cdot 1 + a_2 (1+t) + a_3 (1 + t + t^2) = 0(t)$$
$$= 0 + 0t + 0t^2 \quad \forall t$$

$$(a_1 + a_2 + a_3) 1 + (a_2 + a_3) t + a_3 t^2 = 0$$

$\Rightarrow a_3 = 0 \quad a_2 = 0 \quad a_1 = 0$

$\Rightarrow$ linear independence

EXAMPLE: V: set of continuous functions

$$S = \{\cos t, \sin t, \cos(t - \pi/3)\}$$

Is S linearly ind?

$$a_1 \cos t + a_2 \sin t + a_3 \cos(t - \pi/3) = 0 \quad \forall t$$

$$\cos(t - \pi/3) = \cos t \cos \pi/3 + \sin t \sin \pi/3$$
$$= \frac{1}{2}\cos t + \frac{\sqrt{3}}{2}\sin t$$

$$\frac{1}{2}\cos t + \frac{\sqrt{3}}{2}\sin t - 1 \cos(t - \pi/3) = 0 \quad \forall t$$

linearly dependent

DEFINITION: Let V be a linear space and the set of vectors $B = \{x_1, \ldots, x_n\}$ be a subset of V.

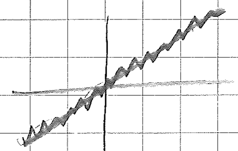
B is said to be a basis for V iff

i) $\text{span}\{B\} = V$

ii) B is a linearly independent set

EXAMPLE: $V = \mathbb{R}^2$

$B_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ $\quad a_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0 \quad B_1 \neq \mathbb{R}^2$



$$\boxed{B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}}$$ B_2 is a basis for $\mathbb{R}^2$

$B_3 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ not ind.

$$B_4 = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \quad \checkmark$$

REMARK: There are infinitely many basis sets for a linear space V.
All of these basis sets have the same number of elements.
That number is called the dimension of the V.
If the dimension is finite, V is said to be finite dimensional.
Otherwise it is called infinite dimensional.

16.10.2015
Friday

$x_1, x_2, \ldots, x_n \in V$

$a_1, a_2, \ldots, a_n \in F$

$$\sum_{i=1}^{n} a_i x_i$$

$$\underbrace{\left\{ \sum_{i=1}^{n} a_i x_i \;/\; a_i \in F \quad i = 1, \ldots, n \right\}}_{\text{span}\{x_1, \ldots, x_n\}}$$

$$\sum_{i=1}^{n} a_i x_i = 0_V \Rightarrow a_i = 0_F \quad \forall i$$

$$B = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \qquad a_1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$a_1 \downarrow 1 \qquad a_2 \downarrow 0$

$\Rightarrow$ B is linearly dependent

* If it contains zero vector, it cannot be independent!

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \qquad a_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_1 + a_2 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow a_1 = 0 \Rightarrow a_2 = 0$$

B is linearly independent

V

$B = \{x_1, \dots, x_n\}$

- $V = \text{span}\{B\}$
- B is lin. ind.

If we satisfy these two condition, then we can say that "B is a basis for V"

$V = \mathbb{R}^2$

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Both are basis for V

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

\# of elements in B is called Dimension of V

DEFINITION: If $y \in V$ and $B = \{x_1, \dots, x_n\}$ is a basis set for V, then there is a unique n-tuple of scalars $(a_1, \dots, a_n)$ such that

$$y = \sum_{i=1}^{n} a_i x_i$$

The scalars $(a_1, \dots, a_n)$ are called the components/coordinates/representations of y with respect the basis B. We show the representation of y wrt B as vector denoted as $[y]B$.

$$[y]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

EXAMPLE: $V=\mathbb{R}^2$ $F=\mathbb{R}$

* ~~$V=\mathbb{R}^2$ $F=\mathbb{C}$~~ → because scalar mult. gives me complex vector space which is not inside of $\mathbb{R}^2$. So, it is NOT possible such combination.

$$B=\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right\}$$

$$y=\begin{bmatrix}2\\3\end{bmatrix} \qquad [y]_B=?$$

$$a_1\begin{bmatrix}1\\0\end{bmatrix}+a_2\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}2\\3\end{bmatrix} \Rightarrow \begin{array}{l} a_1+a_2=2 \\ a_2=3 \Rightarrow a_1=-1\end{array}$$

$$[y]_B=\begin{bmatrix}-1\\3\end{bmatrix}$$

EXAMPLE:

$$V=\text{span}\{\cos t,\ \sin t\}$$

$$B=\{\cos t,\ \sin t\}$$

$$y \triangleq \cos(t-\pi/3)$$

$$y=\frac{1}{2}\cos t+\frac{\sqrt{3}}{2}\sin t$$

$$[y]_B=\begin{bmatrix}1/2\\ \sqrt{3}/2\end{bmatrix}$$

$$\|$$

$$\left[\cos(t-\pi/3)\right]_B=\begin{bmatrix}1/2\\ \sqrt{3}/2\end{bmatrix}$$

Claim: For a given basis $\mathcal{B}$, the representation $[y]_{\mathcal{B}}$ of a vector y is unique.

Proof: Proof by contradiction.

$$p \Rightarrow q$$

If A is a car $\Rightarrow$ A has doors

$$\bar{q} \Rightarrow \bar{p}$$

If A has no doors $\Rightarrow$ A is NOT a car

$p \to$ your assumption

$q \to$ conclusion in the theorem

Let us assume that $[y]_{\mathcal{B}}$ is not unique $(\bar{q})$

$\Rightarrow$ There exists

$$[y]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad [y]_{\mathcal{B}} = \begin{bmatrix} a_1' \\ a_2' \\ \vdots \\ a_n' \end{bmatrix}$$

and $\exists\, i$ s.t $a_i \neq a_i'$

$$\left.\begin{aligned} y &= \sum_{i=1}^{n} a_i x_i \\ y &= \sum_{i=1}^{n} a_i' x_i \end{aligned}\right\} \quad 0 = \sum_{i=0}^{n} (a_i - a_i') x_i$$

$\exists\, i$ s.t $a_i - a_i' \neq 0$

$\Rightarrow x_1, \dots, x_n$ is NOT linearly independent.

$\Rightarrow \mathcal{B}$ is NOT a basis $\Rightarrow$ CONTRADICTION!

EXAMPLE: Let $\mathbb{R}^{2\times2}$ denote the linear space of real valued 2x2 matrices with the standard matrix addition and scalar multiplication operations.

$\mathbb{R}^{2\times2}$

Let the canonical basis be:

$$B = \left\{ \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{x_1}, \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{x_2}, \underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{x_3}, \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}_{x_4} \right\}$$

Hence if we have a matrix $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$[v]_B = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

$$V = ax_1 + bx_2 + cx_3 + dx_4$$

Theorem: $V = \mathbb{R}^2$

$$V = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

$$\left. \begin{array}{l} [V]_{B_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \\ [V]_{B_2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{array} \right\} \quad [V]_{B_1} = P \; [V]_{B_2} \quad \forall V$$

Theorem: Let V be an n-dimensional linear space over $\mathbb{R}$. Let B_1 and B_2 be two basis sets for V. Then there exits an $n \times n$ real invertible matrix P such that

$$[v]_{B_1} = P[v]_{B_2} \quad \forall v \in V$$

Proof: Proof by construction (directly showing)

$$p \Rightarrow q$$

we call the coefficient as p_{ij}

$$B_1 = \{v_1, \dots, v_n\} \qquad w_j \in V$$

$$B_2 = \{w_1, \dots, w_n\} \qquad w_j = \sum_{i=1}^{n} p_{ij} v_i \quad \text{for } j = 1, \dots, n$$

Consider an arbitrary vector $v \in V$

$$v = \sum_{j=1}^{n} \alpha_j w_j \Rightarrow [v]_{B_2} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$v = \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{n} p_{ij} v_i = \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_j p_{ij} v_i$$

$$= \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{n} \alpha_j p_{ij} \right)}_{\alpha_i'} v_i = \sum_{i=1}^{n} \alpha_i' v_i$$

$$\Rightarrow [v]_{B_1} = \begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_n' \end{bmatrix} \qquad \Rightarrow \alpha_i' \triangleq \sum_{j=1}^{n} p_{ij} \alpha_j$$

$$\alpha_1' = \sum_{j=1}^{n} p_{1j}\,\alpha_j$$

$$\begin{bmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_n' \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

<u>EXAMPLE:</u> Consider $V=$ polynomials with degree ≤ 2.

$$B_1 = \{ \overset{v_1}{1},\ \overset{v_2}{1+t},\ \overset{v_3}{1+t+t^2} \}$$

$$B_2 = \{ \overset{w_1}{1},\ \overset{w_2}{t},\ \overset{w_3}{t^2} \}$$

Find the matrix P which converts representations wrt B_2 into rep. wrt B_1.

$$[v]_{B_1} = P\,[v]_{B_2}$$

$$P = ?$$

Choose $v = w_1$

$$\begin{array}{cccc} \text{for } w_1 & \text{for } w_2 & \text{for } w_3 & \\ \end{array}$$

$$[w_1]_{B_1} = \begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$1 = 1\,(1) + 0\,(1+t) + 0\,(1+t+t^2)$

$t = -1\,(1) + 1\,(1+t) + 0\,(1+t+t^2)$

$t^2 = 0\,(1) + -1\,(1+t) + 1\,(1+t+t^2)$

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

NORMED SPACES

DEFINITION: Let (V,F) be a vector space. A norm on V is a function

$$\|\cdot\| : V \to \underbrace{\mathbb{R}_{\geq 0}}_{\substack{\text{non-negative} \\ \text{real numbers}}}$$

satisfying the following properties:

(P1) $\|v\| \geq 0 \quad \forall v \in V$

$\|v\| = 0 \iff v = 0_V$

(P2) $\|\alpha v\| = |\alpha| \, \|v\| \quad \forall \alpha \in \mathbb{R} \quad v \in V$

(P3) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$ } triangul inequality

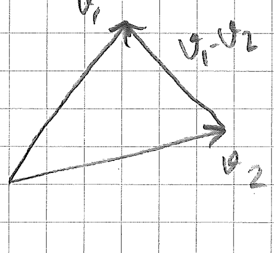
The triplet $(V, F, \|\cdot\|)$ is called a normed space.

REMARK:

1) The vector space is an algebraic structure and with the norm, we start introducing geometry into a linear space.

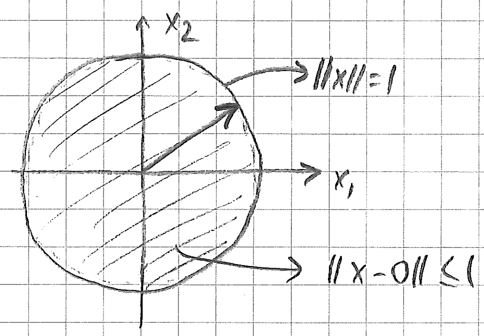
2) Norm defined a distance between two vectors $v_1, v_2 \in V$ with the expression $\|v_1 - v_2\|$

Since $\|v\| = \|v - 0_V\|$, $\|v\|$ is the distance of v to the origin.



$\|v_1 - v_2\|$ = distance

$V=\mathbb{R}^2$ $\quad x=\begin{bmatrix}x_1\\x_2\end{bmatrix}$ $\quad \|x\|=\sqrt{x_1^2+x_2^2}$

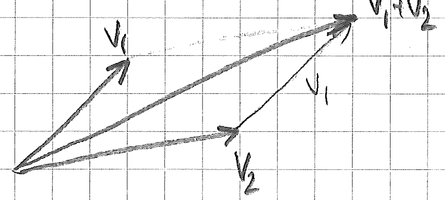


3) We can define a "sphere" in V using the norm concept.

$$S=\left\{v\in V \mid \|v-v_0\|\le r\right\}$$

S is a sphere with center v_0 and radius r.

4) Interpretation of the triangle inequality



$$\underbrace{\|v_1+v_2\|}_{\text{length of one side}} \le \underbrace{\|v_1\|+\|v_2\|}_{\text{length of other sides}}$$

20.10.2015
Tuesday

$x\in V \qquad B=\{x_1,\dots,x_n\}$

$$x=\sum_{i=1}^{n} \underbrace{a_i}_{\text{unique coefficients}} x_i \qquad a_i\in F$$

$$[x]_B=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}$$

$x \in V \quad B_1 = \{x_1, \dots, x_n\}$

$B_2 = \{y_1, \dots, y_n\}$

$$[x]_{B_1} = P \, [x]_{B_2}$$

is an invertable

$$\|\cdot\| : V \to \mathbb{R}_{\geq 0}$$

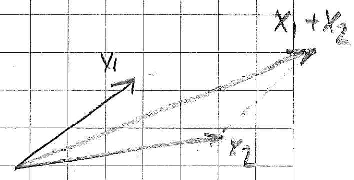
1) $\|x\| \geq 0 \quad \forall x \in V$

$\|x\| = 0 \iff x = 0_V$

2) $\|\alpha x\| = |\alpha| \, \|x\| \quad \forall x \in V \quad \forall \alpha \in F = \mathbb{R}$

3) $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\| \quad \forall x_1, x_2 \in V$

*With addition of norm function $\|\cdot\|$, two algebraic structures V & F becomes geometric also.



$(V, F, \|\cdot\|)$, $\|x_1 - x_2\|$, $\|x - x_0\| \leq r$

sphere in V
center x_0
radius r

EXAMPLE: $V = \mathbb{R}^2 \quad F = \mathbb{R}$

$$x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \qquad \alpha_1, \alpha_2 \in \mathbb{R}$$

$$\|x\|_1 \triangleq |\alpha_1| + |\alpha_2|$$

Is this function $\|\cdot\|_1$ a norm?

! Check the properties !

1) $\|x\|_1 \geq 0 \quad \forall x \in V$ ✓

Let $x = 0_V = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} \alpha_1 = 0 \\ \alpha_2 = 0 \end{matrix} \Rightarrow \|x\|_1 = |0| + |0| = 0$ (backward) ✓

Let $\|x\|_1 = 0 = |\alpha_1| + |\alpha_2| \Rightarrow \begin{matrix} \alpha_1 = 0 \\ \alpha_2 = 0 \end{matrix} \Rightarrow x = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0_V$ (forward) ✓

2) $\alpha x = \alpha \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \end{bmatrix}$

$$\|\alpha x\|_1 \triangleq |\alpha\alpha_1| + |\alpha\alpha_2| = |\alpha||\alpha_1| + |\alpha||\alpha_2| = |\alpha| \underbrace{\left[|\alpha_1| + |\alpha_2| \right]}_{\|x\|_1}$$

$$\Rightarrow \|\alpha x\|_1 = |\alpha| \, \|x\|_1 \ ✓$$

3) $x_1 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \qquad x_2 \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$

$$x_1 + x_2 = \begin{bmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \end{bmatrix}$$

$$\|x_1 + x_2\|_1 = |\alpha_1 + \beta_1| + |\alpha_2 + \beta_2| \leq \underbrace{|\alpha_1| + |\alpha_2|}_{\|x_1\|_1} + \underbrace{|\beta_1| + |\beta_2|}_{\|x_2\|_1}$$

$$\Rightarrow \|x_1 + x_2\|_1 \leq \|x_1\|_1 + \|x_2\|_1 \ ✓$$

$\Rightarrow \|\cdot\|_1$ is actually a norm.

$\boxed{\ell_1 \text{ norm}}$

FABER-CASTELL

EXAMPLE:

$$\|x\|_2 = \left[|\alpha_1|^2 + |\alpha_2|^2 \right]^{1/2}$$

2-norm
l_2-norm
Euclidean Norm

EXAMPLE:

$$\|x\|_\infty = \max \{ |\alpha_1|, |\alpha_2| \}$$

l_∞-norm

$$x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \qquad \|x\|_1 = \sum_{i=1}^{n} |\alpha_i| \qquad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}$$

$$\|x\|_\infty = \max \{ |\alpha_1|, |\alpha_2|, \dots, |\alpha_n| \}$$

These three norms can be generalized into p-norm.

$$\|x\|_p = \left(|\alpha_1|^p + |\alpha_2|^p \right)^{1/p}$$

$$x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \Rightarrow \|x\|_p = \left(\sum_{i=1}^{n} |\alpha_i|^p \right)^{1/p} \quad \text{for} \quad 1 \leq p \leq \infty$$

$$x = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad \|x\|_p = (2^p + 3^p)^{1/p}$$

$p = 1$ $\quad 2+3$

$p = 2$ $\quad 4+9$

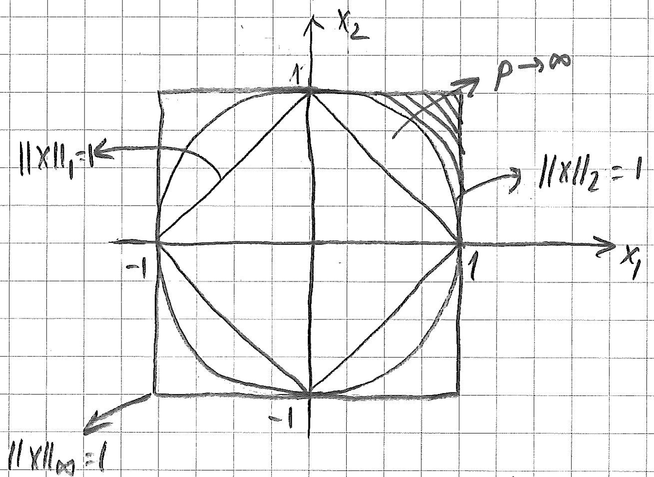
$p = 3$ $\quad 8+27$

$p \to \infty$ $\quad \cancel{2^\infty} + 3^\infty$

$$\lim_{p \to \infty} \|x\|_p = \|x\|_\infty$$

FABER-CASTELL

Consider the unit spheres $\|x\|_p \leq 1$



$$\|x\|_1 = |x_1| + |x_2| = 1$$

$$x_1 + x_2 = 1$$
$$-x_1 + x_2 = 1$$

$$\|x\|_\infty = 1$$

$$\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

$$\begin{bmatrix} -1 \\ -1/2 \end{bmatrix} \quad \begin{bmatrix} -0.5 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

EXAMPLE: $\|x\|_{1/2} = \left(|\alpha_1|^{1/2} + |\alpha_2|^{1/2} \right)^2 \qquad x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$

Is $\|x\|_{1/2}$ a norm?

Check 3rd property $\|x_1 + x_2\| \leq |x_1| + |x_2|$

$$x_1 = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \quad x_2 = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

$$\|x_1 + x_2\|_{1/2} = \left\| \begin{bmatrix} 4 \\ 9 \end{bmatrix} \right\|_{1/2} = 25$$

$$\|x_1\|_{1/2} = 4$$
$$\|x_2\|_{1/2} = 9$$

Violation in 3rd property

$\|\cdot\|_{1/2}$ is NOT a norm!

EXAMPLES: Norms on function spaces.

$$V = \left\{ f(\cdot) \;/\; f:[0,1] \to \mathbb{R} \text{ such that } \int_0^1 |f(t)|^p \, dt < \infty \quad 1 \le p \le \infty \right\}$$

We can define norms as follows:

$$\|f\|_p \triangleq \left[\int_0^1 |f(t)|^p \, dt \right]^{1/p}$$

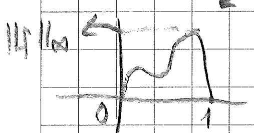
Specific Cases

$$\|f\|_1 = \int_0^1 |f(t)| \, dt \qquad L_1 - \text{norm}$$

$$\|f\|_2 = \sqrt{\int_0^1 |f(t)|^2 \, dt} \qquad L_2 - \text{norm}$$

RMS value of $f(t)$

$$\|f\|_\infty = \max_{0 \le t \le 1} |f(t)| \qquad L_\infty - \text{norm}$$



→ If there is negative value, take the absolute value of it!

23.10.2015

We defined norms on $\mathbb{R}^n$

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

$$\left(|\underset{2}{\cancel{x_1}}|^p + |\underset{3}{x_2}|^p \right)^{1/p} = |x_2| \qquad \|x\|_\infty = \max |x_i|$$

Function Norms

$$\|g\|_p = \left(\int |g(t)|^p \, dt \right)^{1/p}$$

$$\|g\|_\infty = \max |g(t)|$$

MATRIX NORMS

There are two ways to define matrix norms.

The first way is to consider the matrix as a vector.

EXAMPLE.

$$A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix} \Rightarrow a = \begin{bmatrix} 1 \\ -2 \\ 3 \\ -4 \end{bmatrix} \qquad \|A\|_\infty = 4 \qquad \|A\|_1 = 10$$

$$A = [a_{ij}]$$

$$\|A\|_\infty = \max_{i,j} |a_{ij}|$$

$$\|A\|_1 = \sum_i \sum_j |a_{ij}|$$

$$\|A\|_2 = \left(\sum_i \sum_j |a_{ij}|^2 \right)^{1/2} = \sqrt{\operatorname{trace}(A^T A)}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \rightarrow Tr(B) = b_{11} + b_{22}$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \Rightarrow A^T A = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{21}^2 & X \\ X & a_{12}^2 + a_{22}^2 \end{bmatrix}$$

$$tr A^T A = a_{11}^2 + a_{21}^2 + a_{12}^2 + a_{22}^2$$

Frobenius Norm

$$\|A\|_p = \left| \sum_i \sum_j |a_{ij}|^p \right|^{1/p} \qquad 1 \leq p \leq \infty$$

The second way to define a matrix norm is called as "induced norm".

<u>DEFINITION:</u> Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be an $m \times n$ matrix. Let $\|\cdot\|_{\mathbb{R}^n}$ & $\|\cdot\|_{\mathbb{R}^m}$ denote the norms in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. The induced matrix norm is defined as

$$\|A\| \triangleq \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}$$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \end{array} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{x\ n} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{y\ m}$$

$$\max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|y\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}$$

$$\|\alpha x\| = |\alpha| \|x\|$$

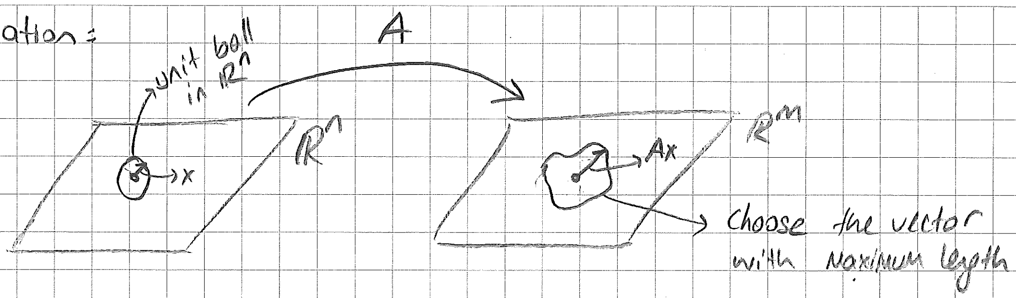
<u>Remark:</u> An equivalent definition is

$$\|A\| \triangleq \max_{\|x\|_{\mathbb{R}^n} = 1} \|Ax\|_{\mathbb{R}^m}$$

$$\frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \Big\| A \underbrace{\frac{x}{\|x\|_{\mathbb{R}^n}}}_{\text{A vector with unit norm}} \Big\|_{\mathbb{R}^m}$$

$$\frac{\|x\|}{\alpha} = \left\|\frac{x}{\alpha}\right\| \quad \alpha > 0$$

Visulation:



EXAMPLE: Consider an $m \times n$ matrix with real elements. Choose the ∞-norm $\|\cdot\|_\infty$ for both $\mathbb{R}^n$ & $\mathbb{R}^m$.

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \\ -2 & 2 \end{bmatrix} \qquad 3 \times 2 \qquad \mathbb{R}^3 \quad \mathbb{R}^2$$

$$\|A\| = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$

first row

$$\bar{A} = \begin{bmatrix} 1 & -2 \end{bmatrix} \qquad \|\bar{A}\| = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$

$$\begin{bmatrix} 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\|\bar{A}x\|_\infty = |x_1 - 2x_2|$$

$$\|x\|_\infty = 1 \rightarrow \max(|x_1|, |x_2|) = 1 \rightarrow x = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow \|\bar{A}\| = 3$$

$$\begin{bmatrix} 1 & -2 \\ -3 & 4 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$

$$\left.\begin{aligned} &\max(|x_1|, |x_2|) = 1 \\ &\|Ax\|_\infty = \max_i |y_i| \end{aligned}\right\} \quad \|A\|_\infty = 7 \qquad x^* = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$y = Ax$

$$\|A\| = \max_i \sum_{j=1}^{n} |a_{ij}| = \alpha$$

* Prove this! (our aim)

$$\left.\begin{aligned} &(\mathrm{I}) \quad \|A\| \leq \alpha \\ &(\mathrm{II}) \quad \|A\| \geq \alpha \end{aligned}\right\} \|A\| = \alpha$$

$y = Ax \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j$

Part I: $\|x\|_\infty = 1$

$$\|A\| = \max_{\|x\|_\infty = 1} \|\underbrace{Ax}_{y}\|_\infty = \max_i |y_i|$$

$$= \max_i \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq \max_i \sum_{j=1}^{n} |a_{ij} x_j| = \max_i \sum_{j=1}^{n} |a_{ij}||x_j|$$

$\max_j |x_j| = 1$

$$\leq \max \sum_{j=1}^{n} |a_{ij}| = \alpha$$

$$\Rightarrow \|A\| \leq \alpha$$

Part II: Prove that $\|A\| \geq \alpha$

$$\|A\| = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$

$\|\xi_1\|_\infty = 1$

$$\geq \|A\xi_1\|_\infty$$

Choose ξ_1 s.t $y_1 = \sum_{j=1}^{n} a_{1j}\, \xi_1(j)$

$$\xi_1 \triangleq \left[\text{sign}(a_{11}) \ \text{sign}(a_{12}) \ \ldots \ \text{sign}(a_{1n}) \right]^T$$

$$A\xi_1 = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} \ \text{sign}\, a_{1j} \\ x \\ x \\ \vdots \\ x \end{bmatrix}$$

$$A\xi_1 = \begin{bmatrix} \sum_{j=1}^{n} |a_{1j}| \\ x \\ \vdots \\ x \end{bmatrix} \qquad z = \begin{bmatrix} 3 \\ x \\ \vdots \\ x \end{bmatrix} \quad \|z\|_\infty \geq 3$$

$$\|A\| \geq \sum_{j=1}^{n} |a_{1j}|$$

Choose $\xi_2 = \left[\text{sign}\, a_{21} \ \text{sign}\, a_{22} \ \ldots \ \text{sign}_{2n} \right]^T$

$$\|\xi_2\|_\infty = 1$$

$$\|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty \geq \|A\xi_2\|_\infty$$

$$A\xi_2 = \begin{bmatrix} x \\ \sum_{j=1}^{n} |a_{2j}| \\ x \\ \vdots \\ x \end{bmatrix} \qquad \geq \sum_{j=1}^{n} |a_{2j}|$$

$$\|A\xi_2\|_\infty \geq \sum_{j=1}^{n} |a_{2j}|$$

$$\|A\| \geq \sum_{j=1}^{n} |a_{2j}|$$

continue in this manner

$$\|A\| \geq \sum_{j=1}^{n} |a_{ij}| \quad \forall i$$

$$\|A\| \geq \max_{i} \sum_{j=1}^{n} |a_{ij}| = \alpha$$

QED

<u>**EXAMPLE:**</u> Choose $\|\cdot\|_1$ norm for both $\mathbb{R}^n$ & $\mathbb{R}^m$.

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

$$\|A\|_1 = \max_{\|x\|_1 = 1} \|Ax\|$$

$$\|\bar{A}\|_1 \qquad \left| \begin{bmatrix} 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right|$$

$$\max_{\|x\|_1 = 1} |x_1 - 2x_2|$$

$$|x_1| + |x_2| = 1$$

$$\|\bar{A}\| = 2$$

$$\begin{array}{c} \begin{bmatrix} 1 & -2 \\ -3 & 4 \\ -2 & 2 \end{bmatrix} \\ \begin{array}{cc} 6 & \;\;8 \end{array} \end{array} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\|A\| = 8$$

$$\|A\| = \max_{j} \sum_{i=1}^{m} |a_{ij}|$$ *Prove this!*

Proof

CONVERGENCE

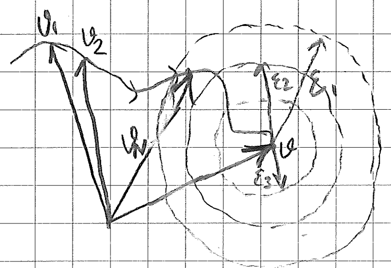
DEFINITION: Let $(V, F, \|\cdot\|)$ be a normed space. Let $\{v_n\}_{n=1}^{\infty}$ be a sequence of vectors in V i.e, $v_n \in V$ for $n=1,2,\ldots$

The sequence $\{v_n\}_{n=1}^{\infty}$ is said to converge to the limit $v \in V$ iff $\|v_n - v\| \to 0$ as $n \to \infty$

OR

If given any $\varepsilon > 0$, $\exists N$ (N can depend on ε) such that $\|v_n - v\| < \varepsilon \quad \forall n \geq N$.

find N, then sequence can converge



Usually as ε gets smaller, N gets larger.

$\varepsilon \downarrow \quad N \uparrow$

Remark: A sequence is called divergent if it is NOT convergent. If $\exists \varepsilon$ st we cannot find an N value, then the sequence is divergent.

EXAMPLE: $V = \mathbb{R} \quad \|v\| = |v|$

Consider the sequence

$$\left\{ \left(\tfrac{1}{2}\right)^n \right\}_{n=1}^{\infty}$$

Prove that the sequence is convergent to ZERO.
$v = 0$

Let $\varepsilon > 0$ be an arbitrary number. We have to find a N st $\|v_n - 0\| \leq \varepsilon \quad \forall n \geq N$

$$\|v_n\| = \left|\left(\tfrac{1}{2}\right)^n\right| = 2^{-n} \leq \varepsilon$$

$$-n \leq \log_2 \varepsilon \Rightarrow n \geq -\log_2 \varepsilon$$

$$N = \text{ceiling}\left(-\log_2 \varepsilon\right)$$

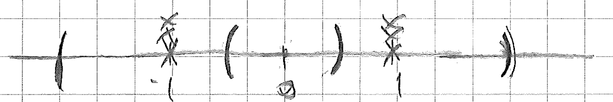
smallest integer greater than $-\log_2 \varepsilon$.

$\Rightarrow \forall n \geq N \quad \|v_n\| \leq \varepsilon$

<u>**EXAMPLE:**</u> $V = \mathbb{R} \qquad \|v\| = |v|$

Consider the sequence $\{(-1)^n\}_{n=1}^{\infty} = \{-1, 1, -1, 1, \dots\}$

Is the sequence convergent?



Suppose we want to check if $v_n \to 0$

$\varepsilon = 2 \Rightarrow N = 1 \quad \rightarrow$ our N must be hold for $\forall \varepsilon$.

$\varepsilon = 1/2$ No N can be found

$\Rightarrow$ The sequence is NOT convergent to ZERO.

Is the sequence convergent to $v = 1$.

$\varepsilon = 3 \quad N = 1$

$\varepsilon = 1$ We cannot find N.

$\Rightarrow$ The sequence is NOT convergent to 1.

$\Rightarrow$ The sequence is NOT convergent to any point in $\mathbb{R}$.

$\Rightarrow$ Sequence is divergent.

27.10.2015
Tuesday

CONVERGENCE

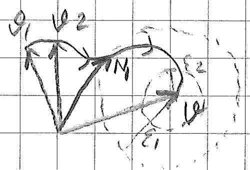
$(V, F, \|\cdot\|)$

$$\{v_n\}_{n=1}^{\infty} \quad v_n \in V$$

converges to $v \in V$ iff $\lim_{n\to\infty} \|v_n - v\| = 0$

OR

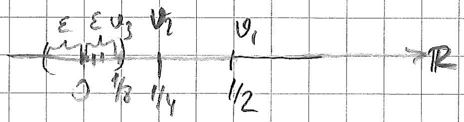
$$\forall \varepsilon > 0 \quad \exists N \quad s.t \quad \|v_n - v\| \leq \varepsilon \quad \forall n \geq N$$



EXAMPLES

$V = \mathbb{R} \quad \|v\| = |v|$

$$\left\{\left(\frac{1}{2}\right)^n\right\}_{n=1}^{\infty}$$



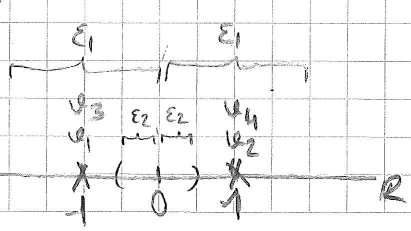
$$\|v_n - v\| = \left|\left(\tfrac{1}{2}\right)^n\right| \leq \varepsilon$$

$$\left.n = -\log_2 \varepsilon \right\} \quad N = \text{ceiling}(-\log_2 \varepsilon)$$

EXAMPLES

$$\{(-1)^n\}_{n=1}^{\infty} \quad \|v\| \triangleq |v|$$

$v_n = (-1)^n$



Is the sequence convergent to $v = 0$

$\varepsilon_1 = 2 \quad N_1 = 1$

$\varepsilon_2 = 1/2 \quad$ No N can be found!

Sequence is NOT convergent / to 0.
/ to any number.

NOTE: In many engineering applications, we are interested in the convergence of an iterative algorithm. But the problem is that we don't now where! Hence we cannot use the standard way to prove convergence. To avoid this problem, we define the concept of a "Cauchy Sequence".

DEFINITION: Let $(V, F, \|\cdot\|)$ be a normed space. A sequence $\{v_n\}_{n=1}^{\infty}$ is set to be a Cauchy Sequence if

$$\|v_n - v_m\| \to 0 \quad \text{as} \quad n, m \to \infty$$

OR

$$\forall \varepsilon > 0 \quad \exists N \text{ (depending on } \varepsilon) \text{ such that } \|v_n - v_m\| \leq \varepsilon \quad \forall n, m \geq N$$

REMARK: Every convergent sequence is a Cauchy Sequence. The reverse is NOT true.

Proof: Suppose the sequence $\{v_n\}_{n=1}^{\infty}$ is convergent and it converges to $v \in V$

$$0 \leq \|v_n - v_m\| = \|v_n - v + v - v_m\| \leq \|v_n - v\| + \|v - v_m\|$$

$$= \|v_n - v\| + \|v_m - v\| = 0 \quad \text{as } \begin{matrix} n \to \infty \\ m \to \infty \end{matrix}$$

$$\left. 0 \leq \|v_n - v_m\| \leq 0 \atop \begin{matrix} n \to \infty \\ m \to \infty \end{matrix} \right\} \Rightarrow \|v_n - v_m\| \to 0 \quad \begin{matrix} n \to \infty \\ m \to \infty \end{matrix}$$

EXAMPLE: A norm space that is defined as

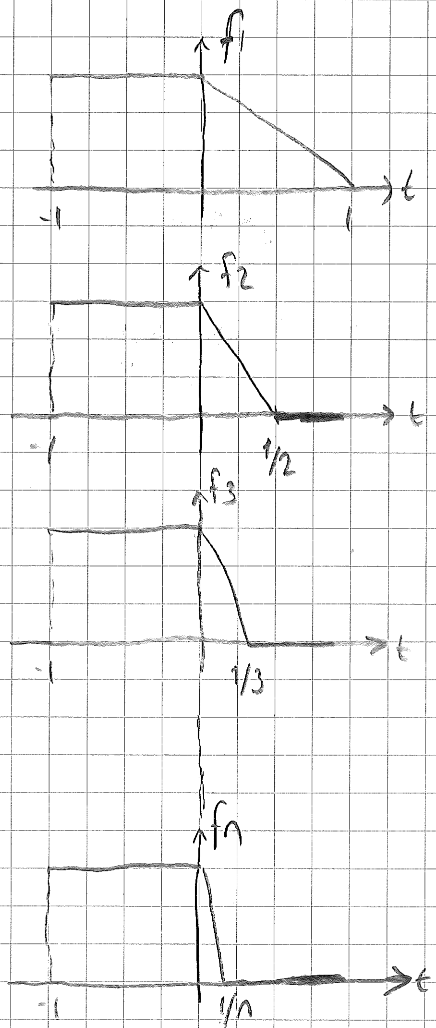
NOT COMPLETE SPACE

$$V = \left\{ f \mid f : [-1, 1] \to \mathbb{R} \ , \ f \text{ is continuous and } \int_{-1}^{1} |f(t)|\, dt < \infty \right\}$$

Define the norm as

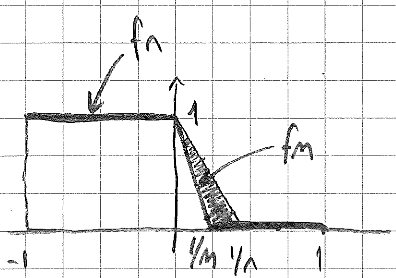
$$\|f\|_1 \triangleq \int_{-1}^{1} |f(t)|\, dt$$

Consider sequence : $\{f_n\}_{n=1}^{\infty}$



Is the sequence Cauchy?

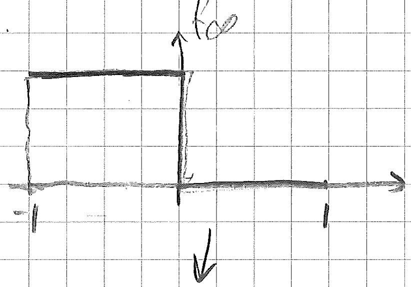
$$\|f_n - f_m\| = \int_{-1}^{1} |f_n(t) - f_m(t)|\, dt$$



$\|f_n - f_m\| = \frac{1}{2}\left|\frac{1}{n} - \frac{1}{m}\right| \to 0$ as $n, m \to \infty$

The sequence $\{f_n\}_{n=1}^{\infty}$ is Cauchy!

Where does the functions go to:



$\cancel{f_n \to f_\infty}$

$v_n \to v \in V$

It is NOT cont. $\Rightarrow$ it is not convergences

<u>DEFINITION:</u> A normed space is said to be <u>complete</u> if every cauchy sequence is convergent.

For ex, this exp is NOT complete. If we don't think cont. requirement, it will be.

A complete normed space is called as "BANACH SPACE".

$\{v_n\}_{n=1}^{\infty}$ $v_n \in V$

03.11.2015
Tuesday

v_n converges to $v \in V$ if $\|v_n - v\| \to 0$ as $n \to \infty$

OR

$\forall \varepsilon > 0 \quad \exists N \text{ s.t } \|v_n - v\| \leq \varepsilon \quad \forall n \geq N$

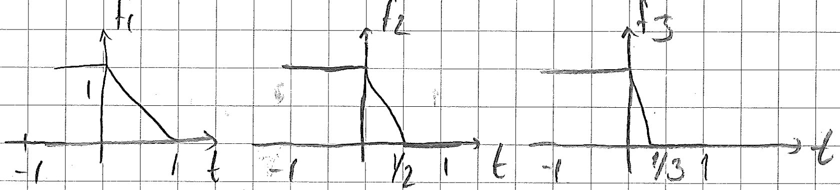
Cauchy Sequence

$\|v_n - v_m\| \to 0 \quad n, m \to \infty$

OR

$\forall \varepsilon > 0 \quad \exists N \text{ s.t } \|v_n - v_m\| \leq \varepsilon \quad \forall n, m \geq N$

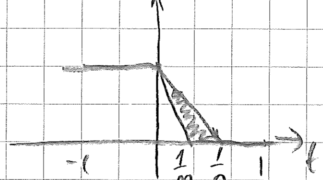
If a Seq. is convergent $\Rightarrow$ It is Cauchy
$\not\Leftarrow$

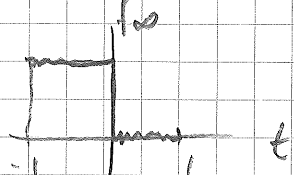


$$V = \left\{ f : [-1, 1] \to \mathbb{R},\ f \text{ is cont.},\ \int_{-1}^{1} |f|\, dt < \infty \right\}$$

$$f_n \in V \qquad \|f\| \stackrel{\Delta}{=} \int_{-1}^{1} |f(t)|\, dt$$

$$\|f_n - f_m\| = \int_{-1}^{1} |f_n(t) - f_m(t)|\, dt \to 0 \to \text{This is Cauchy Sequence}$$





f_n converges to f_∞
f_∞ ~~$\in V$~~

because f_∞ is NOT cont
not inside V.

Spaces where all Cauchy sequences are convergent are called Complete Normed Spaces. (Banach Spaces)

INNER PRODUCT SPACES

An inner product space is a linear space with an additional function called an inner product.

DEFINITION: Let $(V, \mathbb{C})$ be a vector space (ie, the field is complex)

An inner product is a function from $(V \times V)$ into $\mathbb{C}$

$$\langle \cdot , \cdot \rangle : V \times V \to \mathbb{C}$$

Inner product satisfies the following properties;

1) $\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall x, y \in V$ (conjugate symmetry)

2) Linearity w.r.t the first argument:

   a) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad \forall x, y, z \in V$ (additivity)

   b) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \forall x, y \in V \quad \forall \alpha \in \mathbb{C}$ (homogeneity)

3) $\langle x, x \rangle$ is real & nonnegative

$$\langle x, x \rangle \geq 0 \quad \forall x \in V$$

Equality hold iff $x = 0_V$

$$\langle x, x \rangle = 0 \iff x = 0_V$$

(positive definitness)

Linearity wrt second argument:

$$\langle x, y+z \rangle = \overline{\langle y+z, x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle$$

additivity wrt 2^{nd} argument holds

$\sum x_i \bar{y}_i$

$\overline{\sum y \bar{x}}$

$\sum x \bar{y}$

$$\langle x, \alpha y\rangle = \overline{\langle \alpha y, x\rangle} = \overline{\alpha \langle y, x\rangle} = \bar{\alpha}\langle x, y\rangle$$

homogeneity holds with $\bar{\alpha}$.

EXAMPLE: $V = \mathbb{C}^n$ $\quad x = [x_1\ x_2 \dots x_n]^T$

$y = [y_1\ y_2 \dots y_n]^T$

$$\langle x, y\rangle \triangleq \sum_{i=1}^{n} x_i \bar{y}_i$$

Check the properties:

✓ * 1) $\langle x, y\rangle = \sum_{i=1}^{n} x_i \bar{y}_i = \sum_{i=1}^{n} \overline{\overline{x_i \bar{y}_i}} = \sum_{i=1}^{n} \overline{\bar{x}_i y_i} = \overline{\sum_{i=1}^{n} \bar{x}_i y_i} = \overline{\langle y, x\rangle}$

✓ * 2) a) $\langle \underset{x_i}{x} + \underset{y_i}{y}, \underset{z_i}{z}\rangle = \sum_{i=1}^{n} (x_i + y_i)\bar{z}_i = \sum_{i=1}^{n} x_i \bar{z}_i + \sum_{i=1}^{N} y_i \bar{z}_i = \langle x, z\rangle + \langle y, z\rangle$

b) $\langle \alpha x, y\rangle = \sum_{i=1}^{n} (\alpha x_i)\bar{y}_i = \alpha \sum_{i=1}^{n} x_i \bar{y}_i = \alpha \langle x, y\rangle$

✓ * 3) $\langle x, x\rangle = \sum_{i=1}^{n} x_i \bar{x}_i = \sum_{i=1}^{n} |x_i|^2$ is alway real

If $x = 0 \Rightarrow x_i = 0 \Rightarrow |x_i| = 0 \Rightarrow \langle x, x\rangle = 0$

If $\langle x, x\rangle = 0 \Rightarrow |x_i| = 0 \Rightarrow x_i = 0 \Rightarrow x = 0$

EXAMPLE: $V = \mathbb{C}^{n \times n}$

$$\langle A, B \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \overline{b_{ij}} = \text{trace}\,(A \overline{B}^T)$$
$$= \text{trace}\,(A B^H)$$

$A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times n}$

$\cancel{a_{ij}\, b_{ji}}$

EXAMPLE: $V = \left\{ f \mid [0,1] \to \mathbb{C} \text{ such that } \int_0^1 |f(t)|^2 dt < \infty \right\}$

V: square integrable functions.

$$\langle f_1, f_2 \rangle = \int_0^1 f_1(t) \overline{f_2(t)}\, dt$$

$(V, \langle f_1, f_2 \rangle)$ is called L_2 inner product space

THEOREM: (Cauchy - Schwarz Inequality)

Let V be an inner product space. Let $x, y \in V$

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

Proof: Part (I) Case $y = 0_V$

$$\langle x, y \rangle = \langle x, 0_V \rangle = \langle x, \underset{\alpha}{0} . \underset{y}{0_V} \rangle = \overline{0} \langle x . 0_V \rangle = 0$$

$\langle y, y \rangle = 0$

$0 \leq \langle x, x \rangle . 0 = 0$

$0 \leq 0$ Inequality is correct for $y = 0_V$

Part (II) Case $y \neq 0_V$

(2 saat ek ders)

03.11.2015
Tuesday

(Part II) : Case $y \neq 0_V$

$$\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$$

Let $\lambda \in \mathbb{C}$ be arbitrary

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle$$

$$= \langle x, x \rangle - \langle \lambda y, x \rangle - \langle x, \lambda y \rangle + \langle \lambda y, \lambda y \rangle$$

$$= \langle x, x \rangle - \lambda \langle y, x \rangle - \bar{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle$$

$$= \langle x, x \rangle - \lambda \overline{\langle x, y \rangle} - \bar{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle$$

Choose specific $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$

$$= \langle x, x \rangle - \frac{\langle x, y \rangle \overline{\langle x, y \rangle}}{\langle y, y \rangle} - \frac{\overline{\langle x, y \rangle} \langle x, y \rangle}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}$$

$$= \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}$$

$$= \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} \geq 0$$

$$\Rightarrow |\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

FABER-CASTELL

Remark: An inner product induces a norm defined as

$$\|x\|^2 \triangleq \langle x, x \rangle$$

OR

$$\|x\| \triangleq \sqrt{\langle x, x \rangle}$$

$\downarrow$

induced norm

Proof:

1) $\|x\| \geqslant 0$

$\|x\| = 0$ iff $x = 0_V$ by the properties of the inner product

2) $\|\alpha x\| \triangleq \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{|\alpha|^2 \langle x, x \rangle}$

$$= |\alpha| \sqrt{\langle x, x \rangle}$$

$$= |\alpha| \, \|x\|$$

3) $\|x+y\|^2 \triangleq \langle x+y, x+y \rangle$

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

$$= \langle x, x \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} + \langle y, y \rangle$$

$$= \langle x, x \rangle + 2 \operatorname{Re}\{ \langle x, y \rangle \} + \langle y, y \rangle$$

$$\leq \langle x, x \rangle + 2 \underbrace{|\langle x, y \rangle|}_{\leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}} + \langle y, y \rangle$$

$$\leq \langle x, x \rangle + 2 \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} + \langle y, y \rangle$$

$$= \left(\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle} \right)^2$$

$a + bj$

$a \leq \sqrt{a^2+b^2}$

$$\|x+y\|^2 \leq \left(\sqrt{\langle x,x\rangle} + \sqrt{\langle y,y\rangle}\right)^2 = \left(\|x\| + \|y\|\right)^2$$

$$\|x+y\|^2 \leq \left(\|x\| + \|y\|\right)^2 \quad \text{(like triangular equality)}$$

Remark: Every inner product space is also a normed space, converse is not true.

Definition: Inner product space that is complete wrt the norm induced by the inner product is called a complete inner product space.

$$\text{Completeness} \Rightarrow \underbrace{\text{Cauchy}}_{\text{wrt induced norm}} \Rightarrow \underbrace{\text{Convergent}}_{\text{wrt induced norm}}$$

Definition: Two vectors $x, y \in V$ are said to be orthogonal if $\langle x, y\rangle = 0$

Similarly two subsets $S_1, S_2 \subset V$ are orthogonal if $\langle x, y \rangle = 0 \quad \forall x \in S_1 \quad \forall y \in S_2$

EX:

1) $V = \mathbb{R}^n \qquad \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T y = y^T x$

Consider the canonical basic set

$B = \{ e_1, e_2, \ldots, e_n \}$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots\cdots\cdots \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$\langle e_i, e_j \rangle = \left\{ \begin{array}{ll} 0 & , \quad i \neq j \\ 1 & , \quad i = j \end{array} \right\} = \delta_{ij} \rightarrow \text{Kronecker Delta Function}$$

2) $V = \left\{ f \mid f : [-\pi, \pi] \rightarrow \mathbb{C} \text{ should that } \int_{-\pi}^{\pi} |f(t)|^2 dt < \infty \right\}$

Consider the set $\{ e^{jnt} \}_{n \rightarrow -\infty}^{\infty}$

$$\langle f_1, f_2 \rangle \triangleq \int_{-\pi}^{\pi} f_1(t) \overline{f_2(t)} \, dt$$

complex because conjugate in complex plane

$$\langle e^{jnt}, e^{jmt} \rangle = \int_{-\pi}^{\pi} e^{jnt} e^{-jmt} dt$$

$$= \int_{-\pi}^{\pi} \cos((n-m)t) + j \sin((n-m)t) \, dt$$

$$\langle e^{jnt}, e^{jmt} \rangle = \begin{cases} 0 & , n \neq m \\ 2\pi & n = m \end{cases}$$

$\rightarrow$ cos & sin are periodic when $n \neq m$

$$= 2\pi \, \delta_{nm}$$

3) $S_1 = \{\sin nt\}_{n=1}^{\infty}$

$S_2 = \{\cos mt\}_{m=1}^{\infty}$

Are S_1 & S_2 orthogonal?

(We have same V & inner product as the previous example)

$$\langle \sin(nt) \cos(mt) \rangle = \int_{-\pi}^{\pi} \sin nt \cos mt \, dt$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \left(\sin[(n+m)t] + \sin[(n-m)t] \right) dt$$

$$= \begin{cases} 0, & n=m \\ 0, & n \neq m \end{cases} = 0$$

Yes, S_1 & S_2 are orthogonal

$V = \mathbb{R}^n \quad \langle x, y \rangle = x_1 y_1 + x_2 y_2$

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$? \begin{bmatrix} 1 \\ 2 \end{bmatrix} + ? \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \text{Bunu bulamayız}$$

$$\underbrace{\begin{bmatrix} 2 \\ 3 \end{bmatrix}}_{z} = \alpha_1 \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{z_1} + \alpha_2 \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{z_2}$$

$$z = \alpha_1 z_1 + \alpha_2 z_2$$

$$\langle z, z_1 \rangle = \alpha_1 \langle z_1, z_1 \rangle + \alpha_2 \cancel{\langle z_2, z_1 \rangle}$$

0 because z_1 & z_2 are orthogonal

$$\alpha_1 = \frac{\langle z, z_1 \rangle}{\langle z_1, z_1 \rangle}$$

$$\alpha_1 = \frac{5}{2} \qquad \alpha_2 = \frac{-1}{2}$$

$$f(t) = \sum_{n=-\infty}^{\infty} \alpha_n e^{jnt}$$

$$\alpha_n = \frac{\langle f(t), e^{jnt} \rangle}{\langle e^{jnt}, e^{jnt} \rangle} = \underbrace{\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, e^{-jnt}\, dt}_{\text{Fourier Series}}$$

$$\int_{-\pi}^{\pi} e^{jnt} e^{-jnt}\, dt$$

Let V be an inner product space.

Let $S_1 = \{v_1, v_2, \ldots, v_n\}$ be an linearly independent set

Can we obtain another set $S_2 = \{w_1, w_2, \ldots, w_n\}$ such that

i) $w_i \perp w_j$ when $i \neq j$ s.t

$$\langle w_i, w_j \rangle = 0 \quad \text{for} \quad i \neq j$$

ii) $\text{Span } S_1 = \text{Span } S_2$

Gram Schmidt Orthogonalization Algorithm

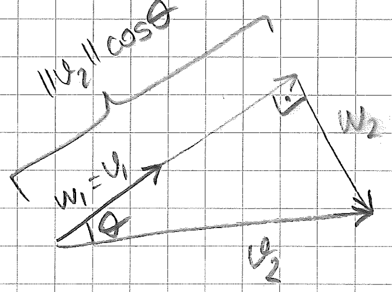
Step 1: $w_1 = v_1$

Step 2: $w_2 = v_2 - \dfrac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1$

$$\langle w_2, w_1 \rangle = \left\langle v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \, , \, w_1 \right\rangle$$

$$= \langle v_2, w_1 \rangle - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} \langle w_1, w_1 \rangle$$

$$= \langle v_2, w_1 \rangle - \langle v_2, w_1 \rangle = 0$$



$$\frac{\langle v_2, w_1 \rangle}{\underbrace{\langle w_1, w_1 \rangle}_{\|w_1\|^2}} w_1 = \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1$$

$\langle x, y \rangle = \|x\| \, \|y\| \cos\theta$

$$= \underbrace{\left\langle v_2, \frac{w_1}{\|w_1\|} \right\rangle}_{\|v_2\| \cdot 1 \cdot \cos\theta} \frac{w_1}{\|w_1\|}$$

$\dfrac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1 \longrightarrow$ projection of V_2 onto V_1

Step 3: $w_3 = v_3 - \underset{w_1}{\text{projection}}(v_3) - \underset{w_2}{\text{projection}}(v_3)$

$$= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2$$

Step n: $w_n = v_n - \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} w_i$

Proof:

Claim 1: $\langle w_i, w_j \rangle = 0 \quad i \neq j \quad w_i \perp w_j$

ie, we are going to prove that $\{w_1, \ldots, w_n\}$ is a pairwise orthogonal set.

Proof of Claim 1: Proof by induction (tümevarım)

Step 1: Prove the statement for some small n

Step 2: Assume that a claim is true for $n=k$

Step 3: Prove that claim is true for $n=k+1$

Step 1: Show that $\{w_1, w_2\}$ is an orthogonal set (we have already done this)

Step 2: Assume that the set $\{w_1, w_2, \ldots, w_k\}$ is an orthogonal set

Step 3: Show that $\{w_1, w_2, \ldots, w_{k+1}\}$ is an orthogonal set.

We assume that

$$\langle w_i , w_j \rangle = 0 \qquad i \neq j \qquad 1 \leq i , j \leq k$$

We need to show $\{w_1, w_2, \ldots, w_k, w_{k+1}\}$ is an orthogonal set.
We only have to show that

$$\langle w_{k+1} , w_j \rangle = 0 \qquad 1 \leq j \leq k$$

$$\left\langle v_{k+1} - \sum_{i=1}^{k} \frac{\langle v_{k+1} , w_i \rangle}{\langle w_i , w_i \rangle} w_i , w_j \right\rangle = 0$$

$$\langle v_{k+1} , w_j \rangle - \sum_{i=1}^{k} \frac{\langle v_{k+1} , w_i \rangle}{\langle w_i , w_i \rangle} \langle w_i , w_j \rangle$$

Assumption is $\{w_1, w_2, \ldots, w_k\}$ is orthogonal so $\langle w_i , w_j \rangle = 0$
for $i \neq j$

$$\langle v_{k+1} , w_j \rangle - \frac{\langle v_{k+1} , w_j \rangle}{\langle w_j , w_j \rangle} \langle w_j , w_j \rangle = 0$$

Claim 1 is proven

Claim 2: $\text{Span}\{w_1, w_2, \ldots, w_n\} = \text{span}\{v_1, v_2, \ldots, v_n\}$
Prove by induction

<u>Step 1:</u> Let us choose $n = 1$

$$\text{span}\{w_1\} = \text{span}\{v_1\}$$

$$w_1 = v_1$$

The statement is trivially correct because $w_1 = v_1$

FABER-CASTELL

Step 2: Assume that

$$\text{span}\{w_1, w_2, \ldots, w_k\} = \text{span}\{v_1, \ldots, v_k\}$$

Step 3: Prove that

$$\underbrace{\text{span}\{w_1, \ldots, w_k, w_{k+1}\}}_{M_1} = \underbrace{\text{span}\{v_1, \ldots, v_k, v_{k+1}\}}_{M_2}$$

$M_1 = M_2$

$$\left.\begin{array}{l} M_1 \subset M_2 \\ M_2 \subset M_1 \end{array}\right\} M_1 = M_2$$

06.11.2015
Friday

$\{v_1, \ldots, v_2\}$ $\{w_1, \ldots, w_2\}$

$v_i \in V$: inner product space $w_i \perp w_j$ $i \neq j$

v_i's are li. $\text{span}\{w_1, \ldots, w_n\} = \text{span}\{v_1, v_2, \ldots, v_n\}$

Proof of Claim 2:

Step 1: Choose $n=1$ $\text{span}\{w_1\} = \text{span}\{v_1\}$

This is trivially correct since $w_1 = v_1$

Step 2: Assume that $\text{span}\{w_1, \ldots, w_k\} = \text{span}\{v_1, \ldots, v_k\}$

Prove that $\underbrace{\text{span}\{w_1, \ldots, w_{k+1}\}}_{M_1} = \text{span}\{\underbrace{v_1, \ldots, v_{k+1}}_{M_2}\}$

$M_1 = M_2$

FABER-CASTELL

Prove: $\left.\begin{array}{l} M_1 \subset M_2 \\ M_2 \subset M_1 \end{array}\right\} \quad M_1 = M_2$

Part 1: Prove $M_1 \subset M_2$

Take an arbitrary element x in M_1 $\Rightarrow$ show that x is also an element of M_2.

Let x be an arbitrary element in M_1.

$$x \in \text{span}\{w_1, \ldots, w_{k+1}\}$$

$$x = \sum_{i=1}^{k+1} \alpha_i w_i = \underbrace{\alpha_1 w_1}_{\in \text{span}\{v_1\}} + \underbrace{\alpha_2 w_2}_{\in \text{span}\{v_1, v_2\}} + \ldots + \underbrace{\alpha_{k+1} w_{k+1}}_{\in \text{span}\{v_1, \ldots, v_{k+1}\}}$$

$$\Rightarrow x \in \text{span}\{v_1, v_2, \ldots, v_{k+1}\} \Rightarrow \boxed{M_1 \subset M_2}$$

Part 2: Prove $M_2 \subset M_1 = \text{span}\{w_1, \ldots, w_{k+1}\}$

Let x be arbitrary in M_2 $\Rightarrow$ $x \in M_2$

$$x \in \text{span}\{v_1, \ldots, v_{k+1}\}$$

$$x = \sum_{i=1}^{k+1} \beta_i v_i = \sum_{i=1}^{k} \beta_i v_i + \beta_{k+1} v_{k+1}$$

$$= \underbrace{\sum_{i=1}^{k} \beta_i v_i}_{\substack{\in \text{span}\{v_1, \ldots, v_k\} \\ = \text{span}\{w_1, \ldots, w_k\}}} + \beta_{k+1} \underbrace{\left(w_{k+1} + \sum_{j=1}^{k} \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} w_j \right)}_{\in \text{span}\{w_1, \ldots, w_{k+1}\}}$$

$$x \in \text{span}\{w_1, \ldots, w_{k+1}\} \Rightarrow x \in M_1 \Rightarrow \boxed{M_2 \subset M_1}$$

FABER-CASTELL

EX: Let $V = \{ f \mid f : [-1,1] \to \mathbb{R}$, f is continuous, $\int_{-1}^{1} |f(t)|^2 dt < \infty \}$

Define the inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(t)\, g(t)\, dt$$

$\{ v_1, v_2, v_3 \}$ $\quad v_1 = 1(t) \quad v_2 = t \quad v_3 = t^2 \longrightarrow$ Legendre Polynomials

Find $\{ w_1, w_2, w_3 \}$ such that $w_i \perp w_j \quad i \neq j$

$\text{span}\{ v_1, v_2, v_3 \} = \text{span}\{ w_1, w_2, w_3 \}$

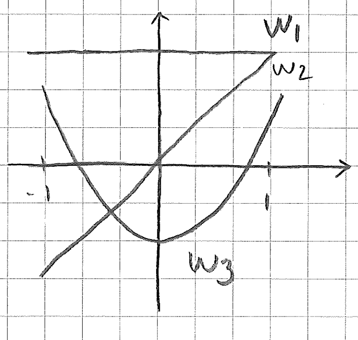
$w_1 = v_1 = 1(t)$

$$w_2 = t - \frac{\int_{-1}^{1} t\, dt}{\int_{-1}^{1} dt} 1(t) = t$$

2

$$w_3 = t^2 - \frac{\int_{-1}^{1} t^2 dt}{\int_{-1}^{1} dt} 1(t) - \frac{\int_{-1}^{1} t^3 dt}{\int_{-1}^{1} t^2 dt} t = t^2 - \frac{\frac{2}{3}}{2} 1(t)$$

$$= t^2 - \frac{1}{3}$$



GRAM-SCHMIT ORTHONORMALIZATION

$$\{v_1, \ldots, v_n\} \rightarrow \{w_1, \ldots, w_n\} \qquad w_i \perp w_j \quad i \neq j$$

$$\text{span}\{v_1, \ldots, v_n\} = \text{span}\{w_1, \ldots, w_n\}$$

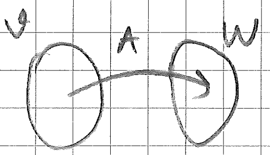
- Apply Gram Schmitt Orthogonalization
- Normalize the resulting vectors

Step 1: $\bar{w}_1 = v_1 \qquad w_1 = \dfrac{\bar{w}_1}{\|\bar{w}_1\|}$

Step 2: $\bar{w}_2 = v_2 - \langle v_2, w_1 \rangle w_1 \quad , \quad w_2 = \dfrac{\bar{w}_2}{\|\bar{w}_2\|}$

Step 3: $\bar{w}_3 = v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2 \quad , \quad w_3 = \dfrac{\bar{w}_3}{\|\bar{w}_3\|}$

Step n: $\bar{w}_n = v_n - \sum_{i=1}^{n-1} \langle v_n, w_i \rangle w_i \qquad w_n = \dfrac{\bar{w}_n}{\|\bar{w}_n\|}$



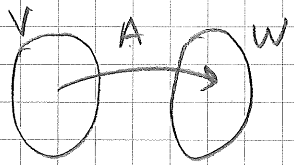
LINEAR TRANSFORMATION

DEFINITION: Let V and W be a vector spaces over the same field F.

A linear transformation A is a mapping from V into W shown as $A: V \to W$ such that

$$A(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 A v_1 + \alpha_2 A v_2 \qquad \forall\, \alpha_1, \alpha_2 \in F$$

$$\forall\, v_1, v_2 \in V$$



$$A\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i A v_i$$

EX

$V = W =$ polynomials degree $\leq n$

$A = \frac{d}{dt}$ V: domain of the L.T

W: Co-domain of the transformation

$$x_1 = \sum_{i=1}^{n} c_i t^i \qquad x_2 = \sum_{i=1}^{n} d_i t^i$$

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \frac{d}{dt}\left[\alpha_1\left(\sum_{i=1}^{n} c_i t^i\right) + \alpha_2\left(\sum_{i=1}^{n} d_i t^i\right)\right]$$

$$= \alpha_1 \frac{d}{dt}\left(\sum_{i=1}^{n} c_i t^i\right) + \alpha_2 \frac{d}{dt}\left(\sum_{i=1}^{n} d_i t^i\right) = \alpha_1 A x_1 + \alpha_2 A x_2$$

$A = \frac{d}{dt}$ is a linear transformation

Ex:

$V = W = \mathbb{R}^2$

Let A be defined as

$$Ax = \begin{bmatrix} \alpha_1 \\ \alpha_1 + \alpha_2 \end{bmatrix} \quad \text{where} \quad x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

$$x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \qquad y = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$ax + by = \begin{bmatrix} a\alpha_1 + b\beta_1 \\ a\alpha_2 + b\beta_2 \end{bmatrix}$$

$$A(ax+by) = \begin{bmatrix} a\alpha_1 + b\beta_1 \\ a\alpha_1 + b\beta_1 + a\alpha_2 + b\beta_2 \end{bmatrix}$$

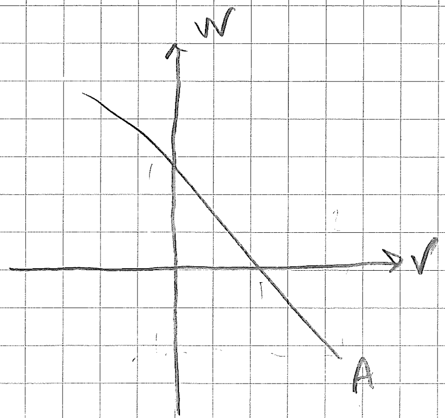
$$= a\begin{bmatrix} \alpha_1 \\ \alpha_1 + \alpha_2 \end{bmatrix} + b\begin{bmatrix} \beta_1 \\ \beta_1 + \beta_2 \end{bmatrix}$$

$$= a\,Ax + b\,Ay$$

A is a linear transformation

Ex:

$V = W = \mathbb{R}$



$y = 1 - x$ $\qquad$ $Ax = 1 - x$

$x_1 = 0 \longrightarrow y_1 = 1$

$x_2 = 2 \longrightarrow y_2 = -1$

$$A(\underbrace{ax_1 + bx_2}_{x_1+x_2}) = a\,Ax_1 + b\,Ax_2$$

$$A2 = A0 + A2$$

$$-1 = 1 + (-1)$$

$$A(ax_1 + bx_2) = a\,Ax_1 + b\,Ax_2$$

$a, b \in \mathbb{R}$

$a = b = 0$

$A0 = 0$

$A0_V = 0_W$

Ex:

$V = \{ f \mid f : [0,1] \to \mathbb{R},\ f \text{ is integrable} \}$

$W = \mathbb{R}$

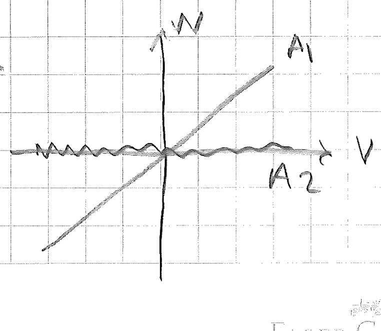
Define A as $A(f) = \int_0^1 f(t)\,dt$

A is a linear transformation thanks to the properties of the integral operation.

DEFINITION: Let $A : V \to W$ be a linear transformation. The null-space of A is defined as

$$N(A) \triangleq \{ v \in V \mid Av = 0_W \}$$

$V = W = \mathbb{R}$



$N(A_1) = 0$

$N(A_2) = \mathbb{R}$

Theorem: $N(A)$ is a subspace of V

Let $x_1, x_2 \in N(A)$

$x_1 + x_2 \in N(A)$

$a x_1 \in N(A)$

$Ax_1 = 0$

$Ax_2 = 0$

$$A(x_1 + x_2) = \underbrace{Ax_1}_{0_W} + \underbrace{Ax_2}_{0_W} = 0_W$$

$\Rightarrow x_1 + x_2 \in N(A)$

$A(ax_1) = a \underbrace{Ax_1}_{0_W} = 0_W \qquad ax_1 \in N(A)$

$N(A)$ is a subspace of V.

DEFINITION: A linear transformation A is called one to one if

$Av_1 = Av_2$ implies $v_1 = v_2$

Proposotion: A linear transformation $A: V \to W$ is one-to-one

iff $N(A) = \{0_V\}$

A is one-to-one $\Leftrightarrow N(A) = \{0_V\}$

$p \Leftrightarrow q$

Part 1: A is one to one $\Rightarrow N(A) = \{0_V\}$

$p \Rightarrow q$

FABER-CASTELL

Proof by contradiction: $\bar{q} \Rightarrow \bar{p}$

Suppose the $N(A) \neq \{0_V\}$

$\exists\, x \in N(A)$ such that $x \neq 0_V$

$$Ax = 0_W$$

Suppose $v_1 \in V$

Consider v_1 & $v_2 = v_1 + x$

$$v_1 \neq v_2$$

Av_1

$Av_2 = A(v_1 + x) = Av_1 + \underbrace{Ax}_{0_W} = Av_1$

$$Av_1 = Av_2$$

$\Rightarrow A$ is NOT one-to-one $\} \bar{p}$

Part 2: $q \Rightarrow p$

$N(A) = \{0_V\} \Rightarrow A$ is one-to-one

$\Rightarrow p$

Proof by Contradiction: $\bar{p} \Rightarrow \bar{q}$

Suppose that A is NOT one-to-one

$\exists\, v_1, v_2$ such that $v_1 \neq v_2$ and $Av_1 = Av_2$

$$A(v_1 - v_2) = 0_W$$

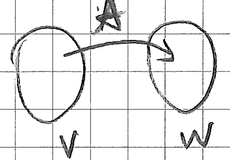
$v_1 - v_2 \in N(A)$ and $v_1 - v_2 \neq 0_V$

$\Rightarrow N(A) = \{0_V\} \} \bar{q}$

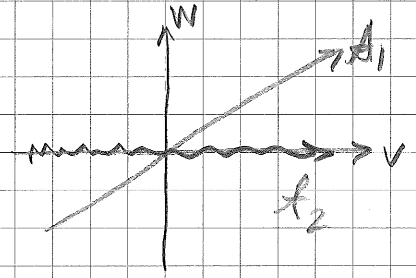
10.11.2015.
Tuesday

Linear Transformation

V, W: Fields are the same



$A(a_1 v_1 + a_2 v_2) = a_1 A(v_1) + a_2 A(v_2)$

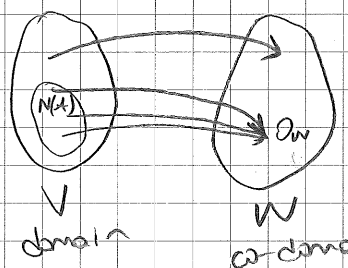


$V = W = \mathbb{R}$

→ eğer zerodan geçmezse LT diyemeyiz.

$N(A_1) = 0_V$

$N(A_2) = V = \mathbb{R}$



null space → $N(A) \triangleq \{ v \in V \mid Av = 0_W \}$

One to one

$Av_1 = Av_2 \Rightarrow v_1 = v_2$

one to one $\Longleftrightarrow N(A) = \{0_V\}$

EXAMPLE:

$V = W = \{\text{polynomials with degree} \leq 2\}$

$A = \frac{d}{dt}$ Is A one-to-one?

$v = at^2 + bt + 5$ $\to$ $v_1 = t+1$, $v_2 = t-1$ $\Big\}$ $v_1 \neq v_2$ while $Av_1 = Av_2$

A is NOT one-to-one!

$N\{A\} = \{\text{constant polynomials}\} \neq \{0_V\}$

EXAMPLE:

$V = \{f \mid f \text{ is integrable } f:[0,1] \to \mathbb{R}\}$

$W = \mathbb{R}$

$Af = \int_0^1 f(t)\,dt$ Is A one-to-one?

$v = 2t-1 \to \int_0^1 (2t-1)\,dt = t^2 - t \Big|_0^1 = 0$

$f = 2t-1 \in N(A)$

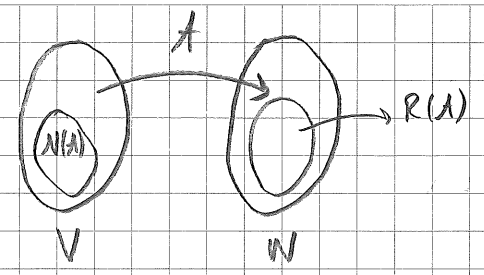
$N(A) \neq \{0_V\} \Rightarrow A$ is NOT one-to-one

EXAMPLE: $V = W = \mathbb{R}^2$

$Ax = \begin{bmatrix} \alpha_1 \\ \alpha_1 + \alpha_2 \end{bmatrix}$ $x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ Is A one-to-one?

$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ $Av = 0 = \begin{bmatrix} v_1 \\ v_1 + v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} v_1 = 0 \\ v_2 = 0 \end{matrix} \Big\} \Rightarrow N(A) = \{0_V\}$

A is one-to-one.



DEFINITION: Let $A: V \to W$ be a linear transformation. The range space of A is defined as

$$R(A) \triangleq \{ w \in W \mid \exists v \in V \text{ st } Av = w \}$$

Theorem: $R(A)$ is a subspace of W.

Proof: Let $w_1, w_2 \in R(A)$ $\exists$ $v_1, v_2 \in V$ st. $Av_1 = w_1$
$Av_2 = w_2$

$\to$ $w_1 + w_2 \overset{?}{\in} R(A)$

$v_1 + v_2 \in V$

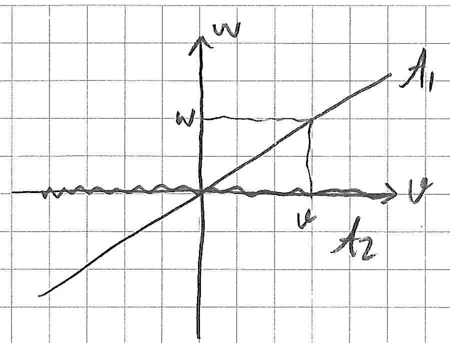
$A(v_1 + v_2) = Av_1 + Av_2 = w_1 + w_2$

$\Rightarrow$ $w_1 + w_2 \in R(A)$

$\to$ $a w_1 \in R(A)$ $\quad a v_1 \in V$ $\quad A a v_1 = a A v_1 = a w_1$

$\Rightarrow$ $a w_1 \in R(A)$

$\therefore$ $R(A)$ is a subspace of W.



$V = W = \mathbb{R}$

$R(A_1) = W = \mathbb{R}$

$R(A_2) = \{0_W\}$

DEFINITION: A linear transformation $A : V \to W$ is called onto if $R(A) = W$.

Otherwise if $R(A) \subset W$, A is called into.

In previous example, A_1 is onto & A_1 is into.

EXAMPLE: $V = W = \{$polynomials with degree $\leq 2\}$

$A = \frac{d}{dt}$ Is A is onto?

$w = t^2$ $\nexists\, v \in V$ st $Av = t^2$

$\Rightarrow A$ is NOT onto.

EXAMPLE: $V = \{ f : [0,1] \to \mathbb{R}$ (f is integrable$\}$

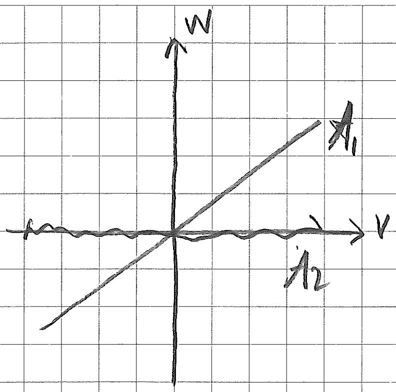
$W = \mathbb{R}$

$Af = \int_0^1 f(t)\, dt$ Is A is onto?

$a \in \mathbb{R}$

$f(t) = a \quad \forall t \in [0,1] \to \int_0^1 f(t)\, dt = a$

A is onto.



$V = W = \mathbb{R}$

$\mathcal{N}(A_2) = \{0\}$ dim = 0

$\mathcal{N}(A_2) = \mathbb{R}$ dim = 1

$R(A_1) = \mathbb{R}$ dim = 1

$R(A_2) = \{0\}$ dim = 1

$$\boxed{\dim \mathcal{N}(A) + \dim R(A) = \dim V}$$

<u>Theorem:</u> If V is finite dimensional & A is linear transformation from V to W. Then $R(A)$ is a finite dimensional subspace of W and

$$\dim \mathcal{N}(A) + \dim R(A) = \dim V.$$

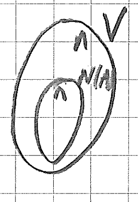
$$v = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2 \xrightarrow{A} Av = a_1 t + a_2 t^2$$

<u>Proof:</u> ~~Part 1)~~ Let V be n-dimensional

Since $\mathcal{N}(A) \subset V$

Suppose $\dim \mathcal{N}(A) = m$

Part 1) Suppose $m = n$



$\mathcal{N}(A) \subset V$

$\dim \mathcal{N}(A) = \dim V$

$\mathcal{N}(A) = V$

What about the range space?

$$\underset{v \in V}{A v} = 0_W \Rightarrow R\{A\} = 0_W$$

$\dim N(A) = n$

$\dim R(A) = 0$

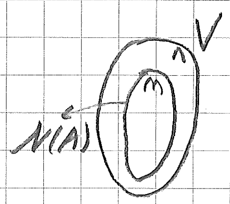
$$\dim N(A) + \dim R(A) = n = \dim V$$

Part 2) Suppose $m < n$

Suppose $\{v_1, \ldots, v_m\}$ be a basis for $N(A)$

$$\text{span}\{v_1, \ldots, v_m\} = N(A)$$

$$N(A) \subsetneq V$$

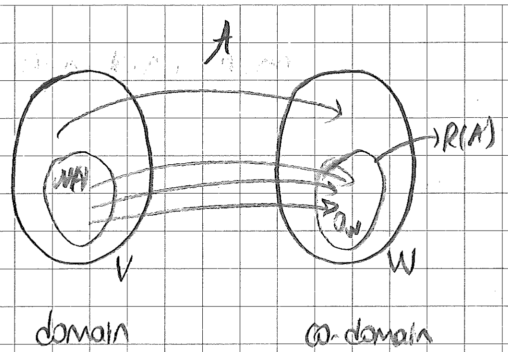


We can add vectors in to basis $\{v_1, \ldots, v_m\}$ such that the resulting set is a basis for V.

$$\{v_1, v_2, \ldots, v_m, w_1, w_2, \ldots w_{n-m}\}$$

Let v be an arbitrary element in V

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m + \beta_1 w_1 + \beta_2 w_2 + \ldots + \beta_{n-m} w_{n-m}$$



13.11.2015
Friday

$N(A) \triangleq \{x \in V \mid Ax = 0_W\}$

$R(A) \triangleq \{y \in W \mid \exists x \in V \text{ st } Ax = y\}$

$$\dim N(A) + \dim R(A) = \dim V$$

Proof: $\dim V = n \quad \dim N(A) = m \qquad m \leq n$

Part 1) Case $m = n$

$$\left.\begin{array}{l} \dim N(A) = n \\ N(A) \subset V \end{array}\right\} N(A) = V$$

$R(A) = \{0_W\}$

$A\underset{\in V}{v} = 0_W \qquad \dim R(A) = 0$

Equality $\dim N(A) + \dim R(A) = n$ is satisfied.

Part 2) $m < n$

Let $\{v_1, v_2, \ldots, v_m\}_{\subset V}$ be a basis for $N(A)$

We can find $n-m$ vectors in V s.t the new set is a basis for V.

$\{v_1, v_2, \ldots, v_m, w_1, w_2, \ldots, w_{n-m}\}$: A basis for V.

We can write an arbitrary vector $v \in V$ as a linear combination of these vectors.

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m + \beta_1 w_1 + \cdots + \beta_{n-m} w_{n-m}$$

Apply A to both sides of this equation.

$$\underbrace{Av}_{\substack{\text{An arbitrary element} \\ \text{of } R(A)}} = \beta_1 \underbrace{Aw_1}_{\in W} + \beta_2 \underbrace{Aw_2}_{\in W} + \cdots + \beta_{n-m} \underbrace{Aw_{n-m}}_{\in W}$$

We can write an arbitrary element of $R(A)$ as a linear combination of $n-m$ vectors.

$$sp\{Aw_1, Aw_2, \ldots, Aw_{n-m}\} \Rightarrow \dim R(A) \leq n-m$$

$R(A)$ is finite dimensional!

Check if these vectors are l.i or not.

$$c_1 Aw_1 + c_2 Aw_2 + \cdots + c_{n-m} Aw_{n-m} = 0_W$$

$$A(c_1 w_1 + c_2 w_2 + \cdots + c_{n-m} w_{n-m}) = 0_W$$

$$c_1 w_1 + c_2 w_2 + \cdots + c_{n-m} w_{n-m} \in N(A)$$

$$c_1 w_1 + c_2 w_2 + \cdots + c_{n-m} w_{n-m} = d_1 v_1 + d_2 v_2 + \cdots + d_m v_m$$

$$d_1 v_1 + d_2 v_2 + \cdots + d_m v_m - c_1 w_1 - c_2 w_2 - \cdots - c_{n-m} w_{n-m} = 0_V$$

$$d_1 = d_2 = \cdots = d_m = c_1 = c_2 = \cdots = c_{n-m} = 0$$

$$\{\underbrace{Aw_1, Aw_2, \ldots, Aw_{n-m}}_{\text{A basis for } R(A)}\} \text{ is a l.i. set}$$

FABER-CASTELL

$\dim R(A) = n-m$

$$\underbrace{\dim N(A)}_{=m} + \underbrace{\dim R(A)}_{=n-m} = n$$

Rank Nullity Theorem

Suppose V is an n-dimensional vector space $v \in V$

Suppose $\underbrace{\{v_1, \ldots, v_n\}}_{B}$ is a basis for V

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \Rightarrow [v]_B = \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}}_{R^n}$$

$$V \xrightarrow{A} W$$

V: $\dim = n$, B — W: $\dim = m$, C

$$\underbrace{[v]_B \in R^n}_{R^n} \qquad \underbrace{[w]_C \in R^m}_{R^m}$$

$$[w]_C = A[v]_B \quad \text{if} \quad w = Av$$

MATRIX REPRESENTATIONS OF LINEAR TRANSFORMATIONS

DEFINITION: Let V ($\dim V = n$) and W ($\dim W = m$) be two linear spaces over the same field F.

Let $A: V \to W$ be a linear transformation. Given the basis $B = \{v_1, \ldots, v_n\}$ and $C = \{w_1, \ldots, w_m\}$ for V & W, the transformation A can be represented by an $m \times n$ matrix $A = [a_{ij}]$ ($a_{ij} \in F$) such that for $v \in V$ & $w \in W$ satisfying $w = Av$

Then we have $[w]_C = \underbrace{A}_{m \times n} [v]_B$

Let us have $w = Av$ & suppose that

$$[w]_C = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \qquad [v]_B = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$w = \sum_{i=1}^{m} \beta_i w_i = Av = A \sum_{j=1}^{n} \alpha_j \underbrace{v_j}_{\in V} = \sum_{j=1}^{n} \alpha_j \underbrace{Av_j}_{\in W}$$

$$= \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{m} a_{ij} w_i$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \alpha_j \right) w_i$$

$$= \sum_{i=1}^{m} \beta_i w_i$$

$$\boxed{\beta_i = \sum_{j=1}^{n} a_{ij} \alpha_j} \qquad i = 1, \ldots, m$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Procedure to find A:

1) Take each vector v_j in B
2) Take the transform of v_j : Av_j
3) Express Av_j in terms of w_i in C
4) Set the j^{th} column of A to the resulting coefficients

$$[w]_C = A[v]_B$$

$$B = \{v_1, \ldots, v_n\} \qquad C = \{w_1, \ldots, w_m\}$$

$$[Av_1]_C = A[v_1]_B$$

$$[Av_1]_C = \begin{bmatrix} & & \\ & A & \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \text{First column of } A \text{ is } [Av_1]_C$$

Take v_2

$$[Av_2]_C = A[v_2]_B = \begin{bmatrix} & & \\ & & \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\Downarrow$

2^{nd} column of A is $[Av_2]_C$

EXAMPLE: $V = \{\text{polynomials with degree} \le 3\}$; V
$W = \{ \quad " \quad " \quad " \le 2 \}$; W

$A : V \to W \qquad A_f = \frac{df}{dt}$

$$B = \{ \overset{v_1}{1}, \overset{v_2}{1+t}, \overset{v_3}{1+t+t^2}, \overset{v_4}{1+t+t^2+t^3} \}$$

$$C = \{ \overset{w_1}{1}, \overset{w_2}{1+t}, \overset{w_3}{1+t+t^2} \}$$

$$[w]_C = A[v]_B$$

$$A = \begin{bmatrix} 0 & 1 & -1 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix}_{3\times 4}$$

$Av_1 = 0 = 0\times 1 + 0\times(1+t) + 0\times(1+t+t^2)$

$Av_2 = 1 = 1\times 1 + 0\times(1+t) + 0\times(1+t+t^2)$

$Av_3 = 2t+1 = -1\times 1 + 2\times(1+t) + 0\times(1+t+t^2)$

$Av_4 = 3t^2+2t+1 = -1\cdot 1 - 1(1+t) + 3\times(1+t+t^2)$

$v = t^3 + 2t^2 + 3t + 4$

$$[v]_B = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \to [w]_C = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}$$

derivative

$w = -1(1) + 1(1+t) + 3(1+t+t^2)$
$= 3t^2 + 4t + 3$

Ex

Let $V = \mathbb{R}^{2\times 2}$

$A : V \to V$

$$Ax = SX + XS^T \quad \text{where} \quad S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \quad v$$

$$C = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\} \quad w$$

$$[w]_C = A[v]_B$$

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

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

$$= 1w_1 + 0w_2 + (-1)w_3 + 0w_4$$

$$A\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = -w_4 + w_3 + w_1$$

$$A\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = -w_4 + w_3 + w_1$$

$$A\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = w_3 - w_1$$

$$\begin{array}{ccc} V & \xrightarrow{A} & W \\ \downarrow & & \downarrow \\ \mathbb{R}^n & & \mathbb{R}^m \\ B & & C \end{array}$$

$$[v]_B \xrightarrow{A} [w]_C$$

$$\begin{array}{ccc} \bar{B} & & \bar{C} \\ [v]_{\bar{B}} & \xrightarrow{\bar{A}} & [w]_{\bar{C}} \end{array}$$

EFFECT OF BASIS CHANGE ON THE MATRIX REPRESENTATION

1) Basis change in the Domain

Suppose that we change the basis in the domain & select $\bar{B}$ instead of B.

$$[w]_C = A\,[v]_B$$

$$[w]_C = \underset{?}{\overset{\|}{\bar{A}}}\,[v]_{\bar{B}}$$

$[v]_B = P\,[v]_{\bar{B}}$ we already know this

$$[w]_C = \underbrace{AP}_{\bar{A}}\,[v_{\bar{B}}]$$

$$\bar{A} = AP$$

2) Basis change in Co-domain (Range)

We use $\bar{C}$ instead of C

$$[w]_C = A\,[v]_B$$

$$[w]_{\bar{C}} = \bar{A}\,[v]_B$$

$$\boxed{[w]_C = Q\,[w]_{\bar{C}}}$$

FABER-CASTELL

$$Q[w]_{\bar{C}} = A[v]_B$$

$$[w]_{\bar{C}} = \underbrace{Q^{-1}A}_{\bar{A}}[v]_B$$

$$\bar{A} = Q^{-1}A$$

3) Basis Change in Both Domain & Co-domain

$$\left.\begin{aligned} [w]_C &= A[v]_B \\ [w]_{\bar{C}} &= \bar{A}[v]_{\bar{B}} \end{aligned}\right\} \quad \bar{A} = Q^{-1}AP$$

SPECIAL CASE: If $W = V$, A is a linear transformation from V to itself, then the basis B can be taken to be the same as C. Then a basis change in both domain & co-domain would give

$$\bar{A} = P^{-1}AP \quad \text{since } Q = P$$

<u>EXAMPLE:</u> $V = \{\text{polynomials with degree} \leq 3\}$

$W = \{\text{polynomials with degree} \leq 2\}$

$B = \{1,\ 1+t,\ 1+t+t^2,\ 1+t+t^2+t^3\}$

$C = \{1,\ 1+t,\ 1+t+t^2\}$

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

Change the basis in domain as $\bar{B} = \{1, t, t^2, t^3\}$

$\bar{A} = AP \qquad [v]_B = P[v]_{\bar{B}}$

$$[v]_B = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} [v]_{\bar{B}}$$

$$\Rightarrow \bar{A} = AP = \begin{bmatrix} 0 & 1 & -1 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -2 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Change the basis in the co-domain to canonical basis while keeping the basis of the domain as the canonical basis.

$\bar{C} = \{1, t, t^2\}$

$[w]_C = \bar{A}[v]_{\bar{B}}$

$[w]_{\bar{C}} = \bar{\bar{A}}[v]_{\bar{B}}$

$\bar{\bar{A}} = Q^{-1}\bar{A} \qquad$ which $\qquad [w]_C = Q[w]_{\bar{C}}$

$Q^{-1} = R \qquad \bar{\bar{A}} = R\bar{A}$

$Q^{-1}[w]_C = [w]_{\bar{C}}$

$[w]_{\bar{C}} = R[w]_C$

$\bar{C} = \{1, t, t^2\} \qquad C = \{1, 1+t, 1+t+t^2\}$

FABER-CASTELL

$$R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \quad \bar{\bar{A}} = R\bar{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -2 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

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

$$[w]_{\bar{C}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} [v]_{\bar{B}}$$

17.11.2015
Tuesday

$$\begin{tikzcd} \hat{V} \arrow[r, "A"] \arrow[d, phantom, "\cong"] & \hat{W} \arrow[d, phantom, "\cong"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \\ {[v]_B} \arrow[d, leftrightarrow, "P"'] & {[w]_C} \arrow[d, leftrightarrow, "Q"] \\ {[v]_{\bar{B}}} \arrow[r, bend left, "\bar{A}"] & {[w]_{\bar{C}}} \end{tikzcd} \quad \Bigg\} \quad \boxed{\bar{A} = Q^{-1}AP}$$

$$[v]_B = P\,[v]_{\bar{B}}$$

$$[w]_C = Q\,[w]_{\bar{C}}$$

REMARK: In general, a basis for $N(A)$ & $R(A)$ can be found by finding the sets for the $N(A)$ & $R(A)$

EXAMPLE: $V = \mathbb{R}^{2\times 2}$ $\quad W = V$

$$Ax = Sx + xS^T \qquad S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

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

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$$C: \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

Task 1: Find a basis for $R(A)$

$$R(A) \triangleq \{ w \in W \mid \exists v \in V \;\; Av = w \}$$

$$R(A) = \{ y \in \mathbb{R}^4 \mid \exists x \in \mathbb{R}^4 \;\; Ax = y \}$$

- Find a basis for $R(A)$
- Convert the vectors in the basis into matrices using the basis C.

$$y = \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \end{array} \\ \begin{bmatrix} | & | & | & | \\ | & | & | & | \end{bmatrix} \end{array} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad R(A) = \{ \text{column space of } A \}$$

$$y = x_1 c_1 + x_2 c_2 + x_3 c_3 + x_4 c_4$$

Procedure: Make elementary column operations.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis for } R(A) = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ -1 \end{bmatrix} \right\}$$

$$\begin{bmatrix}1\\0\\-1\\0\end{bmatrix} = [w_1]_C \Rightarrow \quad w_1 = \begin{bmatrix}1&0\\0&0\end{bmatrix} - \begin{bmatrix}1&1\\1&0\end{bmatrix} = \begin{bmatrix}0&-1\\-1&0\end{bmatrix}$$

$$\begin{bmatrix}0\\0\\2\\-1\end{bmatrix} = [w_2]_C \Rightarrow \quad w_2 = \begin{bmatrix}2&2\\2&0\end{bmatrix} - \begin{bmatrix}1&1\\1&1\end{bmatrix} = \begin{bmatrix}1&1\\1&-1\end{bmatrix}$$

$$\text{Basis for } R(A) = \left\{ \begin{bmatrix}0&-1\\-1&0\end{bmatrix}, \begin{bmatrix}1&1\\1&-1\end{bmatrix} \right\}$$

$$\dim R(A) = 2 \Rightarrow \dim N(A) = \dim V - \dim R(A) = 2$$

Task 2: Find a basis for $N(A)$

Solution 1: $N(A) = \{ v \in V \;/\; Av = 0_W \}$

$$AX = 0_{2\times 2}$$

$$X = \begin{bmatrix}a&b\\c&d\end{bmatrix}$$

$$AX = \begin{bmatrix}0&1\\-1&0\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix} + \begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}0&-1\\1&0\end{bmatrix} = \begin{bmatrix}c&d\\-a&-b\end{bmatrix} + \begin{bmatrix}b&-a\\d&-c\end{bmatrix} = \begin{bmatrix}c+b&d-a\\-a+d&-b-c\end{bmatrix} = \begin{bmatrix}0&0\\0&0\end{bmatrix}$$

$$\left.\begin{array}{l} b+c=0 \\ d-a=0 \\ \cancel{d-a=0} \\ \cancel{-b-c=0} \end{array}\right\} \quad \begin{array}{l} c=-b \\ d=a \end{array}$$

$$\begin{bmatrix}a&b\\c&d\end{bmatrix} = a\begin{bmatrix}1&0\\0&1\end{bmatrix} + b\begin{bmatrix}0&1\\-1&0\end{bmatrix}$$

$$\text{Basis for } N(A) = \left\{ \begin{bmatrix}1&0\\0&1\end{bmatrix}, \begin{bmatrix}0&1\\-1&0\end{bmatrix} \right\}$$

Solution 2: Use the matrix representation

- Find a basis for $N(A)$
- Convert the vectors into matrices using the basis B.

$$N(A) = \{ x \in \mathbb{R}^4 \mid Ax = 0 \}$$

$$\begin{bmatrix} 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \\ 0 & -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Elementary row operation to simplify the matrix.

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

$x_1 - x_4 = 0 \qquad x_4 = x_1$

$x_2 + x_3 = 0 \qquad x_3 = -x_2$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

Basis for $N(A)$ = $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}$

Basis for $N(A)$ = $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$

DIRECT SUMS

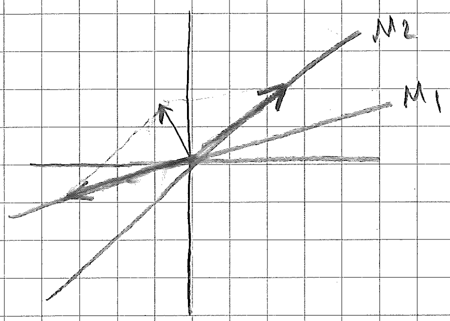
Remember the definition for the sum of subspaces

DEFINITION: Let V be a vector space and let $M_1, M_2, \ldots, M_k$ be subspaces of V. The sum of these subspaces is defined as

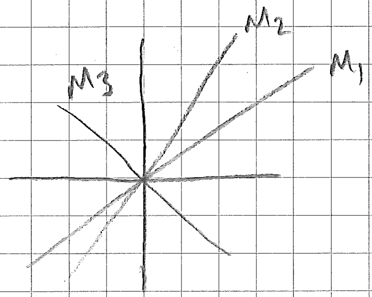
$$M = \{ m \in V \mid m = m_1 + m_2 + \cdots + m_k \text{ where } m_i \in M_i \ \forall i \}$$

$$M = M_1 + M_2 + \cdots + M_k$$

$V = \mathbb{R}^2$



$M_1 \oplus M_2 = \mathbb{R}^2$ $\longrightarrow$ LI



$\cancel{M_1 \oplus M_2 \oplus M_3} = \mathbb{R}^2$ → there is no uniqueness for this case

Not LI

$\oplus$ → to use this symbol, M_i must be linearly ind & span all V.

So $M_1 \oplus M_2$ but not $\cancel{M_1 \oplus M_2 \oplus M_3}$

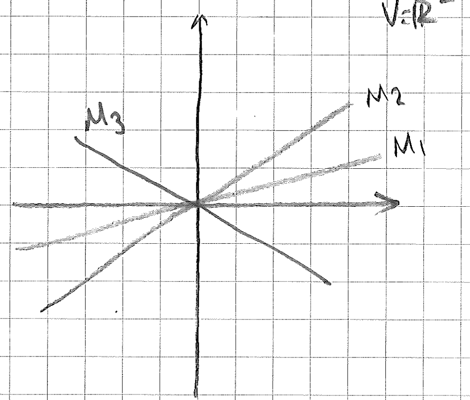
20. 11. 2015
Friday

DIRECT SUM

V, $\underbrace{M_1, M_2, \ldots, M_k}_{\text{subspaces}}$

$M = M_1 + M_2 + \ldots + M_k$

$M = \{ m \in V \mid m = m_1 + m_2 + \ldots + m_k,\ m_i \in M_i\ \forall i \}$



$M_1 + M_2 = \mathbb{R}^2 \rightarrow$ there is uniqeness

$M_1 + M_2 + M_3 = \mathbb{R}^2 \rightarrow$ but for this case, it is NOT unique

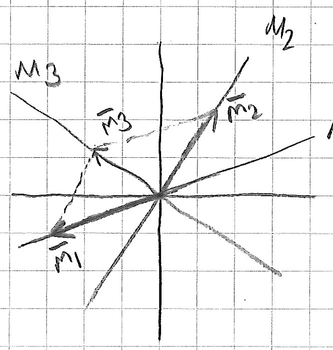
DEFINITION: Let $M_1, M_2, \ldots, M_k$ be subspaces of V. The subspaces are said to be linearly independent if

$$m_1 + m_2 + \ldots + m_k = 0_V \quad \text{where} \quad m_i \in M_i \quad i = 1, 2, \ldots, k$$

implies

$$m_1 = m_2 = \ldots = m_k = 0_V$$

$$a_1 x_1 + a_2 x_2 + \ldots + a_k x_k = 0_V \Rightarrow a_1 = a_2 = \ldots = a_k = 0$$



$$0_V = \underbrace{\overset{0}{\overset{\|}{m_1}}}_{\in M_1} + \underbrace{\overset{0}{\overset{\|}{m_2}}}_{\in M_2}$$

$$0_V = m_1 + m_2 + m_3$$

$\overline{m}_3 = \overline{m}_1 + \overline{m}_2 \Rightarrow \overline{m}_1 + \overline{m}_2 - \overline{m}_3 = 0 \rightarrow$ not linearly ind.

REMARK: Let $V = M_1 + M_2 + \dots + M_k$ and let $M_1, M_2, \dots, M_k$ are linearly independent. Let $x \in V$

$$x = m_1 + m_2 + \dots + m_k \qquad m_i \in M_i \quad \forall i$$

$m_1, m_2, \dots, m_k$ are unique

Proof: Assume that $m_1, \dots m_k$ are not unique.

$$\Rightarrow \quad x = m_1 + \dots + m_k \qquad m_i \in M_i \quad \forall i$$

$$x = \overline{m}_1 + \dots + \overline{m}_k \qquad \overline{m}_i \in M_i \quad \forall i \qquad \exists i \text{ where } m_i \neq \overline{m}_i$$

$$0_V = \underbrace{\overbrace{(m_1 - \overline{m}_1)}^{0_V}}_{\in M_1} + \underbrace{\overbrace{(m_2 - \overline{m}_2)}^{0_V}}_{\in M_2} + \dots + \underbrace{\overbrace{(m_k - \overline{m}_k)}^{0_V}}_{\in M_k}$$

Due to linearly independent
$$\left\{\begin{array}{l} m_1 - \overline{m}_1 = 0_V \\ m_2 - \overline{m}_2 = 0_V \\ \vdots \\ m_k - \overline{m}_k = 0_V \end{array}\right\} \quad \begin{array}{l} m_1 = \overline{m}_1 \\ m_2 = \overline{m}_2 \\ \vdots \\ m_k = \overline{m}_k \end{array} \Bigg\}$$

Contradiction

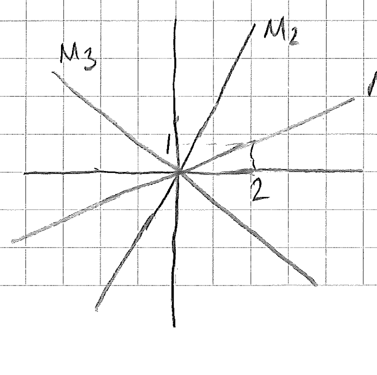
Decomposition $x = m_1 + m_2 + \dots + m_k$ is unique.

DEFINITION: Let $M_1, M_2, \dots, M_k$ be subspaces of V. Let

i) $M = M_1 + M_2 + \dots + M_k$

ii) $M_1, \dots, M_k$ are linearly independent

Then M is said to be direct sum of $M_1, \dots, M_k$. This is shown as $M = M_1 \oplus M_2 \oplus \dots \oplus M_k$



$$\underbrace{\mathbb{R}^2}_{\dim 2} = \underbrace{M_1}_{\dim 1} \oplus \underbrace{M_2}_{\dim 1}$$

$$\cancel{\underbrace{\mathbb{R}^2}_{\dim 2} = \underbrace{M_1}_{\dim 1} \oplus \underbrace{M_2}_{\dim 1} \oplus \underbrace{M_3}_{\dim 1}}$$

$\begin{bmatrix}2\\1\end{bmatrix} \in M_1$ $\quad \left\{ \cancel{\begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}} \right\}$ $\quad B_1 = \left\{ \begin{bmatrix}2\\1\end{bmatrix} \right\} \rightarrow \dim M_1 = 1$

$\dim M_2 = 1$

When we have a direct sum, summation of dimension of M_i is equal to dimension of direct sum.

Ex:

$V = \mathbb{R}^4$ $\qquad x = \begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} \in \mathbb{R}^4$

$M_1 \triangleq \{ x \in \mathbb{R}^4 \;/\; x_3 = x_4 = 0 \}$ $\quad \begin{bmatrix}x_1\\x_2\\0\\0\end{bmatrix}$ $\quad \dim M_1 = 2$

$M_2 \triangleq \{ x \in \mathbb{R}^4 \;/\; x_1 = x_2 = 0 \}$ $\quad \begin{bmatrix}0\\0\\x_3\\x_4\end{bmatrix}$ $\quad \dim M_2 = 2$

$M_3 \triangleq \{ x \in \mathbb{R}^4 \;/\; x_1 = 0 \}$ $\quad \begin{bmatrix}0\\x_2\\x_3\\x_4\end{bmatrix}$ $\quad \dim M_3 = 3$

$M_2 \subset M_3$

$\underset{2}{M_1} + \underset{2}{M_2} = \underset{4}{\mathbb{R}^4} \longrightarrow \begin{bmatrix}x_1\\x_2\\0\\0\end{bmatrix} + \begin{bmatrix}0\\0\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$ $\quad x_1 = x_2 = x_3 = x_4 = 0 \rightarrow$ Direct Sum

$M_1 \oplus M_2 = \mathbb{R}^4$ ✓

$\underset{2}{M_1} + \underset{3}{M_3} \neq \underset{4}{\mathbb{R}^4}$

$\underset{2}{M_2} + \underset{3}{M_3} \neq \underset{3}{M_3}$

$\begin{bmatrix}x_1\\x_2\\0\\0\end{bmatrix} + \begin{bmatrix}0\\\bar{x}_2\\\bar{x}_3\\\bar{x}_4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$

$\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\-1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$

M_1 & M_3 are not l.i.

$\cancel{M_1 \oplus M_3 = \mathbb{R}_4}$

$\begin{bmatrix}0\\0\\x_3\\x_4\end{bmatrix} + \begin{bmatrix}0\\\bar{x}_2\\\bar{x}_3\\\bar{x}_4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$

$\begin{bmatrix}0\\0\\1\\1\end{bmatrix} + \begin{bmatrix}0\\0\\-1\\-1\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$

M_2 & M_3 are not l.i.

$\cancel{\underset{2}{M_2} \oplus \underset{3}{M_3} = M_3}$

DEFINITION: Let V be a inner product space. Two subspaces are said to be orthogonal if

$$\langle m_1, m_2 \rangle = 0 \qquad \begin{array}{l} \forall m_1 \in M_1 \\ \forall m_2 \in M_2 \end{array}$$

We denote this as $M_1 \perp M_2$

DEFINITION: Let M is equal to $M_1 + M_2 + \dots + M_k$ and $M_i \perp M_j$ for $i \neq j$, then M is said to be an orthogonal direct sum of $M_1, \dots, M_k$. We show this as

$$M = M_1 \overset{\perp}{\oplus} M_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} M_k$$

DEFINITION: If $M = V$

$V = M_1 \oplus M_2 \oplus \dots \oplus M_k$ is called to be a direct sum decomposition of V.

$V = M_1 \overset{\perp}{\oplus} M_2 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} M_k$ is called an orthogonal direct sum decomposition of V.

$M \subset V$

$M^\perp \to$ perp

DEFINITION: Let M be a subspace of an inner product space V. The orthogonal complement $M^\perp$ of the subspace M is defined as

$$M^\perp \triangleq \{ x \in V \mid \langle x, m \rangle = 0 \ \forall m \in M \}$$

THEOREM: $M^\perp$ is a subspace of V.

Proof: Take two elements $x, y \in M^\perp$

$$\langle y, m \rangle = 0 \quad \forall\, m \in M$$
$$\langle x, m \rangle = 0 \quad \forall\, m \in M$$

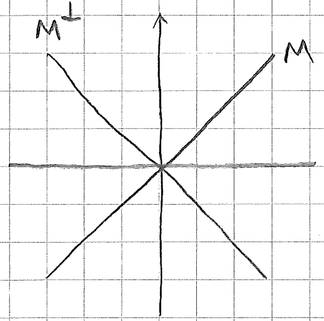
$$\langle x+y, m \rangle = \underbrace{\langle x, m \rangle}_{=0} + \underbrace{\langle y, m \rangle}_{=0} = 0 \quad \forall\, m \in M \Rightarrow x+y \in M^\perp$$

$$\langle \alpha x, m \rangle = \alpha \langle x, m \rangle = 0 \quad \forall\, m \in M \Rightarrow \alpha x \in M^\perp$$

$\Rightarrow M^\perp$ is subspace.

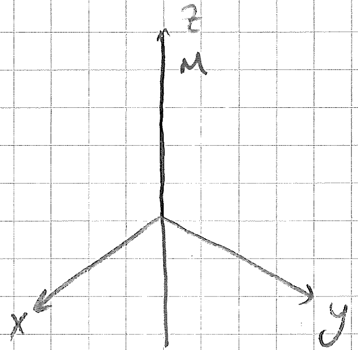
Theorem: Let V be an inner product space and M be a subspace of V. Then,

$$V = M \overset{\perp}{\oplus} M^\perp$$



$V = \mathbb{R}^2$

$$\mathbb{R}^2 = M \overset{\perp}{\oplus} M^\perp$$



$M^\perp$ = xy-plane

$$\mathbb{R}^3 = \underset{\text{axis}}{z} \overset{\perp}{\oplus} \underset{\text{plane}}{x\text{-}y}$$

Proof:

$$V = M \overset{\perp}{\oplus} M^{\perp} \left\{ \begin{array}{l} V = M + M^{\perp} \\ \text{lineer independence} \\ M \perp M^{\perp} \end{array} \right.$$

$V = M + M^{\perp}$, lineer independence } needs to be proved !

$M \perp M^{\perp}$ } by definition

Lineer Independence:

$x \in M$

$y \in M^{\perp}$ $\quad x + y = 0_V$

$$\langle \underbrace{x+y}_{0_V}, x \rangle = 0 = \langle x, x \rangle + \underbrace{\cancel{\langle y, x \rangle}}_{0}$$

$0 \to y \in M^{\perp}$, $x \in M$

$$\Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0_V$$

$$\langle x+y, y \rangle = 0 = \underbrace{\cancel{\langle x, y \rangle}}_{0} + \langle y, y \rangle$$

$$\Rightarrow \langle y, y \rangle = 0 \Rightarrow y = 0_V$$

M & $M^{\perp}$ are l.i

Proof: $V = M + M^{\perp}$

We have to show that every $x \in V$ can be written as the sum of $x_1 \in M$ & $x_2 \in M^{\perp}$

$$x = \underbrace{x_1}_{\in M} + \underbrace{x_2}_{\in M^{\perp}}$$

FABER-CASTELL

$\dim V = n$

Part-I: $\dim M = n$ $\quad M \subset V$

$\dim M = \dim V$

$\Rightarrow M = V$ $\qquad$ Let $x \in V$ be arbitrary

$M^{\perp} = \{0_V\}$

$$x = \underbrace{x}_{\substack{\in V \\ \in M}} + \underbrace{0_V}_{\in M^{\perp}}$$

! $0_V \begin{matrix} \in M \\ \in M^{\perp} \end{matrix} \Big\}$ does not violate linearly ind. case !

Part II: $\underbrace{\dim M}_{k} < n$

$k < n$

Let $B = \{m_1, m_2, \ldots, m_k\}$ be a basis for M.

Let $C = \{m_1, m_2, \ldots, m_k, u_1, u_2, \ldots, u_{n-k}\}$ be a basis for V

Apply Gram-Schmidt orthonormalization process to C. Suppose we obtain

$$\bar{C} = \{e_1, e_2, \ldots, e_k, f_1, f_2, \ldots, f_{n-k}\}$$

$\bar{C}$ is a basis for V.

$\bar{B} = \{e_1, e_2, \ldots, e_k\}$ is basis for M.

Let x be an arbitrary element of V.

$$x = \underbrace{\sum_{i=1}^{k} \alpha_i e_i}_{= x_1} + \underbrace{\sum_{j=1}^{n-k} \beta_j f_j}_{= x_2} = x_1 + x_2$$

$x_1 \in M$ & if we show $x_2 \in M^{\perp}$, we are done!

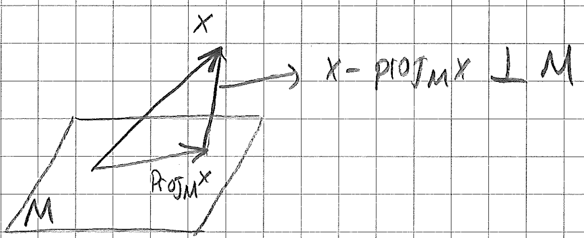
FABER-CASTELL

$$\langle x_2, \underset{y \in M}{m} \rangle = 0 \quad \forall\, m \in M$$

$$\left\langle \sum_{j=1}^{n-k} \beta_j f_j, \sum_{i=1}^{k} \gamma_i e_i \right\rangle = \sum_{j=1}^{n-1} \sum_{i=1}^{k} \beta_j \overline{\gamma_i} \underbrace{\langle f_j, e_i \rangle}_{\substack{=0 \ \forall i,j \\ (\bar{E} \text{ is orthonormal})}} = 0$$

$$\Rightarrow x_2 \in M^{\perp}$$

$$\Rightarrow \quad V = M + M^{\perp} = M \overset{\perp}{\oplus} M^{\perp}$$



THEOREM: (Projection Theorem)

Let H be a Hilbert space and let M a finite dimensional subspace of H. For any $x \in H$, the following minimization problem has a solution.

$$\min_{m \in M} \| x - m \|$$

In other words, we can find $\underbrace{\text{the closest vector}}^{\text{call } m^*}$ to x lying in the subspace M. If $m^* \in M$ is a solution, then $x - m^*$ is orthogonal to M. Furthermore, the solution m^* is unique and it is the orthogonal projection of x onto M.

$$m^* = \mathrm{Proj}_M x$$

Proof: $H = M \overset{\perp}{\oplus} M^\perp$

Any $x \in H$ can be written as $x = \underbrace{x_1}_{\in M} + \underbrace{x_2}_{\in M^\perp}$

$$\|x-m\|^2 = \|x_1+x_2-m\|^2 = \|\overbrace{\underbrace{x_1}_{\in M}-\underbrace{m}_{\in M}}^{\in M}+\underbrace{x_2}_{\in M^\perp}\|^2$$

$$= \langle x_1 - m + x_2 , x_1 - m + x_2 \rangle$$

$$= \langle x_1 - m, x_1 - m \rangle + \cancelto{0}{\langle x_1 - m, x_2 \rangle} + \cancelto{0}{\langle x_2, x_1 - m \rangle} + \langle x_2, x_2 \rangle$$

$$= \langle x_1 - m, x_1 - m \rangle + \langle x_2, x_2 \rangle$$

$$\Rightarrow \|x-m\|^2 = \|x_1 - m\|^2 + \|x_2\|^2 \geq \|x_2\|^2$$

$\boxed{m^* = x_1}$ → Decomp. is unique ⇒ m^* is unique

($H = M \overset{\perp}{\oplus} M^\perp$ is a direct sum decomp.)

$$x - m^* = x - x_1 = x_2 \perp M$$

24. 11. 2015
Tuesday

$M_1, M_2, \ldots, M_k \subset V$

$M_1 \oplus M_2 \oplus \ldots \oplus M_k = M$

$M_1, M_2, \ldots, M_k$ are linearly ind.

$\downarrow$

$m_1 + m_2 + \ldots + m_k = 0_V \Rightarrow m_i = 0 \quad \forall i$

If we have a direct sum, any element $m \in M$

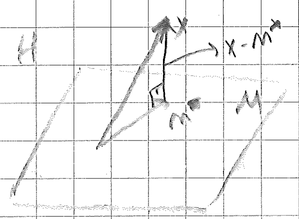
$$m = \underbrace{m_1 + m_2 + \ldots + m_k}_{\text{decomp. is unique}}$$

$M_i \perp M_j$

$\langle m_i, m_j \rangle = 0 \quad \forall\, i \neq j$

$M \subset V \quad M^\perp \triangleq \{ x \in V \,/\, \langle x, m \rangle = 0 \quad \forall m \in M \}$

$$V = M \overset{\perp}{\oplus} M^\perp$$



$$\min_{m \in M} \|x - m\|$$

m^* | always exists
| is unique

$x - m^* \perp M$

This is true for any inner product space. (Convergence of sequence)

Proof: $H = M \overset{\perp}{\oplus} M^{\perp}$ Any vector $x \in H$ $x = \underset{\in M}{x_1} + \underset{\in M^{\perp}}{x_2}$

$$\|x-m\|^2 = \|x_1+x_2-m\|^2 = \|x_1-m+x_2\|^2 = \langle x_1-m+x_2, x_1-m+x_2 \rangle$$

$$= \langle x_1-m, x_1-m \rangle + \underbrace{\langle x_1-m, x_2 \rangle}_{=0} + \underbrace{\langle x_2, x_1-m \rangle}_{=0} + \langle x_2, x_2 \rangle$$

$$= \|x_1-m\|^2 + \|x_2\|^2 \geq \|x_2\|^2$$

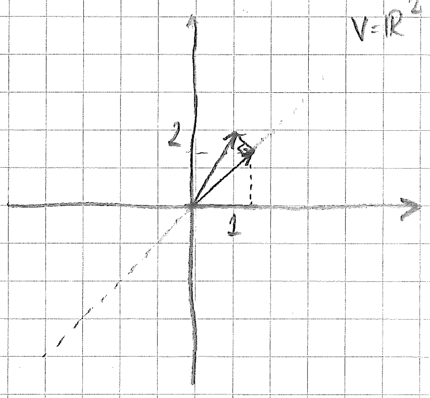
$\boxed{m^* = x_1}$ x_1 is unique $\Rightarrow$ m^* is unique.

$$x - m^* = x - x_1 = x_1 + x_2 - x_1 = x_2$$

$$\underbrace{x - m^* \perp M}_{\text{Orthogonality Condition}}$$

EXAMPLE: $V = \mathbb{R}^2$ $M = \{ \text{span} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \}$

$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ Find m^* which minimizes $\|x-m\|$



$m = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\langle x, y \rangle = x_1 y_1 + x_2 y_2$

$x - m^* \perp M$

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} - \alpha^* \begin{bmatrix} 1 \\ 1 \end{bmatrix} \perp \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1-\alpha^* \\ 2-\alpha^* \end{bmatrix} \perp \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$1 - \alpha^* + 2 - \alpha^* = 0$$

$$\Rightarrow \alpha^* = 3/2$$

FABER-CASTELL

$m^* = 3/2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

<u>EXAMPLE:</u> $M = \text{span}\{v_1\}$, x , find $m^* \in M$ minimizing $\|x - m\|$

$$m = \alpha v_1$$

$$x - \alpha^* v_1 \perp v_1$$

$$\langle x - \alpha^* v_1 , v_1 \rangle = 0$$

$$\langle x, v_1 \rangle - \alpha^* \langle v_1, v_1 \rangle = 0$$

$$\alpha^* = \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} \Rightarrow m^* = \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$$

<u>EXAMPLE:</u> $M = \text{span}\{\underbrace{v_1, v_2}_{LI}\}$, x

$$m = \alpha_1 v_1 + \alpha_2 v_2$$

$$x - \alpha_1 v_1 - \alpha_2 v_2 \perp v_1$$
$$x - \alpha_1 v_1 - \alpha_2 v_2 \perp v_2$$

$$\langle x - \alpha_1 v_1 - \alpha_2 v_2 , v_1 \rangle = 0$$

$$\langle x, v_1 \rangle - \alpha_1 \langle v_1, v_1 \rangle - \alpha_2 \langle v_2, v_1 \rangle = 0$$

$$\alpha_1 \langle v_1, v_1 \rangle + \alpha_2 \langle v_2, v_1 \rangle = \langle x, v_1 \rangle$$
$$\alpha_1 \langle v_1, v_2 \rangle + \alpha_2 \langle v_2, v_2 \rangle = \langle x, v_2 \rangle$$

$$\underbrace{\begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_2, v_1 \rangle \\ \langle v_1, v_2 \rangle & \langle v_2, v_2 \rangle \end{bmatrix}}_{G\text{ : Gramian Matrix}} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}}_{\alpha} = \underbrace{\begin{bmatrix} \langle x, v_1 \rangle \\ \langle x, v_2 \rangle \end{bmatrix}}_{b}$$

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

$$v_1 = v_2 \rightarrow G = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_1 \rangle \\ \langle v_1, v_1 \rangle & \langle v_1, v_1 \rangle \end{bmatrix} \rightarrow G \text{ is not invertable!}$$

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

$$\text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

G is invertable iff v_1 & v_2 are linearly independent. Other case, we canot apply this process.

GENERAL CASE:

$$M = \text{span}\{ \underbrace{v_1, v_2, \ldots, v_k}_{\text{a basis for } M} \}$$

$$m = \sum_{i=1}^{k} \alpha_i v_i \qquad x - \sum_{i=1}^{k} \alpha_i v_i \perp v_j \quad \forall j$$

$$\left\langle x - \sum_{i=1}^{k} \alpha_i v_i \, , \, v_j \right\rangle = 0 \quad \forall j$$

$$\langle x, v_j \rangle = \sum_{i=1}^{k} \langle v_i, v_j \rangle \alpha_i \quad \forall j$$

$$\underbrace{\begin{bmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_k, v_1 \rangle \\ \vdots & & \vdots \\ \langle v_1, v_k \rangle & \cdots & \langle v_k, v_k \rangle \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}}_{\alpha} = \underbrace{\begin{bmatrix} \langle x, v_1 \rangle \\ \langle x, v_2 \rangle \\ \vdots \\ \langle x, v_k \rangle \end{bmatrix}}_{b}$$

G is invertable iff $v_1, v_2, \dots, v_k$ are l.i.

$$\boxed{\alpha = G^{-1} b}$$

SPECIAL CASE $\{v_1, v_2, \dots, v_k\}$ is an orthonormal basis.

$$\langle v_i, v_j \rangle = 0 \quad \text{for} \quad i \neq j$$

$$\langle v_i ; v_i \rangle = 1$$

$$\underbrace{\begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{bmatrix}}_{k \times k} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \begin{bmatrix} \langle x, v_1 \rangle \\ \langle x, v_2 \rangle \\ \vdots \\ \langle x, v_k \rangle \end{bmatrix}$$

$$\alpha_i^* = \langle x, v_i \rangle$$

$$m^* = \sum_{i=1}^{k} \alpha_i^* v_i = \sum_{i=1}^{k} \langle x, v_i \rangle v_i$$

EXAMPLE: H: space of square integrable functions defined over $[-\pi, \pi]$ with the inner product

$$\langle f_1, f_2 \rangle = \int_{-\pi}^{\pi} f_1(t)\, \overline{f_2(t)}\, dt$$

$$M = \mathrm{span}\left\{ \underbrace{\frac{e^{jnt}}{\sqrt{2\pi}}}_{v_n} \right\}_{n=-N}^{n=N}$$

$$\langle v_n, v_m \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{jnt}\, e^{-jmt}\, dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j(n-m)t}\, dt = \delta_{nm} = \begin{cases} 1 & , n=m \\ 0 & , n\neq m \end{cases}$$

$$\alpha_n = \langle x, v_n \rangle$$

$$\boxed{\alpha_n = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} x(t)\, e^{-jnt}\, dt}$$

Finite Fourier Series Coefficients for X

$$\mathrm{Proj}_M\, x = \sum_{n=-N}^{N} \alpha_n^* x_n = \sum_{n=-N}^{N} \frac{1}{\sqrt{2\pi}} \left(\int_{-\pi}^{\pi} x(t)\, e^{-jnt}\, dt \right) \frac{e^{jnt}}{\sqrt{2\pi}}$$

$N \to \infty$ we have Fourier Series Representation of x.

PROJ THM

27. 11. 2015
Friday

$$V = M \overset{\perp}{\oplus} M^{\perp}$$

$$v = m_1 + m_2$$

$m_1 \in M$, $m_2 \in M^{\perp}$

m_1 & m_2 are unique

$$m_1 \perp m_2$$

$$\min_m \| x - m \| = x_1 = Proj_M\, x$$

$x \in V$, $m \in M$

$$M = sp\{v_1, v_2\}$$

$x \in V$

$$x - \alpha_1 v_1 - \alpha_2 v_2 \perp \begin{matrix} v_1 \\ v_2 \end{matrix}$$

$$\langle x - \alpha_1 v_1 - \alpha_2 v_2, v_j \rangle = 0 \quad j = 1, 2$$

$$\underbrace{\begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_2, v_1 \rangle \\ \langle v_1, v_2 \rangle & \langle v_2, v_2 \rangle \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}}_{\alpha} = \underbrace{\begin{bmatrix} \langle x, v_1 \rangle \\ \langle x, v_2 \rangle \end{bmatrix}}_{b} \Rightarrow \alpha = G^{-1} b$$

$$x_1 = \alpha_1^* v_1 + \alpha_2^* v_2$$

For $v_2 = 2v_1$

$$G = \begin{bmatrix} \langle v_1, v_1 \rangle & 2\langle v_1, v_1 \rangle \\ 2\langle v_1, v_1 \rangle & 4\langle v_1, v_1 \rangle \end{bmatrix} = \langle v_1, v_1 \rangle \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \to \text{not invertable}$$

APPLICATION OF PROJ. THM TO $\mathbb{C}^N$

$\mathcal{M} = \text{span}\{m_1, m_2, \ldots, m_k\}$

$m_i \in \mathbb{C}^n$ $\qquad$ $m_1, m_2, \ldots, m_k$ are l.i.

n elements $\left\{ \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix} \right.$ $\rightarrow \in \mathbb{C}$ $\qquad$ $m_1, m_2, \ldots, m_k$ is a basis for $\mathcal{M}$.

$\boxed{x \in \mathbb{C}^n}$ $\qquad$ $\boxed{k < n}$

So strictly true!

$\text{Proj}_{\mathbb{C}^n} x = x \rightarrow$ it's inside $\mathbb{C}^n$ itself.

$\text{Proj}_M x = ?$

Assume that it is projection of x!

$$\underbrace{m}_{\in M} = \alpha_1 m_1 + \alpha_2 m_2 + \ldots + \alpha_k m_k = \underbrace{\left[m_1 \,\vdots\, m_2 \,\vdots\, \ldots \,\vdots\, m_k \right]}_{\triangleq B} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}}_{\alpha}$$

$$m = B\alpha$$

$\langle x, y \rangle = y^H x$ $\qquad$ $\langle x - m, m_i \rangle = 0 \quad i = 1, \ldots, k$

$x, y \in \mathbb{C}^n$

$\langle x - B\alpha, m_i \rangle = 0 \quad i = 1, \ldots, k$

$$\left.\begin{array}{l} m_1^H (x - B\alpha) = 0 \\ m_2^H (x - B\alpha) = 0 \\ \vdots \\ m_k^H (x - B\alpha) = 0 \end{array}\right\} \quad \left.\underbrace{\begin{bmatrix} m_1^H \\ m_2^H \\ \vdots \\ m_k^H \end{bmatrix}}_{B^H} (x - B\alpha) = 0 \right\} \quad B^H (x - B\alpha) = 0$$

$$\underbrace{B^H B}_{G}\, \alpha = \underbrace{B^H x}_{b}$$

$$\begin{bmatrix} \langle m_1, m_1\rangle & \cdots & \langle m_k, m_1\rangle \\ \langle m_1, m_2\rangle & & \vdots \\ \vdots & & \vdots \\ \langle m_1, m_k\rangle & \cdots & \langle m_k, m_k\rangle \end{bmatrix} \alpha = \begin{bmatrix} \langle x, m_1\rangle \\ \langle x, m_2\rangle \\ \vdots \\ \langle x, m_k\rangle \end{bmatrix}$$

$$\alpha^* = G^{-1} b = (B^H B)^{-1} B^H x$$

$$\mathrm{Proj}_M\, x = B\alpha^* = \underbrace{B\,(B^H B)^{-1} B^H}_{P}\, x$$

P is the projection matrix onto M.

$$\underbrace{x - \mathrm{Proj}_M\, x}_{M^\perp} = x - B(B^H B)^{-1} B^H x = \Big(\underbrace{I - B(B^H B)^{-1} B^H}_{\text{projection matrix onto } M^\perp}\Big)\, x$$

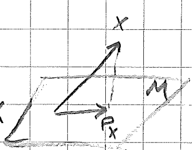
<u>**REMARK:**</u> An orthogonal projection matrix $P \in \mathbb{C}^{k\times k}$ satisfies

i) $P^H = P$

ii) $P^2 = P$

$x \in \mathbb{C}^n$

$$\underbrace{P \underbrace{x}_{\in \mathbb{C}^n}}_{\in M}\, P = Px$$



$P^2 x = Px$

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

$$P = B(B^H B)^{-1} B^H$$

$$P^2 = B\,(B^H B)^{-1} \underbrace{B^H B\,(B^H B)^{-1}}_{= I} B^H$$

$$= B\,(B^H B)^{-1} B^H = P$$

$$P^3 = \underbrace{P^2}_{P} P = P^2 = P$$

$P^l = P \quad \forall l$

EXAMPLE: Find the orthogonal projection of the vector $\begin{bmatrix}1\\0\\0\end{bmatrix}$ onto the subspace spanned by $\underbrace{\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\}}_{\text{l.i. set}}$

$$B = \begin{bmatrix}1 & 1\\1 & -1\\1 & 1\end{bmatrix} \qquad B^H B = B^T B = \begin{bmatrix}1 & 1 & 1\\1 & -1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\1 & -1\\1 & 1\end{bmatrix} = \begin{bmatrix}3 & 1\\1 & 3\end{bmatrix}$$

$$(B^T B)^{-1} = \frac{1}{8}\begin{bmatrix}3 & -1\\-1 & 3\end{bmatrix}$$

$$P = B(B^H B)^{-1} B^H = B(B^T B)^{-1} B^T = \frac{1}{8}\begin{bmatrix}1 & 1\\1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}3 & -1\\-1 & 3\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\1 & -1 & 1\end{bmatrix}$$

$$= \begin{bmatrix}1/2 & 0 & 1/2\\0 & 1 & 0\\1/2 & 0 & 1/2\end{bmatrix}$$

$$x_1 = \mathrm{Proj}_M x = Px = \begin{bmatrix}1/2\\0\\1/2\end{bmatrix}$$

$$x - x_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} - \begin{bmatrix}1/2\\0\\1/2\end{bmatrix} = \begin{bmatrix}1/2\\0\\-1/2\end{bmatrix} \Rightarrow \begin{array}{l}\left\langle \begin{bmatrix}1/2\\0\\-1/2\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\rangle = 0\\[2ex] \left\langle \begin{bmatrix}1/2\\0\\-1/2\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\rangle = 0\end{array}$$

APPLICATION OF PROJ THEOREM TO SOLUTIONS OF LINEAR EQNS.

$$\underbrace{A}_{\in \mathbb{C}^{m\times n}} x = \underbrace{b}_{\in \mathbb{C}^{m}}$$

A & b are known

$x \in \mathbb{C}^n$ is unknown

1) Is there a solution?

2) If there is a solution, is it unique?

EXAMPLE:

$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad b = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$m = 2$
$n = 1$

$\underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{A} x = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \rightarrow$ There is a solution
The solution is unique.

$R(A) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ dim = 1

$N(A) = \text{span}\{ 0 \}$ dim = 0

$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \rightarrow$ Is there a solution? NO

Same $R(A)$ & $N(A)$

Is $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ in $R(A)$? NO

$$\underbrace{\begin{bmatrix}1 & 2\\ 1 & 2\end{bmatrix}}_{A} \underset{\begin{bmatrix}x_1\\ x_2\end{bmatrix}}{\underset{\downarrow}{x}} = \underbrace{\begin{bmatrix}2\\ 2\end{bmatrix}}_{b} \longrightarrow x = \begin{bmatrix}0\\ 1\end{bmatrix}$$

$R(A) = \text{span}\left\{\begin{bmatrix}1\\ 1\end{bmatrix}\right\}$ dim=1

$N(A) = \text{span}\left\{\begin{bmatrix}-2\\ 1\end{bmatrix}\right\}$ dim=1

$$\underbrace{\dim R(A)}_{1} + \underbrace{\dim N(A)}_{1} = \underbrace{\dim V}_{2}$$

There is a solution. $\longrightarrow \begin{bmatrix}2\\ 0\end{bmatrix}, \begin{bmatrix}0\\ 1\end{bmatrix}, \begin{bmatrix}1\\ 1/2\end{bmatrix}$

x & (x+a) are soln.

$$Ax = b$$

$$A(\underset{\in N(A)}{x+a}) = Ax + \underbrace{Aa}_{0} = b$$

$\Downarrow$

The solution is NOT unique. (dim N(A) is diff. than 0, so solution is NOT unique)

ANSWERS

1) Yes, iff $b \in R(A)$

2) Yes, iff $N(A) = \{0\}$

$\dim N(A) = 0$

$$\underbrace{\begin{bmatrix}1\\ 1\end{bmatrix}}_{A} x = \underbrace{\begin{bmatrix}2\\ 1\end{bmatrix}}_{b}$$

$b \notin R(A) \Rightarrow$ There is no solution.

If there is no exact solution, we go for a "best approximation" i.e, we look for an $\hat{x}$ that would solve the following minimization problem

$$\min_{x \in \mathbb{C}^n} \|Ax - b\| \geq 0$$

$$\Big\downarrow \text{equivalent problem}$$

$$\min_{y \in R(A)} \|y - b\|$$

$$b = \underbrace{b_1}_{\in R(A)} + \underbrace{b_2}_{\in R(A)^{\perp}}$$

$$V = M \overset{\perp}{\oplus} M^{\perp}$$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \end{array} \qquad R(A) \subset \underset{\text{co-domain}}{\mathbb{C}^m}$$

$$\mathbb{C}^m = R(A) \overset{\perp}{\oplus} R(A)^{\perp}$$

$$b \in \mathbb{C}^m$$

$$b = \underbrace{b_1}_{\in R(A)} + \underbrace{b_2}_{\in R^{\perp}(A)}$$

$$\|Ax - b\|^2 = \|Ax - b_1 - b_2\|^2 = \langle \underbrace{Ax - b_1}_{\in R(A)} \underbrace{- b_2}_{\in R^{\perp}(A)}, Ax - b_1 - b_2 \rangle$$

$$= \langle Ax - b_1, Ax - b_1 \rangle - \cancelto{0}{\langle Ax - b_1, b_2 \rangle} - \cancelto{0}{\langle b_2, Ax - b_1 \rangle} + \langle b_2, b_2 \rangle$$

$$= \|Ax - b_1\|^2 + \|b_2\|^2 \geq \|b_2\|^2$$

We can always find an $\hat{x}$ values which satisfies $A\hat{x} = b_1$ since

$$b_1 \in R(A)$$

The minimum value is obtained at $x = \hat{x}$.

$$\begin{bmatrix}1\\1\end{bmatrix} x = \begin{bmatrix}2\\1\end{bmatrix} \qquad b_1 = \mathrm{Proj}_{R(A)}\, b$$

$$\mathrm{Proj}_{R(A)} = B(B^H B)^{-1} B^H \quad \& \quad B = \begin{bmatrix}1\\1\end{bmatrix}$$

$$R(A) = \mathrm{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$$

$$B^T B = 2 \Rightarrow \mathrm{Proj}_{R(A)} = \frac{\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}}{2} = \frac{1}{2}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$$

$$b_1 = \mathrm{Proj}_{R(A)}\, b = \frac{1}{2}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}3/2\\3/2\end{bmatrix}$$

$A\hat{x} = b_1 \rightarrow \begin{bmatrix}1\\1\end{bmatrix}\hat{x} = \begin{bmatrix}3/2\\3/2\end{bmatrix} \Rightarrow \boxed{\hat{x} = 3/2} \rightarrow$ The solution $\hat{x}$ is unique

<u>EXAMPLE:</u>

$$A = \begin{bmatrix}1 & 2\\1 & 2\end{bmatrix} \qquad b = \begin{bmatrix}2\\1\end{bmatrix}$$

$$\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}2\\1\end{bmatrix}$$

$R(A) = \mathrm{Span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$ & $b \notin R(A) \Rightarrow$ There is NO exact solution!

$$b_1 = \mathrm{Proj}_{R(A)}\, b = \begin{bmatrix}3/2\\3/2\end{bmatrix}$$

$A\hat{x} = b_1 \Rightarrow \begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\begin{bmatrix}\hat{x}_1\\\hat{x}_2\end{bmatrix} = \begin{bmatrix}3/2\\3/2\end{bmatrix} \rightarrow$ There is a solution but it is NOT unique!

$$\begin{bmatrix}1/2\\1/2\end{bmatrix}, \begin{bmatrix}3/2\\0\end{bmatrix}$$

Because $\dim N(A) \neq 0$

$$A(\hat{x} + a) = A\hat{x} + \underbrace{Aa}_{0} = b_1$$

$$a \in N(A)$$

$$A(\hat{x} + a) = b_1$$

$N(A) = \mathrm{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\} \neq \{0\} \Rightarrow \hat{x}$ is NOT unique.

Suppose $\hat{x}$ is a solution of $A\hat{x} = b_1$

Suppose that m is any vector in $N(A)$

Then $\hat{x}+m$ is another solution.

$$\left.\begin{array}{l} A\hat{x} = b_1 \\ m \in N(A) \end{array}\right\} \quad A(\hat{x}+m) = b_1$$

In the case of non-uniqeness, we are going to look for the solution with min norm.

$$\min_{A\hat{x}=b} \|\hat{x}\|$$

$$N(A) \subset \mathbb{C}^n \Rightarrow \mathbb{C}^n = N(A) \overset{\perp}{\oplus} N^{\perp}(A)$$

Take an arbitrary solution $\hat{x}$ satisfying $A\hat{x} = b_1$. We can decompose $\hat{x}$ as,

$$x = \underbrace{\hat{x}_1}_{\in N^{\perp}(A)} + \underbrace{\hat{x}_2}_{\in N(A)}$$

We are going to find min norm of solution.

$$\underbrace{\|\hat{x}+m\|^2}_{m \in N(A)} = \|\underbrace{\hat{x}_1}_{\in N^{\perp}(A)} + \underbrace{\hat{x}_2 + m}_{\in N(A)}\|^2 = \|\hat{x}_1\|^2 + \|\hat{x}_2 + m\|^2 \geq \|\hat{x}_1\|^2$$

$$m^* = -\hat{x}_2$$

$\hat{x}+m = \hat{x} + (-\hat{x}_2) = \hat{x}_1 \rightarrow$ The minimum norm solution is $\hat{x}_1$

$$\hat{x}_1 = \mathrm{Proj}_{N^{\perp}(A)}\, \hat{x}$$

$$b_1 = \mathrm{Proj}_{R(A)}\, b$$

01.12.2015
Tuesday

$M \subset V = \mathbb{C}^n$

$M = \text{span}\{m_1, m_2, \ldots, m_k\}$

$\in \mathbb{C}^n$

$x \in \mathbb{C}^n$

$$\text{Proj}_M\, x = B(B^H B)^{-1} B^H x$$

$$B = [m_1 \mid m_2 \mid \ldots \mid m_k]$$

$B^H B$ is invertable iff $\{m_1, \ldots, m_k\}$ is l.i.

$$Ax = b$$

$A \in \mathbb{C}^{m \times n}$, $x \in \mathbb{C}^n$, $b \in \mathbb{C}^m$

- There exist an exact solution iff $b \in R(A)$

If b is not in $R(A)$

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|$$

Solve: $A\hat{x} = b_1$ & $b_1 = \text{Proj}_{R(A)}\, b$

- The solution is unique iff $N(A) = \{0\}$

$\dim N(A) = 0$

$$A(\hat{x} + m) = b_1$$

$m \in N(A)$

$\min \|\hat{x}\|$ $A\hat{x} = b_1$ $\hat{x}_1 = \text{Proj}_{N^\perp(A)}\, \hat{x}$

It is difficult to make a projection onto $N^{\perp}(A)$. Therefore, we are going to use the following theorem

Theorem: $N^{\perp}(A) = R(A^H)$

Proof: We are going to prove $N(A) = R^{\perp}(A^H)$

Part-I: Prove $N(A) \subset R^{\perp}(A^H)$

Let $x \in N(A) \Rightarrow Ax = 0$

$\langle x, y \rangle = y^H x \Rightarrow \langle Ax, y \rangle = 0$ y is arbitrary

$y^H A x = 0$

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

$\langle x, A^H y \rangle = 0$

$\Downarrow$

$x \perp A^H y \Rightarrow x \perp R(A^H) \Rightarrow x \in R^{\perp}(A^H)$

Part-II: Prove $R^{\perp}(A^H) \subset N(A)$

$x \in R^{\perp}(A^H)$

$x \perp A^H y$ y is arbitrary

$\langle x, A^H y \rangle = 0$

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

$\langle Ax, y \rangle = 0 \Rightarrow Ax = 0 \Rightarrow x \in N(A)$

Example:

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$R(A) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \notin R(A)$

$$b_1 = \text{Proj}_{R(A)}\, b = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}3/2\\3/2\end{bmatrix}$$

$$A\hat{x} = b_1$$

$$\begin{bmatrix}1 & 2\\ 1 & 2\end{bmatrix}\hat{x} = \begin{bmatrix}3/2\\3/2\end{bmatrix} \qquad \hat{x} = \begin{bmatrix}3/2\\0\end{bmatrix}$$

$$N(A) = \text{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\} \qquad A^H = \begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$$

$$\underset{\substack{\| \\ R(A^H)}}{N^\perp(A)} = \text{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\} \Rightarrow R(A^H) = \text{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$$

$$\hat{x}_1 = \underbrace{\text{Proj}_{R(A^H)}}_{B(B^HB)^{-1}B^H}\hat{x}$$

$$B = \begin{bmatrix}1\\2\end{bmatrix} \qquad B^HB = 5 \qquad \text{Proj}_{R(A^H)} = \frac{BB^H}{5} = \frac{1}{5}\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix} \quad \& \quad \hat{x} = \begin{bmatrix}3/2\\0\end{bmatrix}$$

$$\hat{x}_1 = \frac{1}{5}\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}\begin{bmatrix}3/2\\0\end{bmatrix} = \frac{1}{5}\begin{bmatrix}3/2\\3\end{bmatrix}$$

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} & \\ & \end{bmatrix}\end{array}$$

$$\underset{\substack{\| \\ (n-k)}}{\dim N(A)} + \underset{\substack{\| \\ (k)}}{\dim R(A)} = \dim \mathbb{C}^n$$

$$(n-k) + (k) = n$$

$$N(A) \subset \mathbb{C}^n \qquad \dim N^\perp(A) = k$$

$$\|$$

$$\dim R(A^H) = k$$

$\dim R(A) = k$ # of linearly independent columns of A
$\dim R(A^H) = k$ # of linearly independent rows of A
} k is called to be the rank of matrix.

Special Cases:

1) Columns of A are l.i

A is full column rank

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} A \end{bmatrix} \end{array} x = \begin{bmatrix} \\ \end{bmatrix} \qquad m \geq n$$

There are more equations than unknowns (Over determined)

$$Ax = b$$

$$\dim R(A) = n \Rightarrow \dim N(A) + \underbrace{\dim R(A)}_{n} = \underbrace{\dim V}_{n}$$

$$\Rightarrow \dim N(A) = 0 \Rightarrow N(A) = 0$$

If there exist a solution $\Rightarrow$ It is unique.

$$b_1 = \underbrace{\mathrm{Proj}_{R(A)}}_{B(B^HB)^{-1}B^H} b$$

$$B = A \Rightarrow b_1 = A(A^HA)^{-1}A^H b$$

$$A\hat{x} = b_1 = A(A^HA)^{-1}A^H b$$

$$\Rightarrow A\left(\hat{x} - (A^HA)^{-1}A^Hb\right) = 0 \Rightarrow \hat{x} = (A^HA)^{-1}A^H b$$

$\hat{x}$ is unique

$$\underbrace{\mathrm{Proj}_{N^{\perp}(A)}}_{\mathbb{C}^n} \hat{x} = \hat{x}$$

$(A^HA)^{-1}A^H$ is called to be Left Pseudo Inverse of A.

04.12.2015
Friday

$$\underset{m\times n}{A}\ \underset{n\times 1}{x} = \underset{m\times 1}{b}$$

- $\exists$ a solution iff $b \in \underbrace{R(A)}_{\text{column space}}$

- A solution is unique iff $N(A) = \{0\}$
  $\dim N(A) = 0$

$Ax = b$

There is no exact solution:

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|$$

$$b_1 = \mathrm{Proj}_{R(A)}\, b$$

$$A\hat{x} = b_1$$

If $N(A) = \{0\}$, $\hat{x}$ is unique

If not, solution is not unique

$$\min_{A\hat{x} = b_1} \|\hat{x}\|$$

$$\hat{x}_1 = \mathrm{Proj}_{N^{\perp}(A)}\, \hat{x}$$ any solution

$$N^{\perp}(A) = R(A^H)$$

$$\hat{x}_1 = \mathrm{Proj}_{R(A^H)}\, \hat{x}$$

FABER-CASTELL

$$\dim \underbrace{R(A)}_{\substack{\text{column}\\ \text{space}}} = \dim \underbrace{R(A^H)}_{\substack{\text{row space}\\ \text{of } A.}}$$

$$\begin{array}{cc} & n \\ m & \left[\begin{array}{ccc} & & \\ & & \\ & & \end{array}\right] \end{array}$$

# of l.i columns of $A = k$ = # of l.i rows of A.

k = Rank of the matrix A.

SPECIAL CASE

1) Column of A are l.i

$$\begin{array}{cc} & n \\ m & \left[\begin{array}{cc} & \\ & \end{array}\right] \end{array} \quad m \geq n$$

A is $\underbrace{\text{full column rank}}_{\text{over determined}}$

$\dim R(A) = n$

$$\underbrace{\dim N(A)}_{=0} + \underbrace{\dim R(A)}_{n} = n \Rightarrow N(A) = \{0\}$$

If we find a solution $\Rightarrow$ it has to be unique

$Ax = b$

$$b_1 = \underbrace{\mathrm{Proj}_{R(A)}}_{B(B^HB)^{-1}B^H} b$$

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

$B = [m_1 \mid m_2 \mid \quad \mid m_k]$

$B = A$

$b_1 = A(A^HA)^{-1}A^H b$

$$A\hat{x} = A(A^H A)^{-1} A^H b$$

$$A\underbrace{\left(\hat{x} - (A^H A)^{-1} A^H b\right)}_{\in N(A)} = 0$$

$$\underbrace{\hat{x}}_{\text{unique solution}} = \underbrace{(A^H A)^{-1} A^H}_{\text{left pseudo-inverse of } A} b$$

$$\hat{x}_1 = \mathrm{Proj}_{\underbrace{N^\perp(A)}_{\mathbb{C}^n}} \hat{x} = \hat{x}$$

SPECIAL CASE II

Row of A are l.i.

$\Updownarrow$

A is full row rank

$$m\underbrace{\begin{bmatrix} & \overset{n}{} & \end{bmatrix}}\begin{bmatrix} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \end{bmatrix}$$

$m < n$

Underdetermined system of equations.

$\dim R(A^H) = m$

$\Vert$

$\dim R(A) = m$

$R(A) \subset \mathbb{C}^m \longrightarrow m\begin{bmatrix} & \overset{n}{} & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$R(A) = \mathbb{C}^m$

m dimensional

$Ax = b$

$\downarrow \mathbb{C}^m$

Solution always exists. Because $b \in R(A)$

$$b_1 = \text{Proj}_{R(A)} b$$

$$= \text{Proj}_{\mathbb{C}^m} b = b$$

$$A\hat{x} = b$$

$$\hat{x}_1 = \text{Proj}_{N^\perp(A)} \hat{x} = \underbrace{\text{Proj}_{R(A^H)}}_{B(B^HB)^{-1}B^H} \hat{x} \longrightarrow B = A^H$$

$$= A^H (AA^H)^{-1} A \hat{x}$$

$$\boxed{\hat{x}_1 = A^H (AA^H)^{-1} b}$$

↳ Right-pseudo inverse of A.

SPECIAL CASE III

A is both full column rank & full row rank. Columns are rows of A are l.i.

$$\begin{array}{cc} & n \\ m & \begin{bmatrix} & & \end{bmatrix} \end{array}$$

$m = n$

We have square matrix

$$\begin{array}{cc} & 5 \\ 3 & \begin{bmatrix} & & \end{bmatrix} \end{array}$$

↳ 3 ind. lin row ⇒ Rank = 3 ⇒ # of ind col. = 3 ⇒ n = 3 ⇒ n = m

$$Ax = b$$

$$m\begin{bmatrix} \overset{n}{\vdots} \end{bmatrix}$$

$N(A) = \{0\}$

$R(A) = \mathbb{C}^m$

$$x = A^{-1} b$$

solution always exists & it is unique.

$$x = (A^H A)^{-1} A^H b$$
$$= A^{-1} (A^H)^{-1} A^H b$$
$$= A^{-1} b$$

Left & right pseudo-inverse both turn into A^{-1}!

SPECIAL CASE IV

(there is no short cut, we have to apply all procedure)

A is neither full column rank nor full row rank

EXAMPLE:

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}_{b}$$

Obtain the unique minimum norm solution which minimize $\overset{\text{error}}{\|Ax - b\|}$.

$\hat{x}_1 = ?$

$$2\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

Rank of $A = 2$

$b \notin R(A)$

$$b_1 = \underbrace{\mathrm{Proj}_{R(A)}}_{B(B^H B)^{-1} B^H} b$$

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix}$$

$$B^H B = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 7 \\ 7 & 11 \end{bmatrix}$$

$$(B^H B)^{-1} = \frac{1}{72} \begin{bmatrix} 11 & -7 \\ -7 & 11 \end{bmatrix}$$

$$\underbrace{B(B^H B)^{-1} B^H}_{P} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix} \frac{1}{72} \begin{bmatrix} 11 & -7 \\ -7 & 11 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \end{bmatrix}$$

$$\Rightarrow \quad P = \frac{1}{72} \begin{bmatrix} 8 & 16 & 16 \\ 16 & 68 & -4 \\ 16 & -4 & 68 \end{bmatrix} \longrightarrow \text{Any proj matrix must be symmetric}$$

$$b_1 = Pb = P \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \frac{1}{72} \begin{bmatrix} 24 \\ 256 \\ -60 \end{bmatrix}$$

$$A\hat{x} = b$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{72} \begin{bmatrix} 24 \\ 256 \\ -60 \end{bmatrix}$$

$\dim(R(A)) = 2$

$\dim(N(A)) = 1$

$x_2 = 0$

$$\left.\begin{aligned} x_1 + x_3 &= \frac{1}{72} 24 \\ 3x_1 + x_3 &= \frac{-60}{72} \end{aligned}\right\} \quad \left.\begin{aligned} x_3 &= \frac{11}{12} \\ x_1 &= \frac{-7}{12} \end{aligned}\right\} \quad \hat{x} = \begin{bmatrix} -7/12 \\ 0 \\ 11/12 \end{bmatrix}$$

$$\hat{x}_1 = \mathrm{Proj}_{R(A^H)} \hat{x}$$

$$A^H = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 1 & 3 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$B^H B = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$P = \begin{bmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{bmatrix}$$

$$\hat{x}_1 = P \begin{bmatrix} -7/12 \\ 0 \\ 11/12 \end{bmatrix}$$

$$= \frac{1}{72} \begin{bmatrix} -46 \\ 8 \\ 62 \end{bmatrix} \Bigg\} \text{ Unique minimum norm minimum error solution}$$

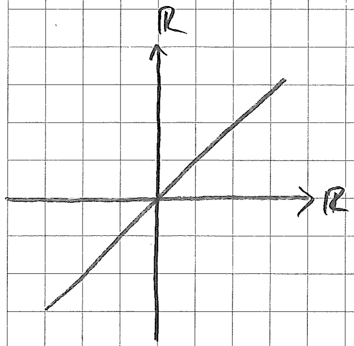
$$\begin{array}{c} A \\ \begin{bmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix} \end{array} \qquad A : V \to V$$

$$\left[\begin{array}{cc|c|c} x & x & 0 & 0 \\ x & x & & \\ \hline 0 & & x & 0 \\ \hline 0 & & 0 & x \end{array}\right]$$

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$\begin{bmatrix} x & x \\ x & x \end{bmatrix} \qquad \begin{bmatrix} x & x \\ x & x \end{bmatrix}$$

Invariance



$$M = \text{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$$

$$A\alpha\begin{bmatrix}1\\1\end{bmatrix} = \alpha\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \alpha\begin{bmatrix}3\\3\end{bmatrix} = 3\alpha\underbrace{\begin{bmatrix}1\\1\end{bmatrix}}_{\text{again subspace}}$$

M is invariant under A

SPECTRAL ANALYSIS OF LINEAR OPERATORS

DEFINITION: Let $A : V \to V$

be a linear transformation defined over vector space V.

A subspace M of V is said to be invariant under A if

$Ax \in M$ for all $x \in M$.

EXAMPLE: $R(A)$ is invariant under A

Let $x \in R(A)$ be an arbitrary vector in $R(A)$

$Ax \in R(A) \Rightarrow R(A)$ is invariant under A.

EXAMPLE: $N(A)$ is invariant under A.

$x \in N(A)$

$Ax = 0_V \in N(A) \Rightarrow N(A)$ is invariant under A.

DEFINITION: Powers of a linear transformation A are defined as

$$A^k x = \underbrace{A(A(\ldots\ldots(Ax)\ldots))}_{A \text{ is applied } k \text{ times}}$$

We define $A : V \to V$, otherwise $V \to W$, we cannot apply A more than one

$\cancel{A}\;\underbrace{Ax}_{\in V,\ \in W}$

In this way, polynomials of A can be constructed using linear combinations of power of A.

$$p(s) = \alpha_0 s^n + \alpha_1 s^{n-1} + \dots + \alpha_{n-1} s + \alpha_n$$

$$P(A) = \alpha_0 A^n + \alpha_1 A^{n-1} + \dots + \alpha_{n-1} A + \alpha_n I$$

identity transformation $I(x) = x$

Property: $\underbrace{(A^2 + A + 2I)}_{p(A)} Ax = A^3 x + A^2 x + 2Ax$

$$= A(A^2 x + Ax + 2x)$$

$$= A \underbrace{(A^2 + A + 2I)}_{p(A)} x$$

$p(A) . Ax = A\, p(A) x \Rightarrow p(A)$ & A commutes

EXAMPLE: $R(p(A))$ & $N(p(A))$ are invariant under A.

- $R(p(A))$ $\quad x \in R(p(A))$

$\exists y$ st $\quad x = p(A) y$

$$Ax = A\, p(A) y = p(A) \underbrace{Ay}_{y'} = p(A) y' \in R(p(A))$$

$R(p(A))$ is invariant under A.

- $N(p(A))$ $\quad x \in N(p(A))$

$$p(A) x = 0_V$$

$$P(A) Ax = A \underbrace{p(A) x}_{0_V} = 0_V \Rightarrow Ax \in N(p(A))$$

$\Rightarrow N(p(A))$ is invariant under A.

DEFINITION: Let A denote the matrix representation of a linear transformation from V to V. (A is a square matrix)

The eigenvalues of A denoted by λ_i are defined as the n-roots of the equation

$$\det(sI-A)=0$$

where $d(s)\triangleq\det(sI-A)$ is called as the characteristic polynomials of A.

Vectors $e_i\in\mathbb{C}^n$ satisfying $e_i\neq 0$ and $Ae_i=\lambda_i e_i$ are called as the eigenvectors of A.

$$A=\begin{bmatrix}2&1\\1&2\end{bmatrix}\qquad sI-A=\begin{bmatrix}s&0\\0&s\end{bmatrix}-\begin{bmatrix}2&1\\1&2\end{bmatrix}=\begin{bmatrix}s-2&-1\\-1&s-2\end{bmatrix}$$

$$I \to \begin{bmatrix}1&0\\0&1\end{bmatrix}$$

$$d(s)=\det(sI-A)=(s-2)^2-1=s^2-4s+3=0$$

$$\lambda_1=3\quad\lambda_2=1$$

$$\begin{bmatrix}2&1\\1&2\end{bmatrix}e_1=3e_1$$

$\downarrow$

We have to find non-zero vector. $\quad e_1=\begin{bmatrix}1\\1\end{bmatrix}$

$$\begin{bmatrix}2&1\\1&2\end{bmatrix}e_2=1e_2\longrightarrow e_2=\begin{bmatrix}1\\-1\end{bmatrix}$$

$$\begin{array}{ccc}V&\xrightarrow{A}&V\\\downarrow&&\downarrow\\\dim=n&&\dim=n\\\cong&&\cong\\\mathbb{C}^n&\xrightarrow{A}&\mathbb{C}^n\\{}[v]_B&&\end{array}$$

From this point on in this course, we are going to work on only the matrix representation of a linear transformation. However, we need to keep in mind that the matrix A represents a linear transformation and vectors we work with are representations wrt some basis sets.

$N(p(A))$ is invariant under A.

$N(p(A))$ is invariant under A.

$$y = \underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}_{A} x \qquad \mathbb{R}^2 = \text{span} \underbrace{\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}}_{B}$$

$$[y]_B = \bar{A} \, [x]_B$$

$$\underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}_{\bar{A}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \Rightarrow \quad \bar{A} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A e_1 = \lambda_1 e_1 \rightarrow (\lambda_1 I - A) e_1 = 0$$

$$e_1 \in N(\lambda_1 I - A)$$

$$e_2 \in N(\lambda_2 I - A)$$

$$\mathbb{R}^2 = N(\lambda_1 I - A) \oplus N(\lambda_2 I - A)$$

08.12.2015
Tuesday

$A : V \to V$

$\quad\ \ n \qquad n$

|||

$\underbrace{A}_{n\times n} \quad \mathbb{C}^n \to \mathbb{C}^n$

$M \subset V$

$x \in M \xrightarrow{Ax} Ax \in M$

$A : n \times n$

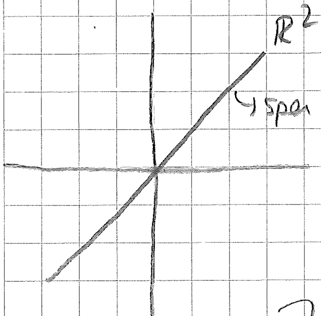
$\left.\underbrace{\det(sI - A)}_{\substack{d(s) \\ \text{char. poly}}} = 0 \quad \right\}$ char. Eqn.

$s^n + d_1 s^{n-1} + \dots + d_{n-1} s + d_n = 0$

$(s-\lambda_1)(s-\lambda_2) \dots (s-\lambda_n) = 0$

$A\underbrace{e_i}_{e_i \neq 0} = \lambda_i e_i$

span$\{e_i\}$



$e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\left.\begin{array}{l} N(p(A)) \\ R(p(A)) \end{array}\right\}$ invariant under A

EX Cont'd

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \Big\{ \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{e_1}, \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{e_2} \Big\}$$

$$Ae_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \overset{\lambda_1}{3} e_1$$

This means e_1 is eigen vector

$$Ae_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \overset{\lambda_2}{1} e_2$$

$\text{span}\Big\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big\}$ & $\text{span}\Big\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\Big\}$ are invariant under linear transf. A.

e_1 & e_2 are linearly independent $\Rightarrow$ spans are l. ind. $\Longrightarrow$ $\Rightarrow$ $\mathbb{R}^2 = \underbrace{\text{span}\Big\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big\}}_{M_1} \overset{\perp}{\oplus} \underbrace{\Big\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\Big\}}_{M_2}$

$$[y]_B = \bar{A}\,[x]_B$$

Choose $B = \Big\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\Big\}$

For $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \bar{A} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \bar{A} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \Bigg\} \Rightarrow \bar{A} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

$\Downarrow$ First column of $\bar{A}$ $\qquad$ $\Downarrow$ Second column of $\bar{A}$

$$\bar{A} \overset{?}{=} f(A)$$

$3e_1 = Ae_1$

$e_2 = Ae_2$

$$\underbrace{[e_1 \ \ e_2]}_{P} \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}_{\bar{A}} = A \underbrace{[e_1 \ \ e_2]}_{P} \Rightarrow \boxed{\bar{A} = P^{-1} A P}$$

Theorem: Consider the linear transformation $y = Ax$ where $A \in \mathbb{C}^{n\times n}$, $x, y \in \mathbb{C}^n$

Suppose that

i) $\mathbb{C}^n = M_1 \oplus M_2 \oplus \ldots \oplus M_k$

ii) Each subspace M_i is invariant under A.

Then, the transformation A has a block diagonal representation

$$\bar{A} = \begin{bmatrix} \bar{A}_1 & & & \\ & \bar{A}_2 & & \\ & & \ddots & \\ & & & \bar{A}_k \end{bmatrix}$$

where $\bar{A} = P^{-1}AP$

$$P = \left[P_1 \,|\, P_2 \,|\, \cdots \,|\, P_k \right]$$

$$P_i = \left[b_i^1 \;\; b_i^2 \;\; \cdots \;\; b_i^{n_i} \right]$$

n_i is the dimension of M_i and $b_i^1, \ldots, b_i^{n_i} \in \mathbb{C}^n$ are basis vectors for M_i. $\bar{A}_i$ is of size $n_i \times n_i$.

A specific case:

$\mathbb{C}^n = M_1 + M_2$ $\qquad$ $M_1 = \text{span}\{\overbrace{\underbrace{b_1^1}_{\mathbb{C}^5}, b_1^2}^{\text{l.i}}\}$

$n = 5$

$\qquad\qquad M_2 = \text{span}\{\overbrace{\underbrace{b_2^1, b_2^2, b_2^3}_{\in \mathbb{C}^5}}^{\text{l.i}}\}$

$\left.\begin{array}{l} n_1 = \dim M_1 = 2 \\ n_2 = \dim M_2 = 3 \\ A \in \mathbb{C}^{5\times 5} \end{array}\right\}$ M_1 & M_2 are both invariant under A.

$$P = [\, P_1 \mid P_2 \,]$$

$$P_1 = [\, b_1^1 \mid b_1^2 \,]$$

$$P_2 \, [\, b_2^1 \mid b_2^2 \mid b_2^3 \,]$$

$$P = [\, b_1^1 \mid b_1^2 \mid b_2^1 \mid b_2^2 \mid b_2^3 \,]$$

$$\bar{A} = P^{-1} A P \quad \text{or} \quad P\bar{A} = AP$$

$$[\, b_1^1 \; b_1^2 \mid b_2^1 \; b_2^2 \; b_2^3 \,] \begin{bmatrix} \alpha_1^1 & \beta_1^1 & 0 & 0 & 0 \\ \alpha_1^2 & \beta_1^2 & 0 & 0 & 0 \\ 0 & 0 & \alpha_2^1 & \beta_2^1 & \gamma_2^1 \\ 0 & 0 & \alpha_2^2 & \beta_2^2 & \gamma_2^2 \\ 0 & 0 & \alpha_2^3 & \beta_2^3 & \gamma_2^3 \end{bmatrix} = A \, [\, \underbrace{b_1^1 \; b_1^2}_{M_1} \;\; \underbrace{b_2^1 \; b_2^2 \; b_2^3}_{M_2} \,]$$

$\bar{A}_1$ (upper left block), $\bar{A}_2$ (lower right block)

b_1^1 must be written in terms of M_1 again. (invariant under A)

EXAMPLE:

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

$$M_1 = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{b_1^1}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{b_1^2} \right\}$$

$$M_2 = \text{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}_{b_2^1} \right\}$$

$\Rightarrow$ all vectors are l.i. so call these b_1^1, b_1^2, b_2^1

i) Is M_1 invariant under A?

$$A b_1^1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = 2b_1^1 + 0b_1^2 \in M_1$$

$$A b_1^2 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = 1b_1^1 + 2b_1^2 \in M_1$$

M_1 is invariant under A.

ii) Is M_2 invariant under A?

$$A b_2^1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 1b_2^1 \in M_2 \Rightarrow M_2 \text{ is invariant under } A.$$

$$R^3 = \underbrace{M_1}_{dim=2} \oplus \underbrace{M_2}_{dim=1}$$

iii) Change the basis in both domain & codomain to $\{b_1^1\ b_1^2\ b_2^1\}$

$$\bar{A} = \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$$

EXAMPLE:

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

- Find eigenvalues

$$sI - A = \begin{bmatrix} s-1 & 0 & 0 \\ 1 & s-2 & -1 \\ 0 & 0 & s-3 \end{bmatrix}$$

$$\det(sI - A) = (s-1)(s-2)(s-3) = 0$$

$\lambda_1 = 1$

$\lambda_2 = 2$

$\lambda_3 = 3$

$Ae_1 = e_1$

$(A - I)e_1 = 0$

$(A - \lambda_1 I)e_1 = 0 \quad \rightarrow \quad A - \lambda_1 I = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \qquad e_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

Column Rank $= 2 = \dim R(A)$

$\Rightarrow \dim N(A) = 1 \quad \Rightarrow N(A - \lambda_1 I) = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$

$$(A-\lambda_2 I) = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Rightarrow N(A-\lambda_2 I) = \text{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

$$(A-\lambda_3 I) = \begin{bmatrix} -2 & 0 & 0 \\ -1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad e_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \Rightarrow N(A-\lambda_3 I) = \text{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$$

$$\mathbb{R}^3 = \underbrace{N(A-\lambda_1 I)}_{1} \oplus \underbrace{N(A-\lambda_2 I)}_{1} \oplus \underbrace{N(A-\lambda_3 I)}_{1}$$

$$\bar{A} = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \Rightarrow A \text{ is diagonalizable}$$

11. 12. 2015
Friday

$$\mathbb{C}^7 = \underbrace{M_1}_{\substack{4 \\ \lambda_1}} \oplus \underbrace{M_2}_{\substack{2 \\ \lambda_2}} \oplus \underbrace{M_3}_{\substack{1 \\ \lambda_3}}$$

M_1, M_2 & M_3 are invariant under A.

$$\bar{A} = \left[\begin{array}{c|c|c} \overset{4\times 4}{\bar{A}_1} & & \\ \hline & \overset{2\times 2}{\bar{A}_2} & \\ \hline & & \overset{1\times 1}{\bar{A}_3} \end{array}\right]$$

$$\bar{A} = P^{-1} A P$$

$$P = \begin{bmatrix} P_1 & P_2 & P_3 \end{bmatrix}$$

$$P_i = \begin{bmatrix} b_i^1 & b_i^2 & \dots & b_i^{n_i} \end{bmatrix}$$

EXAMPLE:

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

$$d(s) = \det(sI - A) = (s-1)(s-2)(s-3) = 0$$

$$\lambda_1 = 1$$
$$\lambda_2 = 2$$
$$\lambda_3 = 3$$

$$N(A - \lambda_1 I) = N\left(\begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\right) = \text{span}\left\{\underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{e_1}\right\}$$

$$A e_1 = \lambda_1 e_1$$
$$(A - \lambda_1 I) e_1 = 0$$

$$A - \lambda_2 I = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow N(A - \lambda_2 I) = \text{span}\left\{\underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{e_2}\right\}$$

$$A - \lambda_3 I = \begin{bmatrix} -2 & 0 & 0 \\ -1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow N(A - \lambda_3 I) = \text{span}\left\{\underbrace{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}_{e_3}\right\}$$

$$\underset{\in \mathbb{R}^3}{x} = \alpha_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbb{R}^3 = \text{span}\{e_1\} \oplus \text{span}\{e_2\} \oplus \text{span}\{e_3\}$$

$$\mathbb{R}^3 = \underbrace{\underbrace{N(A - \lambda_1 I)}_{n_1 = 1}}_{M_1} \oplus \underbrace{\underbrace{N(A - \lambda_2 I)}_{n_2 = 1}}_{M_2} \oplus \underbrace{\underbrace{N(A - \lambda_3 I)}_{n_3 = 1}}_{M_3}$$

$\bar{A} = \begin{bmatrix} & \end{bmatrix}$ $\quad P = \begin{bmatrix} P_1 & P_2 & P_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}$

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

$$\begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = A\begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix}$$

A is diagonalizable.

THEOREM: Let A be an $n \times n$ matrix with n distinct eigenvalues. $\lambda_1, \lambda_2, \ldots, \lambda_n$.

$$\lambda_i \neq \lambda_j \text{ when } i \neq j$$

Then the set of eigenvectors $\{e_1, \ldots, e_n\}$ for a l.i. set. Moreover,

$$\text{Span}\{e_i\} = N(A - \lambda_i I)$$

Hence

$$\mathbb{C}^n = \underbrace{N(A-\lambda_1 I)}_{d_1=1} \oplus \underbrace{N(A-\lambda_2 I)}_{d_2=1} \oplus \cdots \oplus \underbrace{N(A-\lambda_n I)}_{d_n=1}$$

Since $N(A-\lambda_i I)$ is invariant under A, A has a diagonal representation.

$$\bar{A} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

$\bar{A} = E^{-1}AE$ where $E = [e_1 \mid e_2 \mid \cdots \mid e_n]$

Proof: We have to show that $\{e_1, \dots, e_n\}$ is a l.i. set

Consider e.g. A is 3×3 and $\{e_1, e_2, e_3\}$ are its eigen vectors

$$\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$$

Multiply both sides with $A - \lambda_1 I$ from the left

$$(A - \lambda_i I) e_j = \underbrace{A e_j}_{\lambda_j e_j} - \lambda_i e_j$$

$$= (\lambda_j - \lambda_i) e_j$$

$$\alpha_1 \underbrace{\cancel{(\lambda_1 - \lambda_1)}}_{0} e_1 + \alpha_2 (\lambda_2 - \lambda_1) e_2 + \alpha_3 (\lambda_3 - \lambda_1) e_3 = 0$$

$$\alpha_2 (\lambda_2 - \lambda_1) e_2 + \alpha_3 (\lambda_3 - \lambda_1) e_3 = 0$$

Let us multiply both sides with $A - \lambda_2 I$

$$\alpha_2 (\lambda_2 - \lambda_1) \underbrace{\cancel{(\lambda_2 - \lambda_2)}}_{0} e_2 + \alpha_3 (\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2) e_3 = 0$$

$$\alpha_3 \underbrace{(\lambda_3 - \lambda_1)}_{\neq 0} \underbrace{(\lambda_3 - \lambda_2)}_{\neq 0} \underbrace{e_3}_{\neq 0} = 0 \Rightarrow \alpha_3 = 0$$

distinct eigen values

from eigen vector definition

$$\Rightarrow \left.\begin{matrix} \alpha_3 = 0 \\ \alpha_2 = 0 \\ \alpha_1 = 0 \end{matrix}\right\} e_1, e_2, e_3 \text{ are l.i.}$$

$e_1, e_2, \ldots, e_n$ are l.i & $\{e_1, \ldots, e_n\}$ is a basis for $\mathbb{C}^n$.

$\downarrow$

$\in \mathbb{C}^n$

$$\mathbb{C}^n = \underbrace{\mathrm{span}\{e_1\}}_{\substack{\parallel \\ N(A-\lambda_1 I)}} \oplus \ldots \oplus \underbrace{\mathrm{span}\{e_n\}}_{\substack{\parallel \\ N(A-\lambda_2 I)}}$$

Prove that

$N(A-\lambda_i I) = \mathrm{span}\{e_i\} \rightarrow$ to prove that first prove $\mathrm{span}\{e_i\} \subset N(A-\lambda_i I)$

then $N(A-\lambda_i I) \subset \mathrm{span}\{e_i\}$

Part I) Let x be arbitrary in $\mathrm{span}\{e_i\}$ $\quad x \in \mathrm{span}\{e_i\}$

$$x = \alpha e_i \qquad (A-\lambda_i I)x = (A-\lambda_i I)\alpha e_i$$

$$= \alpha (A-\lambda_i I) e_i = \alpha (\underbrace{\lambda_i - \lambda_i}_{=0}) e_i = 0$$

$$\Rightarrow \mathrm{span}\{e_i\} \subset N(A-\lambda_i I)$$

Part II) $\underset{\in \mathbb{C}^n}{\underbrace{x}} \in N(A-\lambda_i I)$

$$x = \sum_{j=1}^{n} \alpha_j e_j$$

$$\underbrace{x}_{\in N(A-\lambda_i I)} - \underbrace{\alpha_i e_i}_{\in N(A-\lambda_i I)} = \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j e_j$$

$$0 = \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j \underbrace{(A-\lambda_i I) e_j}_{(\lambda_j - \lambda_i) e_j}$$

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_j (\lambda_j - \lambda_i) e_j = 0$$

$$\alpha_j \underbrace{(\lambda_j - \lambda_i)}_{\neq 0} = 0 \quad \forall j \neq i$$

$$\alpha_j = 0 \quad \forall j \neq i \quad \Rightarrow \quad x - \alpha_i e_i = 0 \quad \Rightarrow \quad x = \alpha_i e_i$$

$$\Rightarrow \quad x \in \text{span}\{e_i\}$$

$$\Rightarrow \quad N(A - \lambda_i I) \subset \text{span}\{e_i\}$$

$$\Rightarrow \quad N(A - \lambda_i I) = \text{span}\{e_i\}$$

$$\mathbb{C}^n = \underbrace{N(A - \lambda_1 I)}_{d_1 = 1} \oplus \cdots \oplus \underbrace{N(A - \lambda_n I)}_{d_n = 1}$$

$d(s) = \det(sI - A)$

$d(s) = s^n + d_1 s^{n-1} + \cdots + d_{n-1} s + d_n$

$\deg(d(s)) = n$

Distinct eigenvalue case

$$d(s) = (s - \lambda_1)(s - \lambda_2) \cdots (s - \lambda_n)$$

General Case:

$$d(s) = (s - \lambda_1)^{r_1} (s - \lambda_2)^{r_2} \cdots (s - \lambda_\sigma)^{r_\sigma}$$ r_i: multiplicity of λ_i in $d(s)$

$$r_1 + r_2 + \cdots + r_\sigma = n$$

↳ ALGEBRAIC MULTIPLICITY

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad d(s) = \det(sI - A)$$

$$= \det\left(\begin{bmatrix} s-2 & -1 \\ -1 & s-2 \end{bmatrix}\right)$$

$$= s^2 - 4s + 3$$

$$d(A) = A^2 - 4A + 3I$$

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

$$= \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$d(A) = 0_{n \times n}$$

Theorem: (Cayley Hamilton Theorem)

Every $n \times n$ matrix satisfies its characteristic eqn.

$$d(A) = 0_{n \times n}$$

$$A^n + d_1 A^{n-1} + \dots + d_n I = 0_{n \times n}$$

REMARK: Cayley Hamilton Thm basicly says that n^{th} and higher order powers of an $n \times n$ matrix can be written as a linear combination of its lower powers.

$$A^n = -d_1 A^{n-1} - d_2 A^{n-2} - \dots - d_n I$$

For previous example,

$$A^2 - 4A + 3I = 0$$

$$A^2 = 4A - 3I$$

$$A^3 = AA^2 = 4A^2 - 3A$$
$$= 4(4A - 3I) - 3A$$
$$= 13A - 12I$$

$$\underset{n \geq 2}{A^n} = \alpha A + \beta I$$

$$\exp(s) = \sum_{n=0}^{\infty} \frac{s^n}{n!}$$

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

$$= \alpha A + \beta I$$

$$A^{-1} = ?$$

$$A^2 - 4A + 3I = 0$$

$$A - 4I + 3A^{-1} = 0 \Rightarrow A^{-1} = \frac{-A + 4I}{3}$$

$$f(A) = \alpha A + \beta I$$

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad A^{100} = ?$$

$$A^{100} = \alpha A + \beta I$$

$\lambda_1 = 3 \quad e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ → from earlier lectures

$\lambda_2 = 1 \quad e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$A^{100} = \alpha A + \beta I$

$\rightarrow A^{100} e_1 = \alpha \underbrace{A e_1}_{\lambda_1 e_1} + \beta e_1 \longrightarrow A e_1 = \lambda_1 e_1$

$A^2 e_1 = \lambda_1 A e_1 = \lambda_1^2 e_1$

$3^{100} e_1 = 3\alpha e_1 + \beta e_1$

$3^{100} e_1 = (3\alpha + \beta) e_1 \qquad A^n e_1 = \lambda_1^n e_1$

$\boxed{3^{100} = 3\alpha + \beta}$

$\rightarrow A^{100} e_2 = \alpha A e_2 + \beta e_2$

$e_2 = \alpha e_2 + \beta e_2$

$\boxed{\alpha + \beta = 1}$

$\Rightarrow \quad \alpha = \frac{3^{100} - 1}{2}$

$\beta = \frac{3 - 3^{100}}{2}$

$\Rightarrow \quad A^{100} = \frac{3^{100} - 1}{2} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \frac{3 - 3^{100}}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$d(A) = O_{n \times n}$

$A^n + d_1 A^{n-1} + \dots + d_n I = 0$

$m(A) = O_{n \times n}$ → sometimes yes!

FABER-CASTELL

EXAMPLE:

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad d(s) = (s-2)^2(s-1)$$

$$d(A) = (A-2I)^2(A-I)$$
$$= (A-2I)(A-2I)(A-I)$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$m(s) = (s-2)(s-1)$$
$$m(A) = (A-2I)(A-I)$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = [0]$$

$$m(A) = 0_{n \times n}$$

called minimal polynomial

MINIMAL POLYNOMIAL

DEFINITION: A monic polynomial is a polynomial whose highest degree coefficient is unity, $=1$

Ex

$s^n + a_1 s^{n-1} + \dots \rightarrow$ monic

$2s^n + a_1 s^{n-1} + \dots \rightarrow$ not monic

DEFINITION: For an $n \times n$ matrix A, the minimal polynomial $m(s)$ is the monic polynomial with the smallest degree such that

$$m(A) = O_{n \times n}$$

Theorem: Given $A \in \mathbb{C}^{n \times n}$, let $m(s)$ be its minimal polynomial

i) $m(s)$ is unique

ii) $m(s)$ divides $d(s)$ without any remainder

$$\exists\, q(s) \quad \text{s.t} \quad d(s) = m(s)\, q(s)$$

iii) Every root of $d(s)$ is also a root of $m(s)$

$d(s) = (s-2)^2 (s-1)$

$\cancel{m(s) = (s-2)^2} \longrightarrow$ Prop iii) is violated

$m(s) = (s-2)(s-1)$ ✓

$m(s) = (s-2)^2 (s-1)$ ✓

$m(s) = (s-2)(s-1)(s-3) \rightarrow$ prop ii) is not satisfied

$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $d(s) = (s-2)^2 (s-1)$

$m(s) = (s-2)(s-1)$ ✓

$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $d(s) = (s-2)^2 (s-1)$

$\cancel{m(s) \overset{?}{=} (s-2)(s-1)}$ $\qquad m(s) = (s-2)^2 (s-1)$ ✓

$\downarrow$

Check

$(A-2I)(A-I)$

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

Proof:

i) Suppose $m(s)$ is not unique!

$m_1(s)$ monic

$m_2(s)$ monic

$$m_1(A) = 0_{n \times n}$$

$$m_2(A) = 0_{n \times n}$$

$$\Rightarrow \deg m_1 = \deg m_2 = k$$

$$m_1(s) = s^k + a_1 s^{k-1} + \dots + a_k$$

$$m_2(s) = s^k + b_1 s^{k-1} + \dots + b_k$$

$a_i \neq b_i$ for some i

$$m_1(s) - m_2(s) = (a_1 - b_1) s^{k-1} + (a_2 - b_2) s^{k-2} + \dots + (a_k - b_k)$$

Suppose $a_1 - b_1 \neq 0$

$$\underbrace{\frac{m_1(s) - m_2(s)}{a_1 - b_1}} = s^{k-1} + \frac{a_2 - b_2}{a_1 - b_1} s^{k-2} + \dots + \frac{a_k - b_k}{a_1 - b_1}$$

$$\frac{m_1(A) - m_2(A)}{a_1 - b_1} = 0_{n \times n}$$

$$\underbrace{\deg \frac{m_1(s) - m_2(s)}{a_1 - b_1}}_{\text{monic}} < k$$

$k-1$

contradiction to the starting assumption that m_1 & m_2 are minimal.

FABER-CASTELL

ii) $m(s)$ divide $d(s)$ without a remainder

$$d(s) = m(s)\, q(s)$$

$$\begin{array}{r|l} s^3+2s^2+3s+4 & s^2+2s+1 \\ \cline{2-2} -s^3+2s^2+s & s \\ \hline 2s+4 & \end{array}$$

$$(s^3+2s^2+3s+4) = (s^2+2s+1)\, s + (2s+4)$$

$$d(s) = m(s)\, q(s) + \underset{\sim}{r(s)}$$

degree $(r(s)) <$ degree $m(s)$

Assume $r(s) \neq 0$

$$d(s) = m(s)\, q(s) + r(s)$$

$$\underset{\sim}{d(A)} = m(A)\, q(A) + r(A)$$

$$0_{n\times n} = 0_{n\times n} + r(A)$$

$$r(A) = 0_{n\times n}$$

$$\deg(r(s)) < \deg(m(s))$$

$m(s)$ is not minimal $\Rightarrow$ contradiction

15.12.2015
Tuesday

$A \in \mathbb{C}^{n\times n}$ $\quad \lambda_1, \lambda_2, \ldots, \lambda_n$

$\lambda_i \neq \lambda_j$ when $i \neq j$

$$\mathbb{C}^n = \underbrace{N(A-\lambda_1 I)}_{\dim=1} \oplus \underbrace{N(A-\lambda_2 I)}_{\dim=1} \oplus \ldots \oplus \underbrace{N(A-\lambda_n I)}_{\dim=1}$$

$$\left(\begin{array}{l} x \in N(A-\lambda_1 I) \\ (A-\lambda_1 I)x = 0 \\ \quad \rightarrow Ax = \lambda_1 x \\ \\ Ax = \lambda_1 x \in N(A-\lambda_1 I) \end{array} \right)$$

$= \text{span}\{e_1\} \oplus \text{span}\{e_2\} \oplus \ldots \oplus \text{span}\{e_n\}$

n linearly ind. eigenvector

$\bar{A} = E^{-1} A E \qquad E = [e_1 \,\vdots\, e_2 \,\vdots\, \ldots \,\vdots\, e_n]$

$$\bar{A} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \text{ is diagonal}$$

A is diagonalizable

How can we generalize this result to matrices with repeated eigen-values?

$$d(s) = \det(sI - A) = s^n + a_1 s^{n-1} + \ldots + a_n = 0$$
$$= (s-\lambda_1)^{r_1}(s-\lambda_2)^{r_2} \ldots (s-\lambda_\sigma)^{r_\sigma}$$

$$r_1 + r_2 + \ldots + r_\sigma = n$$

[In the distinct eigen value case: $d(s) = (s-\lambda_1)^1 (s-\lambda_2)^1 \ldots (s-\lambda_n)^1$]

Characteristic Polynomial $= d(s)$

$d(s) = 0$: char. eqn.

$d(A) = 0_{n\times n}$ (Cayley Hamilton Thm)

FABER-CASTELL

Minimal Polynomial : Monic polynomial with the least degree s.t $m(s)$

$$m(A) = 0_{n\times n}$$

Properties:

1) $m(s)$ is unique

2) $m(s)$ divides $d(s)$ without any reminder

$$\exists\, q(s) \quad s.t \quad d(s) = m(s)\, q(s)$$

3) Every root of $d(s)$ is a root of $m(s)$.

$$d(s) = (s-2)^4 (s-4)^2$$

$\cancel{m(s) = (s-3)}$ (property #2)

$\cancel{m(s) = (s-2)^5 (s-4)}$ (Property #2)

$\cancel{m(s) = (s-4)^2}$ (Property #3)

$$m(s) = s^l + c_1 s^{l-1} + \dots + c_l$$

$$m(A) = A^l + c_1 A^{l-1} + \dots + c_l I = 0_{n\times n}$$

$$\lambda_i \quad i = 1, \dots, \sigma$$

Consider the eigen vector e_i

$$m(A)\, e_i = 0_{n\times 1}$$

$[\quad][\,]$ Vector

$$A^l e_i + c_1 A^{l-1} e_i + \dots + c_l e_i = 0_{n\times 1}$$

$$\left.\begin{aligned} A e_i &= \lambda_i e_i \\ A^2 e_i &= A A e_i = A \lambda_i e_i = \lambda_i A e_i = \lambda_i^2 e_i \\ A^3 e_i &= \lambda_i^3 e_i \end{aligned}\right\} A^k e_i = \lambda_i^k e_i$$

$$\lambda_i^{l} e_i + c_1 \lambda_i^{l-1} e_i + \cdots + c_l e_i = 0_{n\times 1}$$

$$\left(\lambda_i^{l} + c_1 \lambda_i^{l-1} + \cdots + c_l\right) \underbrace{e_i}_{\substack{\neq 0_{n\times 1} \\ \text{by} \\ \text{definition}}} = 0_{n\times 1}$$

$$\Rightarrow \quad \lambda_i^{l} + c_1 \lambda_i^{l-1} + \cdots + c_l = 0$$

$$\Rightarrow \quad m(\lambda_i) = 0$$

λ_i is a root of $m(s)$

EXAMPLE:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad \left.\begin{aligned} d(s) &= (s-1)(s-2)(s-3) \\ m(s) &= (s-1)(s-2)(s-3) \end{aligned}\right\}$$

When A has distinct eigen values we have
$d(s) = m(s) = (s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)$

$$\mathbb{R}^3 = N(A-\lambda_1 I) \oplus N(A-\lambda_2 I) \oplus N(A-\lambda_3 I)$$

$$d(s) = (s-\lambda_1)^{r_1}(s-\lambda_2)^{r_2} \cdots (s-\lambda_\sigma)^{r_\sigma}$$

$$m(s) = (s-\lambda_1)^{m_1}(s-\lambda_2)^{m_2} \cdots (s-\lambda_\sigma)^{m_\sigma}$$

$$1 \leq m_i \leq r_i$$

$$r_1 + r_2 + \cdots + r_\sigma = n$$

$$A - I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad N(A-I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{e_1} \right\}$$

$$A - 2I = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad N(A-2I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{e_2} \right\}$$

$$A - 3I = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad N(A-3I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{e_3} \right\}$$

3 l.i eigen vectors

What happens when there is repeated eigen values?

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

$d_1(s) = (s-1)(s-2)^2$ $\qquad$ $d_2(s) = (s-1)(s-2)^2$

$m_1(s) = (s-1)(s-2)$ ✓ $\qquad$ $m_2(s) = (s-1)(s-2)$ X → not satisfy

$(s-1)(s-2)^2$ X $\qquad$ $(s-1)(s-2)^2$ ✓

satisfy eqn. but it has not lower power

$$(A_1 - I)(A_1 - 2I)$$

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

$$(A_2 - I)(A_2 - 2I)$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \neq 0_{3\times 3}$$

$$A_1 - I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad N(A_1 - I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{e_1} \right\}$$

$$A_1 - 2I = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad N(A_1 - 2I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{e_2}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{e_3} \right\}$$

$$A_2 - I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad N(A_2 - I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{e_1} \right\}$$

$$A_2 - 2I = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad N(A_2 - 2I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{e_2} \right\}$$

$$\mathbb{R}^3 = \underbrace{N(A_1 - I)}_{\dim = 1} \oplus \underbrace{N(A_1 - 2I)}_{\dim = 2}$$

$$\cancel{\mathbb{R}^3 = \underbrace{N(A_2 - I)}_{1} \oplus \underbrace{N(A_2 - 2I)}_{1}}$$

$$\mathbb{R}^3 = N(A-\lambda_1 I) \oplus N(A-\lambda_2 I) \oplus N(A-\lambda_3 I)$$

$$(A_2 - 2I)^2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$N((A_2-2I)^2) = \text{span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{e_2}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\mathbb{R}^3 = \underbrace{N(A_2 - I)}_{\dim = 1} \oplus \underbrace{N((A_2-2I)^2)}_{\dim = 2}$$

$$\mathbb{C}^n = \underbrace{N((A-\lambda_1 I)^{m_1})}_{\dim\ r_1} \oplus \underbrace{N((A-\lambda_2 I)^{m_2})}_{\dim\ r_2} \oplus \dots \oplus \underbrace{N((A-\lambda_\sigma I)^{m_\sigma})}_{\dim\ r_\sigma}$$

18.12.2015
Friday

$A \in \mathbb{C}^{n \times n}$

$\lambda_1, \lambda_2, \dots, \lambda_n \qquad \lambda_i \neq \lambda_j$ when $i \neq j$

$d(s) = (s-\lambda_1)(s-\lambda_2) \dots (s-\lambda_n)$

$m(s) = (s-\lambda_1)(s-\lambda_2) \dots (s-\lambda_n)$

$$\mathbb{C}^n = \underbrace{\underbrace{N(A-\lambda_1 I)}_{\text{span}\{e_1\}}}_{\dim = 1} \oplus \underbrace{\underbrace{N(A-\lambda_2 I)}_{\text{span}\{e_2\}}}_{\dim = 2} \oplus \dots \oplus \underbrace{\underbrace{N(A-\lambda_n I)}_{\text{span}\{e_n\}}}_{\dim = n}$$

$\bar{A} = E^{-1} A E$

$E = [e_1 \ e_2 \ \dots \ e_n] \qquad \bar{A} = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{bmatrix}$

FABER-CASTELL

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

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

$d_1(s) = (s-1)(s-2)^2$

$m_1(s) = (s-1)(s-2)$

$d_2(s) = (s-1)(s-2)^2$

$m_2(s) = (s-1)(s-2)^2$

$N(A_1 - I) = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{e_1} \right\}$

$N(A_1 - 2I) = \text{Span}\left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{e_2}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{e_3} \right\}$

$\mathbb{R}^3 = N(A_1 - I) \oplus N(A_1 - 2I)$

$N(A_2 - I) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

$N(A_2 - 2I) = N\left(\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right) = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

$\cancel{\mathbb{R}^3 = N(A_2 - I) \oplus N(A_2 - 2I)}$

$N((A_2 - 2I)^2) = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

eigen vector because it's inside $N(A_2 - 2I)$

NOT eigen vector but it is called generalized eigen vector

$\mathbb{R}^3 = N(A_2 - I) \oplus N((A_2 - 2I)^2)$

* There is some relationship between decomposition form & $m(s)$ in terms of power.

$m_2(s) = (s-1)^{1}(s-2)^{2}$

$\mathbb{R}^3 = N((A_2 - I)^{1}) \oplus N((A_2 - 2I)^{2})$

FABER-CASTELL

THEOREM: $\mathbb{C}^n = N\left((A-\lambda_1 I)^{m_1}\right) \oplus N\left((A-\lambda_2 I)^{m_2}\right) \oplus \ldots \oplus N\left((A-\lambda_\sigma I)^{m_\sigma}\right)$

$$d(s) = (s-\lambda_1)^{r_1}(s-\lambda_2)^{r_2} \ldots (s-\lambda_\sigma)^{r_\sigma}$$

$$r_1 + r_2 + \ldots + r_\sigma = n$$

$$m(s) = (s-\lambda_1)^{m_1}(s-\lambda_2)^{m_2} \ldots (s-\lambda_\sigma)^{m_\sigma}$$

$$1 \leq m_i \leq r_i$$

$$\bar{A} = \begin{bmatrix} \bar{A}_1 & & & 0 \\ & \bar{A}_2 & & \\ & 0 & \ddots & \\ & & & \bar{A}_\sigma \end{bmatrix}$$

$\bar{A} = B^{-1}AB$ where B is composed of basis vectors for $N(A-\lambda_i I)$

$$B = \begin{bmatrix} B_1 & B_2 & \ldots & B_\sigma \end{bmatrix}$$

B_i is composed of basis vectors for $N(A-\lambda_i I)$

Size of $\bar{A}_i$ = $\dim N\left((A-\lambda_i I)^{m_i}\right)$

We now concentrate on the structure of $N\left((A-\lambda_i I)^{m_i}\right)$

$$\boxed{\Sigma_i \triangleq A - \lambda_i I}$$

Theorem:

$$\underset{\dim}{N(\Sigma_i)} \underset{<}{\subsetneq} \underset{\dim}{N(\Sigma_i^2)} \underset{<}{\subsetneq} \underset{\dim}{N(\Sigma_i^3)} \underset{<}{\subsetneq} \ldots \underset{<}{\subsetneq} \underset{\dim}{N(\Sigma_i^{k_i})} \underset{=}{=} \underset{\dim}{N(\Sigma^{k_i+1})} \underset{=}{=} \underset{\dim}{N(\Sigma^{k_i+2})} = \ldots$$

$N(\Sigma_i) \to N(A-\lambda_i I)$

$\subsetneq \to$ proper subset

$$A \subsetneq B \Rightarrow A \subset B \;\&\; A \neq B$$

Proof: We first show that $N(\Sigma_i^l) \subset N(\Sigma_i^{l+1})$

Let x be arbitrary and

$$x \in N(\Sigma_i^l)$$

$$\Rightarrow \Sigma_i^l x = 0$$

$$\Sigma_i \Sigma_i^l x = \Sigma_i 0 = 0$$

$$\Sigma_i^{l+1} x = 0$$

$$\Rightarrow x \in N(\Sigma_i^{l+1})$$

We show that when

$$N(\Sigma_i^l) = N(\Sigma_i^{l+1})$$

then

$$N(\Sigma_i^{l+1}) = N(\Sigma_i^{l+2})$$

- $N(\Sigma_i^{l+1}) \subset N(\Sigma_i^{l+2})$ → this is already proven.
- $N(\Sigma_i^{l+2}) \subset N(\Sigma_i^{l+1})$

$$x \in N(\Sigma_i^{l+2})$$

$$\Sigma_i^{l+2} x = 0$$

$$\Sigma_i^{l+1} \Sigma_i x = 0$$

$$\Sigma_i x \in N(\Sigma_i^{l+1})$$

$$\Sigma_i x \in N(\Sigma_i^l)$$

$$\Sigma_i^l \Sigma_i x = 0 \Rightarrow \Sigma_i^{l+1} x = 0 \Rightarrow x \in N(\Sigma_i^{l+1})$$

Suppose we call the smallest power l where $N(\Sigma_i^l) = N(\Sigma_i^{l+1})$ to be k_i.

$$N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq N(\Sigma_i^3) \subsetneq \dots \subsetneq N(\Sigma_i^{k_i}) = N(\Sigma_i^{k_i+1}) = N(\Sigma_i^{k_i+2}) = \dots$$

EXAMPLE:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad \begin{array}{l} d(s) = (s-1)^3 (s-2) \\ m(s) = ? \;\; = (s-1)^{?} (s-2) \end{array}$$

$$\underbrace{(A - I)}_{\Sigma_1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \dim N(A - I) = 2$$

increase

$$\Sigma_1^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \dim N(\Sigma_1^2) = 3$$

$$\Sigma_1^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \dim N(\Sigma_1^3) = 3 = r_1$$

$$\dim N(\Sigma_1^4) = 3$$

$$\Rightarrow k_1 = 2 = m_1$$

$$\Sigma_2 = A - \lambda_2 I = A - 2I = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \dim N(\Sigma_2^1) = 1 = r_2$$

$$\Sigma_2^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \dim N(\Sigma_2^2) = 1$$

$$\Rightarrow k_2 = 1 = m_2$$

$$\cancel{m(s) = (s-1)(s-2)}$$

$$\cancel{(A-I)(A-2I) = 0}$$

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

$$m(s) = (s-1)^2(s-2)$$

$$\Sigma_1^2 \Sigma_2 = \begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \end{bmatrix} \begin{bmatrix} -1&0&0&0 \\ 0&-1&1&0 \\ 0&0&-1&0 \\ 0&0&0&0 \end{bmatrix} = \begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix}$$

Theorem: $d(s) = (s-\lambda_1)^{r_1}(s-\lambda_2)^{r_2} \dots (s-\lambda_\sigma)^{r_\sigma}$

$m(s) = (s-\lambda_1)^{m_1}(s-\lambda_2)^{m_2} \dots (s-\lambda_\sigma)^{m_\sigma}$

$$1 \leq m_i \leq r_i$$

- $k_i = m_i$
- $\dim\left(\left(A-\lambda_i I\right)^{m_i}\right) = r_i$

$$\dim N\left(\Sigma_i^{m_i-1}\right) < r_i$$

$$N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq \dots \subsetneq \underbrace{N(\Sigma_i^{m_i})}_{\dim = r_i} = \underbrace{N(\Sigma_i^{m_i+1})}_{\dim = r_i} = \dots$$

EXAMPLE:

$$A = \begin{bmatrix} 1&1&0&0 \\ 0&1&1&0 \\ 0&0&1&0 \\ 0&0&0&2 \end{bmatrix} \qquad d(s) = (s-1)^3(s-2)$$

$$m(s) = (s-1)^?(s-2)$$

$$\Sigma_1 \triangleq A - I = \begin{bmatrix} 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \\ 0&0&0&1 \end{bmatrix} \Rightarrow \dim N(\Sigma_1) = 1 \rightarrow \text{not } 3$$

continue to take power of Σ_1

$$\Sigma_1^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \dim\left(N(\Sigma_1^2)\right) = 2$$

$$\Sigma_1^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \dim\left(N(\Sigma_1^3)\right) = 3 = r_1 \Rightarrow m_1 = 3$$

$$\Rightarrow m(s) = (s-1)^3 (s-2)$$

EXAMPLE:

$$A = \begin{bmatrix} 2 & 1 & 0 & & & & & & & \\ 0 & 2 & 1 & & & & & & & \\ 0 & 0 & 2 & & & & & & & \\ & & & 2 & 1 & & & & & \\ & & & 0 & 2 & & & & & \\ & & & & & 2 & & & & \\ & & & & & & 2 & & & \\ & & & & & & & 3 & 1 & \\ & & & & & & & 0 & 3 & \\ & & & & & & & & & 3 \end{bmatrix}$$

(off-diagonal blocks: 0)

A matrix in Jordan Canonical Form

Consider a single Jordan block:

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \qquad d_1(s) = (s-2)^3$$

$$A_1 - 2I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \dim N(A_1 - 2I) = 1$$

$+1$

$$(A_1 - 2I)^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \dim N\left((A_1 - 2I)^2\right) = 2$$

$+1$

$$(A_1 - 2I)^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \dim N\left((A_1 - 2I)^3\right) = 3$$

FABER-CASTELL

$$\dim\left(N\left(A_1-2I\right)^{\ell}\right)=3$$

$$\ell \geq 3$$

Now, let's go back 10x10 matrix,

$$d(s)=(s-2)^7(s-3)^3$$

$$m(s)=(s-2)^?(s-3)^?$$

#of zero column

$\Sigma_1 = A-2I \rightarrow \dim N(\Sigma_1)=4$ $\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$, $\begin{bmatrix}0&1\\0&0\end{bmatrix}$

$\dim N(\Sigma_1^2)=6$ $\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}$

$\dim N(\Sigma_1^3)\ 7$

$\dim N(\Sigma_1^4)=7$

⋮

# of zero column

$\Sigma_2 : A-3I \rightarrow \dim N(\Sigma_2)=2$ $\begin{bmatrix}0&1\\0&0\end{bmatrix}$

$\dim N(\Sigma_2^2)=3$

$\dim N(\Sigma_2^3)=3$

⋮

$$\Rightarrow m(s)=(s-2)^3(s-3)^2$$

$\dim N(\Sigma_i)$: # of Jordan blocks corresponding to eigenvalue λ_i.

= # of Jordan blocks with size ≥ 1

$\dim N(\Sigma_i^2) - \dim N(\Sigma_i)$: # of Jordan blocks with size ≥ 2 corresponding to λ_i.

$\dim N(\Sigma_i^3) - \dim N(\Sigma_i^2)$: # of Jordan blocks with size ≥ 3 corresponding to λ_i

$\dim N(\Sigma_i^k) - \dim N(\Sigma_i^{k-1})$ = # of " " " size $\geq k$ " " "

m_i : size of the largest Jordan block corresponding to λ_i.

$$m(s) = (s-2)^3 (s-3)^2$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \quad 3\times3 \qquad \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \quad 2\times2$$

$$\underbrace{\bar{A}}_{\text{in Jordan canonical form}} = B^{-1} A B \quad \Rightarrow \quad B = ?$$

What should be chosed our basis set such that the simple representation $\bar{A}$ is in J.C.F.

22.12.2015
Tuesday

$A \in \mathbb{C}^{n\times n}$

$$\mathbb{C}^n = N\big(\underbrace{(A-\lambda_1 I)^{m_1}}_{\dim = r_1}\big) \oplus \ldots\ldots \oplus N\big((A-\lambda_\sigma I)^{m_\sigma}\big)$$

$$d(s) = (s-\lambda_1)^{r_1} (s-\lambda_2)^{r_2} \ldots\ldots (s-\lambda_\sigma)^{r_\sigma}$$

$$m(s) = (s-\lambda_1)^{m_1} (s-\lambda_2)^{m_2} \ldots\ldots (s-\lambda_\sigma)^{m_\sigma}$$

$$r_1 + r_2 + \ldots + r_\sigma = n$$

$$1 \leq m_i \leq r_i$$

$$\Sigma_i = A - \lambda_i I$$

$$N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq \ldots \subsetneq \overbrace{N(\Sigma_i^{m_i})}^{\dim r_i} = N(\Sigma_i^{m_i+1}) = \ldots$$

A matrix in JCF

dim $N(A-\lambda_i I)$: # of Jordan block for λ_i with size ≥ 1

dim $N((A-\lambda_i I)^2)$ − dim $N(A-\lambda_i I)$: # of Jordan blocks for λ_i with size ≥ 2

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{(\cdot)^3} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{(\cdot)^2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

dim $N((A-\lambda_i I)^k)$ − dim $N((A-\lambda_i I)^{k-1})$: # of Jordan blocks for λ_i with size $\geq k$

m_i : size of the largest Jordan block for λ_i.

Finding JCF for a matrix A from Subspace Dimensions

$$\bar{A} = B^{-1} A B$$

rank A = rank BA where B is an invertable matrix

rank A = rank AB where " " " " "

$(\bar{A}-\lambda_i I)$ $\quad B^{-1}(A-\lambda_i I)B = B^{-1}AB - \lambda_i B^{-1}B = \bar{A} - \lambda_i I$

dim $N(A-\lambda_i I)$ = dim $N(\bar{A}-\lambda_i I)$

A is not in JCF but we will still find dim of null space of $\bar{A}$ with this.

EXAMPLE: Suppose a matrix $A \in \mathbb{R}^{8\times 8}$ has the following subspace dimensions.

$$\dim N(A-3I) = 5$$
$$\dim N\left((A-3I)^2\right) = 7 \quad 2 \quad \geq 2$$
$$\dim N\left((A-3I)^3\right) = 8 \qquad \mathbb{R}^8 = \left((A-3I)^3\right)$$

a) Find $d(s)$

$$d(s) = (s-3)^8$$

b) Find $m(s)$

$$m(s) = (s-3)^3$$

c) What is JCF for A? $\left(\bar{A} = ?\right)$

$\dim N(A-3I) =$ # of Jordan blocks for $\lambda = 3$ with size ≥ 1

$\Rightarrow$ # of Jordan blocks $= 5$

$\dim N\left((A-3I)^2\right) - \dim N(A-3I) = 2 =$ # of Jordan blocks with size ≥ 2

$\dim N\left((A-3I)^3\right) - \dim N\left((A-3I)^2\right) = 1 =$ # of Jordan blocks with size ≥ 3

Further substractions will give us zero. $\Rightarrow$ # of JB with size ≥ 4 $= 0$

# of Jordan blocks with size $= 3$ $= 1$ $\qquad \geq 1 \rightarrow 5$

$\geq 2 \rightarrow 2$

$\geq 3 \rightarrow 1$

# of Jordan blocks with size $= 2$ $= 1$

# of Jordan blocks with size $= 1$ $= 3$

$$\bar{A} = \left[\begin{array}{cccccccc} 3 & 1 & & & & & & \\ & 3 & 1 & & & & 0 & \\ & & 3 & & & & & \\ & & & 3 & 1 & & & \\ & & & & 3 & & & \\ & & & & & 3 & & \\ & 0 & & & & & 3 & \\ & & & & & & & 3 \end{array}\right]$$

EXAMPLE: Let $A \in \mathbb{R}^{4\times 4}$

Single eigenvalue $\lambda_1 = 7$

$m_1 = 2$

Find all possible JCFs.

$d(s) = (s-7)^4$

$m(s) = (s-7)^2$

$$\underbrace{N(A-7I)}_{\substack{dim = 3 \\ = 2 \\ = 1}} \subsetneq \underbrace{N((A-7I)^2)}_{dim = 4}$$

dim $N(A-7I) = 3$

\# of blocks $= 3$

\# of blocks size ≥ 2 $= 1$

\# of blocks size ≥ 3 $= 0$

\# of blocks with size $= 1$ $= 2$

\# of blocks with size $= 2$ $= 1$

$$\bar{A} = \left[\begin{array}{cc|c|c} 7 & 1 & & 0 \\ & 7 & & \\ \hline & & 7 & 0 \\ \cline{3-4} 0 & & 0 & 7 \end{array}\right]$$

dim $N(A-7I) = 2$

\# of blocks $= 2$

\# of blocks size ≥ 2 $= 2$

$$\bar{A} = \left[\begin{array}{cc|cc} 7 & 1 & & 0 \\ 0 & 7 & & \\ \hline & & 7 & 1 \\ 0 & & 0 & 7 \end{array}\right]$$

dim $N(A-7I) = 1$

\# of blocks $= 1$

\# of blocks size ≥ 2 $= 3$

There cannot be such an A matrix. This case is not possible!

2. Midterm up to!

How to Find a Basis set to Convert a Matrix A into JCF

$A \in \mathbb{C}^{n \times n}$

$J = B^{-1} A B$

$B = ?$

$$\mathbb{C}^n = N\big(\underbrace{(A-\lambda_1 I)^{m_1}}_{r_1}\big) \oplus \dots \oplus N\big(\underbrace{(A-\lambda_\sigma I)^{m_\sigma}}_{r_\sigma}\big)$$

$$B = \left[\, B_1 \,\vdots\, B_2 \,\vdots\, \cdots \,\vdots\, B_\sigma \,\right]$$

$$B_i = \Big[\, \underbrace{b_i^1 \,\vdots\, b_i^2 \,\vdots\, \cdots \,\vdots\, b_i^{r_i}}_{\text{A basis set for } N\left((A-\lambda_i I)^{m_i}\right)} \,\Big]$$

$b_i^j = ?$

$$J = \begin{bmatrix} \bar{A}_1 & & & \\ & \bar{A}_2 & & \\ & & \ddots & \\ & & & \bar{A}_\sigma \end{bmatrix}$$

Since $\bar{A}_i$ are going to be $r_i \times r_i$

$$\bar{A}_i = \begin{bmatrix} \lambda_i & * & & \\ & \lambda_i & * & 0 \\ & & \ddots & * \\ & 0 & & \lambda_i \end{bmatrix}_{r_i \times r_i}$$ $*$ can be 1 or 0.

$$\left[\begin{array}{cc|cc} 7 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ \hline 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 7 \end{array}\right] \rightarrow \text{wrt } 1 \& 0, \text{ we can determine size of JB.}$$

24.12.2015
Thursday

$A \in \mathbb{C}^{n \times n}$

$\bar{A} = J = B^{-1}AB$

$\mathbb{C}^n = N(\underbrace{(A-\lambda_1 I)^{m_1}}) \oplus \ldots \oplus N((A-\lambda_\sigma I)^{m_\sigma})$

we look for basis vectors for $\underbrace{N((A-\lambda_i I)^{m_i})}_{r_i}$

$B = [B_1 \mid B_2 \mid \ldots \mid B_\sigma]$

$B_i = [b_i^1 \mid b_i^2 \mid \ldots \mid b_i^{r_i}]$ $\quad b_i^1, \ldots, b_i^{r_i}$ have to be basis vectors for $N((A-\lambda_i I)^{m_i})$

$$\bar{A} = \begin{bmatrix} \bar{A}_1 & & 0 \\ & \bar{A}_2 & \\ 0 & & \ddots & \\ & & & \bar{A}_\sigma \end{bmatrix}$$

$\bar{A}_i \in \mathbb{C}^{r_i \times r_i}$ & $\bar{A}_i = \begin{bmatrix} \lambda_i & * & & 0 \\ & \lambda_i & * & \\ 0 & & \ddots & * \\ & & & \lambda_i \end{bmatrix}$

$* \in \{0,1\}$

$\Sigma_i \triangleq A - \lambda_i I$

$N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq \ldots \subsetneq \underbrace{N(\Sigma_i^{m_i-1})}_{x \notin N(\Sigma_i^{m_i-1})} \subsetneq \underbrace{N(\Sigma_i^{m_i})}_{x \in N(\Sigma_i^{m_i})} = \ldots$

We can always find a vector x st $x \in N(\Sigma_i^{m_i})$ & $x \notin N(\Sigma_i^{m_i-1})$

Consider the chain of vectors : $\{\Sigma_i^{m_i-1}x, \Sigma_i^{m_i-2}x, \ldots, \Sigma_i x, x\}$

Claim: The set $\{\Sigma_i^{m_i-1}x, \dots, x\}$ is l.i.

Proof:

$$\cancel{\Sigma_i^{m_i-1}} \quad \alpha_1 \Sigma_i^{m_i-1} x + \alpha_2 \Sigma_i^{m_i-2} + \dots + \alpha_{m_i-1} \Sigma_i x + \alpha_{m_i} x = 0$$

$$\alpha_1 \underbrace{\cancel{\Sigma_i^{2m_i-2}}}_{0} + \alpha_2 \underbrace{\cancel{\Sigma_i^{2m_i-3}}}_{0} + \dots \underbrace{\cancel{\ }}_{0} + \alpha_{m_i-1} \underbrace{\cancel{\Sigma_i^{m_i}}}_{0} x + \alpha_{m_i} \Sigma_i^{m_i-1} x = 0$$

$$\alpha_{m_i} \underbrace{\Sigma_i^{m_i-1} x}_{\text{cannot be zero}} = 0 \quad \Rightarrow \quad \alpha_{m_i} = 0$$

$$\cancel{\Sigma_i^{m_i-2}} \quad \alpha_1 \Sigma_i^{m_i-1} x + \alpha_2 \Sigma_i^{m_i-2} + \dots + \alpha_{m_i-1} \Sigma_i x = 0$$

$$\alpha_1 \underbrace{\cancel{\Sigma_i^{2m_i-3}}}_{0} + \alpha_2 \underbrace{\cancel{\Sigma_i^{2m_i-4}}}_{0} \underbrace{\cancel{+}}_{0} \dots \underbrace{\cancel{\ }}_{0} + \alpha_{m_i-1} \Sigma_i^{m_i-1} x = 0$$

$$\Rightarrow \alpha_{m_i-1} = 0$$

We can continue in this manner to show that $\alpha_1, \dots, \alpha_{m_i}$ are all equal to zero

$$\Rightarrow \{\Sigma_i^{m_i-1} x, \dots, \Sigma_i x, x\} \text{ is a l.i. set}$$

$$(\Sigma_i^{m_i-1} x)$$

$$x \in N(\Sigma_i^{m_i}) \qquad x \notin N(\Sigma_i^{m_i-1})$$

$$\Sigma_i(\Sigma_i^{m_i-1} x) = \Sigma_i^{m_i} x = 0$$

$$\Sigma_i^{m_i-1} x \in N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq \dots$$

$N(A - \lambda_i I)$

$\Sigma_i^{m_i - 1} x$ is an eigenvector corresponding to λ_i.

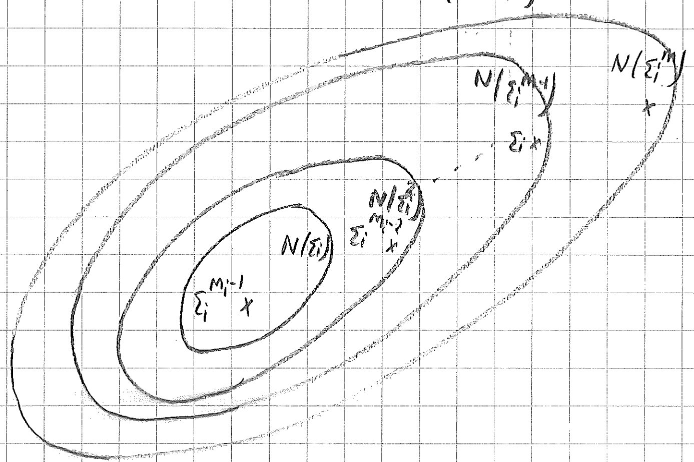
$$\Sigma_i \Sigma_i^{m_i - 2} x = \Sigma_i^{m_i - 1} x \neq 0$$

$$x \notin N(\Sigma_i)$$

$$\Sigma_i^2 \Sigma_i^{m_i - 2} x = \Sigma_i^{m_i} x = 0$$

$$\Rightarrow \Sigma_i^{m_i - 2} x \in N(\Sigma_i^2)$$

$$\Sigma_i^{m_i - k} x \in N(\Sigma_i^k)$$



If $r_i = m_i$ set is sufficient but if it is NOT, we have to add new vectors to set to make $r_i = m_i$.

EXAMPLE:

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

$m(s)$

$\bar{A} = ? = J$

$B = ?$ st $B^{-1} A B = J$

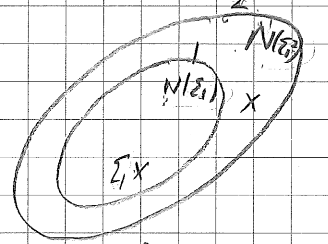
FABER-CASTELL

$$(sI-A) = \begin{bmatrix} s-1 & -1 & 1 \\ 1 & s-3 & -2 \\ 0 & 0 & s-1 \end{bmatrix}$$

$$d(s) = \det(sI-A) = \underset{\lambda_1}{(s-2)^2}\ \underset{\lambda_2}{(s-1)}$$

$$\Sigma_1 = A-2I = \begin{bmatrix} -1 & 1 & -1 \\ -1 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix} \qquad \dim N(A-2I) = 1$$

$$m(s) = \cancel{(s-2)(s-1)}$$



$$\Sigma_1^2 = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \dim N((A-2I)^2) = 2$$

$$\Rightarrow m(s) = (s-2)^2(s-1) = d(s)$$

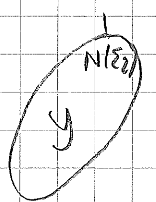
$$x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \underset{\text{eigenvector}}{\Sigma_1 x} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

$$J = \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$$

$$\cancel{x = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}$$ it is inside of $N(\Sigma_1)$ also so we cannot choose x like that

Chain for $\lambda_1 = \{\Sigma_1 x, x\}$

$$\Sigma_2 = A-I = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \qquad \dim (A-I) = 1$$



$$y = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$

eigenvectors

Chain for $\lambda_2 = \{y\}$

$$B = \left[\, \Sigma_i x \;\middle|\; x \;\middle|\; y \,\right]$$

$$J = B^{-1} A B \Rightarrow BJ = AB$$

$$\underbrace{\begin{bmatrix} -1 & 1 & 4 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{B} \underbrace{\left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]}_{J} = \underbrace{\begin{bmatrix} 1 & 1 & -1 \\ -1 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 4 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{\begin{bmatrix} -2 & 1 & 4 \\ -2 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}$$

$$\underbrace{\begin{bmatrix} 1 & -1 & 4 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{B_1} \underbrace{\left[\begin{array}{cc|c} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{array}\right]}_{J_1} = \underbrace{\begin{bmatrix} 1 & 1 & -1 \\ -1 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 4 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{\begin{bmatrix} 1 & -2 & 4 \\ -1 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix}}$$

$$B_2 = \begin{bmatrix} 4 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \rightarrow J_2 = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 1 \\ 0 & 0 & 2 \end{array}\right] \checkmark$$

Special Cases

<u>Case-I:</u> Let A have a single eigenvalue λ_i and we have $m_i = r_i$

$$A \in \mathbb{C}^{r_i \times r_i}$$

$$d(s) = (s-\lambda_i)^{r_i} \qquad m(s) = (s-\lambda_i)^{m_i = r_i}$$

$$N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq \dots \subsetneq N(\Sigma_i^{m_i-1}) \subsetneq \underbrace{N(\Sigma_i^{m_i})}_{r_i = m_i}$$

$$1 < 2 < \dots < m_i - 1 < m_i$$

We have a single chain

$$\underbrace{\{ \Sigma_i^{m_i-1} x,\ \Sigma_i^{m_i-2} x,\ \text{-----},\ \Sigma_i x,\ x \}}_{m_i = r_i} \quad \text{A basis for } N(\Sigma_i^{m_i})$$

$$B = \left[\Sigma_i^{m_i-1} x \;\middle|\; \Sigma_i^{m_i-2} x \;\middle|\; \text{-------} \;\middle|\; \Sigma_i x \;\middle|\; x \right]$$

$$B \begin{bmatrix} \lambda_i & 1 & 0 & & & \\ 0 & \lambda_i & 1 & & 0 & \\ 0 & 0 & \lambda_i & 1 & & \\ \vdots & \vdots & 0 & \ddots & \ddots & \\ & & & & & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} = AB$$

for $\Sigma_i^{m_i-1} x$

For $\Sigma_i^{m_i-2} x \rightarrow \quad \Sigma_i \Sigma_i^{m_i-2} x = \Sigma_i^{m_i-1} x$

$$(A - \lambda_i I)\Sigma_i^{m_i-2} x = \Sigma_i^{m_i-1} x$$

$$A\Sigma_i^{m_i-2} x - \lambda_i \Sigma_i^{m_i-2} x = \Sigma_i^{m_i-1} x$$

$$A\Sigma_i^{m_i-2} x = \Sigma_i^{m_i-1} x + \lambda_i \Sigma_i^{m_i-2} x \Rightarrow \begin{bmatrix} 1 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

For $\Sigma_i^{m_i-3} x \rightarrow \quad \Sigma_i \Sigma_i^{m_i-3} x = \Sigma_i^{m_i-2} x$

$(A - \lambda_i I)$

$$\Rightarrow A\Sigma_i^{m_i-3} x = \Sigma_i^{m_i-2} x + \lambda_i \Sigma_i^{m_i-3} x \Rightarrow \begin{bmatrix} 0 \\ 1 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad J = \begin{bmatrix} \lambda_i & 1 & & & \\ 0 & \lambda_i & 1 & & \\ & 0 & \lambda_i & 1 & \\ & & 0 & \ddots & \ddots \\ & & & & 1 \\ & & & 0 & \lambda_i \end{bmatrix}$$

We have a single block with size $r_i = m_i$

$\Sigma_i^{m_i-1} x$ is the only eigenvector

The other vectors in the chain are called "generalized eigen vectors."

FABER-CASTELL

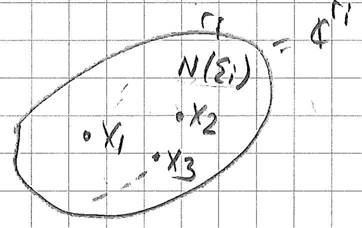
Case II: A has a single eigenvalue λ_i

$m_i = 1$

$d(s) = (s-\lambda_i)^{r_i}$

$m(s) = (s-\lambda_i)^1$

$$\underbrace{N(\Sigma_i)}_{\substack{\dim = r_i \\ \| \\ \mathbb{C}^{r_i}}} = \underbrace{N(\Sigma_i^2)}_{r_i}$$



Chain $= \underbrace{\{x_1\} \; \{x_2\}}_{\text{Choose } x_1 \,\&\, x_2 \text{ are l.i}}$ ...... $\{x_{r_i}\}$

Vectors $x_1, \ldots, x_i$ should be l.i

We have r_i chains with size 1.

r_i l.i. eigenvectors corresponding to λ_i.

$$B = [\, x_1 \mid x_2 \mid \cdots \mid x_{r_i} \,]$$

$$B \underbrace{\begin{bmatrix} \lambda_i & 0 & 0 & & 0 \\ 0 & \lambda_i & 0 & & \\ 0 & 0 & \lambda_i & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & & \lambda_i \end{bmatrix}}_{\text{We have diagonal J.C.F}} = A B$$

It is not surprize $m(s) = (s-\lambda_i) \Rightarrow A - \lambda_i I = 0 \Rightarrow A = \lambda_i I$

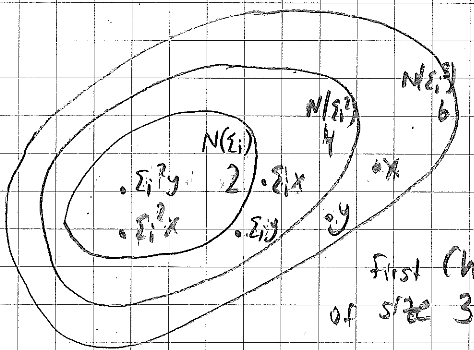
EXAMPLE: A is a 6×6 Matrix, λ_i is the only eigenvalue.

$d(s)=(s-\lambda_i)^6 \quad \dim N(\Sigma_i)=2$

$m(s)=(s-\lambda_i)^3 \quad \dim N(\Sigma_i^2)=4$

$B=?$

$$N(\Sigma_i) \subsetneq N(\Sigma_i^2) \subsetneq N(\Sigma_i^3) = N(\Sigma_i^4) = \dots$$
$$2 \qquad\quad 4 \qquad\quad 6 \qquad\quad 6$$



Find $x \in N(\Sigma_i^3)$

$x \notin N(\Sigma_i^2)$

first chain $=\{\Sigma_i^2 x, \Sigma_i x, x\}$ of size 3

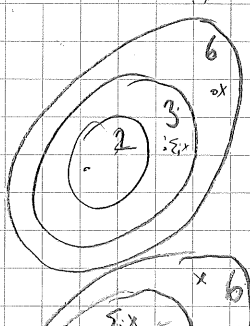
We can find y which is l.i. with x because dimension is increase from 4 to 6.

Find $y \in N(\Sigma_i^3)$

$y \notin N(\Sigma_i^2)$

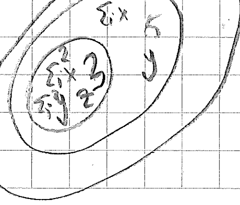
$\Sigma_i^2 x$ & $\Sigma_i^2 y$ should be l.i. (If we garantee this condition, we already have l.i. $\Sigma_i^2 x, \Sigma_i x, x$ & $\Sigma_i^2 y, \Sigma_i y, y$)

The vectors in the inner most space have to be l.i. $\{\Sigma_i^2 y, \Sigma_i y, y\}$ second chain with size 3.



# v	size ≥ 1	$=2$
# v	size ≥ 2	$=1$
# v	size ≥ 3	$=3$

Not possible must be decrease



$\{\Sigma_i^2 x, \Sigma_i x, x, \Sigma_i y, y, z\}$

size $\geq 1 = 3$

size $\geq 2 = 2$

size $\geq 3 = 1$

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

$$B=\left[\Sigma_i^2 x,\ \Sigma_i x,\ x,\ \Sigma_i^2 y,\ \Sigma_i y,\ y\right]$$

$$J=\left[\begin{array}{ccc|ccc} \lambda_i & 1 & 0 & & & \\ & \lambda_i & 1 & & 0 & \\ & & \lambda_i & & & \\ \hline & & & \lambda_i & 1 & 0 \\ & 0 & & & \lambda_i & 1 \\ & & & & & \lambda_i \end{array}\right]$$

! # of Jordan blocks will be equal to # of chains.

! Sizes of the Jordan blocks will be equal to the sizes of the chains.

<u>EXAMPLE:</u> 4x4 matrix

Single eigen value λ_i $m_i=2$

Find all possible chains and JCFs.

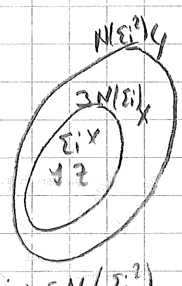
$d(s)=(s-\lambda_i)^4$

$m(s)=(s-\lambda_i)^2$

$N(\Sigma_i) \subsetneq N(\Sigma_i^2)$
$\quad 3,2,1 \qquad 4$

<u>Case I</u> : $\underset{3}{N(\Sigma_i)} \subsetneq \underset{4}{N(\Sigma_i^2)}$

$$\left[\begin{array}{cc|c|c} \lambda_i & 1 & & \\ 0 & \lambda_i & \multicolumn{2}{c}{0} \\ \hline & & \lambda_i & \\ \hline \multicolumn{2}{c|}{0} & & \lambda_i \end{array}\right]$$



Choose $x \in N(\Sigma_i^2)$
$x \notin N(\Sigma_i)$

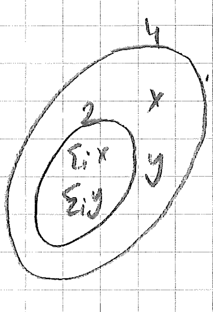
y & z & $\Sigma_i x$ should be l.i !

$\{\Sigma_i x, x\}$ $\{y\}$ $\{z\}$

$B=[\Sigma_i x, x, y, z]$

3 eigen vectors : $\Sigma_i x, y, z$!

<u>Case II</u> : $\underset{2}{N(\Sigma_i)} \subsetneq \underset{4}{N(\Sigma_i^2)}$



Choose $x \in N(\Sigma_i^2)$
$x \notin N(\Sigma_i)$

Choose $y \in N(\Sigma_i^2)$
$y \notin N(\Sigma_i)$

$\Sigma_i x$ & $\Sigma_i y$ should be l.i !

$$\left[\begin{array}{cc|cc} \lambda_i & 1 & & \\ 0 & \lambda_i & \multicolumn{2}{c}{0} \\ \hline & & \lambda_i & 1 \\ \multicolumn{2}{c|}{0} & 0 & \lambda_i \end{array}\right]$$

$\{\Sigma_i x, x\}$ $\{\Sigma_i y, y\}$

2 eigen vectors : $\Sigma_i x, \Sigma_i y$!

29.12.2015
Tuesday

$J = B^{-1}AB \qquad B = ?$

Ex

4x4 Matrix

Single eigenvalue λ_i

$m_i = 2$

Find all JCF's J & B

$$d(s) = (s-\lambda_i)^4$$
$$m(s) = (s-\lambda_i)^2$$

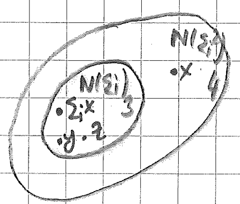
$\Sigma_i = A - \lambda_i I$

$$N(\Sigma_i) \subsetneq N(\Sigma_i^2)$$

$$\begin{array}{ccc} N(\Sigma_i) & \subsetneq & N(\Sigma_i^2) \\ 3 & & 4 \\ 2 & & \\ 1 & & \end{array}$$

Case I:

$$\begin{array}{ccc} N(\Sigma_i) & \subsetneq & N(\Sigma_i^2) \\ 3 & & 4 \end{array}$$



$$J = \left[\begin{array}{cc|c|c} \lambda_i & 1 & & \\ 0 & \lambda_i & \multicolumn{2}{c}{0} \\ \hline & & \lambda_i & \\ \hline & & & \lambda_i \end{array}\right]$$

$x \in N(\Sigma_i^2) \quad x \notin N(\Sigma_i)$

$\{\underbrace{\Sigma_i x}_{\text{eigenvector}}, \underbrace{x}_{\text{generalized eigenvector}}\}$

$\{y\} \quad y \in N(\Sigma_i) \quad y$ is $l.i$ with $\Sigma_i x$

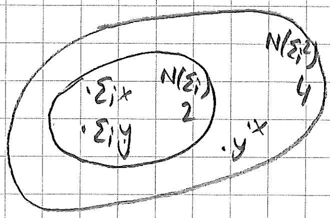
$\{z\} \quad z \in N(\Sigma_i) \quad z$ is $l.i$ with $\Sigma_i x$

$$B = [\, \Sigma_i x \mid x \mid y \mid z \,]$$

$$J = B^{-1} A B$$

Case II

$$\underset{2}{N(\Sigma_i)} \subsetneq \underset{4}{N(\Sigma_i^2)}$$



$$J = \left[\begin{array}{cc|cc} \lambda_i & 1 & & \\ 0 & \lambda_i & \multicolumn{2}{c}{0} \\ \hline \multicolumn{2}{c|}{0} & \lambda_i & 1 \\ & & 0 & \lambda_i \end{array}\right]$$

$x \in N(\Sigma_i^2) \quad x \notin N(\Sigma_i)$

$$\{\underbrace{\Sigma_i x}_{\text{eigen vector}}, x\}$$

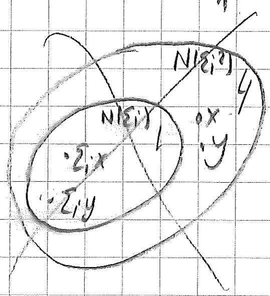
$y \in N(\Sigma_i^2) \quad y \notin N(\Sigma_i)$ & $\Sigma_i y$ & $\Sigma_i x$ have to be l.i.

$$B = [\, \Sigma_i x \mid x \mid \Sigma_i y \mid y \,]$$

$$\{\underbrace{\Sigma_i y}_{\text{eigen vector}}, y\}$$

Case III

$$\underset{1}{N(\Sigma_i)} \subsetneq \underset{4}{N(\Sigma_i^2)}$$



$\cancel{J = [\quad]}$ $\quad$ $A\ \cancel{[\quad]}$ $\to$ cannot possible

$\Sigma_i x$ & $\Sigma_i y$ canot be l.i. So it's impossible.

This case is not possible

There cannot be a matrix A

with these null space dimensions.

FABER-CASTELL

EXAMPLE:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 2 \end{bmatrix} \qquad d(s) = (s-1)^3 (s-2)$$

$J = ?$

$B = ?$ st $J = B^{-1} A B$

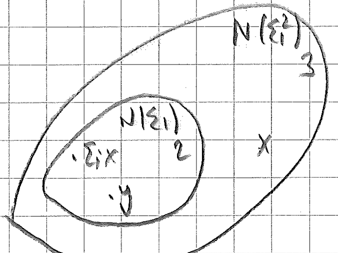
$$\Sigma_1 = A - \lambda_1 I = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix} \qquad \dim N(\Sigma_1) = 2$$

$$\Sigma_1^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \qquad \underset{\downarrow \; m_1}{\dim N(\Sigma_1^2)} = \underset{\substack{\| \\ r_1}}{3} \quad \Rightarrow \quad m(s) = (s-1)^2 (s-2)$$

$$\Sigma_2 = A - 2I = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix} \qquad \dim N(\Sigma_2) = 1$$

$$\underset{2}{N(\Sigma_1)} \subsetneq \underset{3}{N(\Sigma_1^2)}$$

$$J = \left[\begin{array}{cc|c|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right]$$



$x = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ ✓ $\qquad \cancel{x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}}$ ✗ $\qquad x = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ ✓

↓

choose this one

$$\Sigma_1 x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

↓

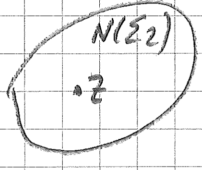
eigen vector.

$$A \, \Sigma_1 x = \underset{\substack{\downarrow \\ \lambda_1}}{1} \, \Sigma_1 x$$

$y \in N(\Sigma_1)$ & y is li with $\Sigma_1 x$ $\qquad y = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

Find z s.t $z \in N(\Sigma_2)$

$$(A - \lambda_2 I) z = 0$$



$$z = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} = [\Sigma_1 x \mid x \mid y \mid z]$$

HERMITIAN MATRICES

Definition: An $n \times n$ complex matrix A is called Hermitian if

$$\bar{A}^T = A,$$

i.e, its complex conjugate transpose equals itself.

$$\bar{A}^T = A^H$$

For Hermitian matrix, we have $A^H = A$

When $A^H = A$ for real matrix, we have $A^T = A$ which means A is symmetric matrix.

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad d(s) = (s-2)^2 - 1 = s^2 - 4s + 3 = (s-3)(s-1)$$

$$\lambda_1 = 3 \quad \lambda_2 = 1$$

For Hermitian matrices, eigen values are always real.

$$e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvectors are going to be orthogonal to each other, for different eigenvalues.

Theorem: Let A be Hermitian. Then, $\langle x, Ax \rangle$ is real for all $x \in \mathbb{C}^n$

Proof: $\langle x, y \rangle = y^H x$

$$\langle x, Ax \rangle = (Ax)^H x = x^H A^H x = x^H A x \quad : 1\times1 \text{ matrix, scalar}$$

$$\langle x, Ax \rangle = x^H A x : \text{scalar (transpose of scalar is itself)}$$

$$\overline{\langle x, Ax \rangle} = x^T \bar{A} \bar{x} \underset{\substack{\downarrow \\ \text{take} \\ \text{transpose}}}{=} \bar{x}^T \bar{A}^T x = x^H A^H x = \langle x, Ax \rangle$$

$\Rightarrow \langle x, Ax \rangle$ is real

Property:

$$\langle x, Bx \rangle = x^H B^H x = (Bx)^H x = x^H (B^H x) = \langle B^H x, x \rangle$$

$$\Rightarrow \langle x, Bx \rangle = \langle B^H x, x \rangle$$

where B is arbitrary square matrix

05.01.2016
Tuesday

$$\langle x, y \rangle \triangleq y^H x$$

$$\langle x, By \rangle = (By)^H x$$

$$= y^H B^H x$$

$$\langle B^H x, y \rangle = y^H B^H x$$

$$\boxed{\langle x, By \rangle = \langle B^H x, y \rangle}$$

Theorem: Let A be Hermitian. Then, $\langle x, Ax \rangle$ is real.

$$\bar{A}^T = A^H = A$$

$$\overline{\langle x, Ax \rangle} = \langle Ax, x \rangle = \langle x, A^H x \rangle = \langle x, Ax \rangle$$

$$\Rightarrow \langle x, Ax \rangle \text{ is real } \forall x$$

Theorem: If A is Hermitian, its eigenvalues are real.

Proof: Let λ_i be an eigenvalue of A.

Let e_i be an eigenvector corresponding to λ_i.

$$\langle e_i, Ae_i \rangle = \langle e_i, \lambda_i e_i \rangle = \bar{\lambda}_i \langle e_i, e_i \rangle = \bar{\lambda}_i \|e_i\|^2$$

$$\langle e_i, Ae_i \rangle = \langle A^H e_i, e_i \rangle = \langle Ae_i, e_i \rangle = \langle \lambda_i e_i, e_i \rangle = \lambda_i \|e_i\|^2$$

$$\Rightarrow \bar{\lambda}_i \cancel{\|e_i\|^2} = \lambda_i \cancel{\|e_i\|^2} \Rightarrow \bar{\lambda}_i = \lambda_i$$

$$\Rightarrow \lambda_i \text{ is always real}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$\lambda_1 = 3 \quad e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = 1 \quad e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$e_1^T e_2 = \langle e_1, e_2 \rangle = 0$$

Theorem: Let A be Hermitian and λ_i, λ_j be two distinct eigenvalues with eigenvectors e_i, e_j $(\lambda_i \neq \lambda_j)$. Two eigenvectors e_i & e_j are orthogonal.

$$\langle e_i, Ae_j \rangle = \langle e_i, \lambda_j e_j \rangle = \lambda_j \langle e_i, e_j \rangle$$

$$\langle e_i, A e_j \rangle = \langle A^H e_i, e_j \rangle = \langle Ae_i, e_j \rangle = \langle \lambda_i e_i, e_j \rangle = \lambda_i \langle e_i, e_j \rangle$$

$$\Rightarrow \underset{\substack{\downarrow \\ 3}}{\lambda_j} \langle e_i, e_j \rangle = \underset{\substack{\downarrow \\ 2}}{\lambda_i} \langle e_i, e_j \rangle$$

$$\underbrace{(\lambda_j - \lambda_i)}_{\neq 0} \underbrace{\langle e_i, e_j \rangle}_{=0} = 0 \Rightarrow \langle e_i, e_j \rangle = 0$$

Theorem: Let A be Hermitian. Then its minimal polynomial is

$$m(s) = (s - \lambda_1)(s - \lambda_2) \ldots (s - \lambda_\sigma)$$

i.e, $m_i = 1$ for $i = 1, 2, \ldots, \sigma$

$$\left.\begin{aligned} d(s) &= (s-\lambda_1)^{r_1} (s-\lambda_2)^{r_2} \ldots (s-\lambda_\sigma)^{r_\sigma} \\ m(s) &= (s-\lambda_1)^{m_1} (s-\lambda_2)^{m_2} \ldots (s-\lambda_\sigma)^{m_\sigma} \end{aligned}\right\} \begin{aligned} &\text{If } A \text{ is Hermitian} \\ &m_1 = m_2 = \ldots = m_\sigma = 1 \end{aligned}$$

FABER-CASTELL

Proof:

$$N((A-\lambda_i I)) \overset{?}{=} N((A-\lambda_i I)^2)$$

We need to prove that

$N((A-\lambda_i I)) \subset N((A-\lambda_i I)^2)$ This was already proven in previous lectures.

So, we have to prove that

$$N((A-\lambda_i I)^2) \subset N((A-\lambda_i I))$$

Let x be arbitrary in $N((A-\lambda_i I)^2)$

$$x \in N((A-\lambda_i I)^2)$$

$$(A-\lambda_i I)^2 x = 0$$

$$\langle x, (A-\lambda_i I)^2 x \rangle = 0$$

$$\langle x, (A-\lambda_i I)(A-\lambda_i I)x \rangle = \langle \underbrace{(A-\lambda_i I)^H}_{\substack{A^H - \lambda_i I \\ \shortparallel \\ A-\lambda_i I}} x, (A-\lambda_i I)x \rangle = \langle (A-\lambda_i I)x, (A-\lambda_i I)x \rangle = 0$$

$$\Rightarrow (A-\lambda_i I)x = 0 \Rightarrow x \in N(A-\lambda_i I)$$

$$\Rightarrow \underbrace{N(A-\lambda_i I)}_{r_i} = N((A-\lambda_i I)^2)$$

$$\Rightarrow m_i = 1 \text{ for } i = 1, 2, \dots, \sigma$$

General Case:

$$\mathbb{C}^n = N(\underbrace{(A-\lambda_1 I)^{m_1}}_{\dim r_1}) \oplus N(\underbrace{(A-\lambda_2 I)^{m_2}}_{\dim r_2}) \oplus \dots \oplus N((A-\lambda_\sigma I)^{m_\sigma})$$

Hermitian Case

$$\mathbb{C}^n = \underbrace{N((A-\lambda_1 I))}_{\dim r_1} \overset{\perp}{\oplus} \underbrace{N((A-\lambda_2 I))}_{\dim r_2} \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} \underbrace{N((A-\lambda_\sigma I))}_{\dim r_\sigma}$$

$$\underbrace{N((A-\lambda_i I))}_{r_i} = \underbrace{N((A-\lambda_i I)^2)}_{r_i}$$

# of blocks corresponding to λ_i (with size ≥ 1) is r_i

" " " " " (with size ≥ 2) is 0

$$J = \begin{bmatrix} \bar{A}_1 & & & \\ & \bar{A}_2 & & \\ & & \ddots & \\ & & & \bar{A}_\sigma \end{bmatrix}$$

$\bar{A}_i$ is of size $r_i \times r_i$

$$\bar{A}_i = \lambda_i I = \begin{bmatrix} \lambda_i & & & \\ & \lambda_i & & \\ & & \ddots & \\ & & & \lambda_i \end{bmatrix}_{r_i \times r_i} \rightarrow r_i \text{ blocks of size } 1.$$

Jordan canonical form for a Hermitian Matrix is DIAGONAL.

$$J = \begin{bmatrix} \lambda_1 I_{r_1 \times r_1} & & & 0 \\ & \lambda_2 I_{r_2 \times r_2} & & \\ & 0 & \ddots & \\ & & & \lambda_\sigma I_{r_\sigma \times r_\sigma} \end{bmatrix}$$

Theorem. Let A be an $n \times n$ Hermitian Matrix with

$d(s) = \prod_{i=1}^{\sigma} (s-\lambda_i)^{r_\sigma}$. Then there exist a unitary matrix

P. (i.e $P^{-1} = P^H$) such that

$$J = P^H A P$$

~~Proof~~: We know that $\exists P$ such that

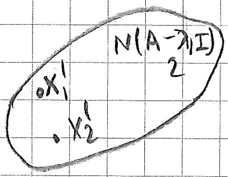
$$J = \underbrace{P^{-1}}_{=P^H} A P$$

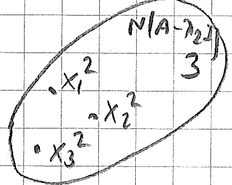
let us have $A \in \mathbb{C}^{5 \times 5}$ $A^H = A$

Suppose that

$$d(s) = (s-\lambda_1)^2 (s-\lambda_2)^3$$

$$m(s) = (s-\lambda_1)^1 (s-\lambda_2)$$





$\{x_1^1\}\ \{x_2^1\}$ $\quad$ $\{x_1^2\}\ \{x_2^2\}\ \{x_3^2\}$

Apply Gram Schmidt Orthonormalization to each set of vectors

$\{e_1^1, e_2^1\}$ $\qquad$ $\{e_1^2, e_2^2, e_3^2\}$

$$P = \left[e_1^1 \;\; e_2^1 \;\; e_1^2 \;\; e_2^2 \;\; e_3^2 \right]$$

$$P^H P = \begin{bmatrix} (e_1^1)^H \\ (e_2^1)^H \\ (e_1^2)^H \\ (e_2^2)^H \\ (e_3^2)^H \end{bmatrix} \left[e_1^1 \;\; e_2^1 \;\; e_1^2 \;\; e_2^2 \;\; e_3^2 \right] = I$$

(2 saat Ek Ders) $\qquad$ 06.01.2016 Wednesday

$A^H = A$

$\langle x, By \rangle = \langle B^H x, y \rangle$

arbitrary

* $\langle x, Ax \rangle$ is real
* λ_i is real
* $\lambda_i \neq \lambda_j$ $\quad e_i$ & e_j are orthogonal

$\langle e_i, e_j \rangle = 0$

$$d(s) = (s-\lambda_1)^{r_1} (s-\lambda_2)^{r_2} \ldots (s-\lambda_\sigma)^{r_\sigma}$$

$$m(s) = (s-\lambda_1)(s-\lambda_2) \ldots (s-\lambda_\sigma)$$

FABER-CASTELL

$$\underbrace{N(A-\lambda_i I)}_{r_i} = N((A-\lambda_i I)^2)$$

$$\mathbb{C}^n = N((A-\lambda_1 I)^{m_1}) \oplus N((A-\lambda_2 I)^{m_2}) \oplus \dots \oplus N((A-\lambda_\sigma I)^{m_\sigma}) \leftarrow \text{general case}$$

$$\mathbb{C}^n = N(A-\lambda_1 I) \overset{\perp}{\oplus} N(A-\lambda_2 I) \overset{\perp}{\oplus} \dots \oplus N(A-\lambda_\sigma I) \leftarrow \text{Hermitian case}$$

$$J = \begin{bmatrix} \bar{A}_1 & & 0 \\ & \bar{A}_2 & \\ 0 & & \ddots \ \bar{A}_\sigma \end{bmatrix} \qquad \bar{A}_i = \lambda_i I_{r_i \times r_i}$$

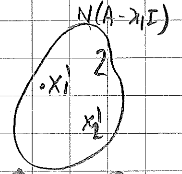
$J = P^{-1}AP$

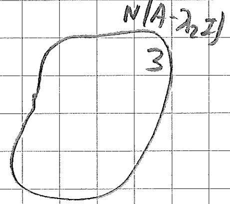
$\exists P$ which satisfies $\underbrace{P^{-1} = P^H}_{\text{unitary matrix}}$ & $J = P^H A P$

Idea of the Proof: $A \in \mathbb{C}^{5\times 5}$

$d(s) = (s-\lambda_1)^2 (s-\lambda_2)^3$

$m(s) = (s-\lambda_1)(s-\lambda_2)$





$\{x_1^1\}$ $\{x_2^1\}$

$\{x_1^2\}$ $\{x_2^2\}$ $\{x_3^2\}$

x_1^1 & x_2^1 are l.i but they don't have to be orthogonal. However, x_1^1 & x_1^2, x_2^2, x_3^2 must be both l.i & orthogonal.

$\underbrace{\{e_1^1 \ e_2^1\}}_{\text{both l.i. \& orthonormal}}$ $\underbrace{\{e_1^2 \ e_2^2 \ e_3^2\}}_{\text{both l.i \& orthonormal}}$

$P = [e_1^1 \ e_2^1 \ e_1^2 \ e_2^2 \ e_3^2]$

$P^H P$

$$\begin{bmatrix} (e_1^1)^H \\ (e_2^1)^H \\ (e_1^2)^H \\ (e_2^2)^H \\ (e_3^2)^H \end{bmatrix} \begin{bmatrix} e_1^1 & e_2^1 & e_1^2 & e_2^2 & e_3^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = I$$

Theorem: Let A be an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_\sigma$.

$$\lambda_{min} \triangleq \min_i \lambda_i \qquad \lambda_{max} \triangleq \max_i \lambda_i$$

Then we have

$$\lambda_{min} \langle x, x \rangle \leq \langle x, Ax \rangle \leq \lambda_{max} \langle x, x \rangle \qquad \forall x \in \mathbb{C}^n$$

<u>Proof:</u> $\mathbb{C}^n = N(A - \lambda_1 I) \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} N(A - \lambda_\sigma I)$

We can decompose an arbitrary $x \in \mathbb{C}^n$ as

$$x = x_1 + x_2 + \cdots + x_\sigma \qquad x_i \in N(A - \lambda_i I)$$

$$\langle x_i, x_j \rangle = 0 \text{ when } i \neq j$$

$$x = \sum_{i=1}^{\sigma} x_i$$

$$\langle x, x \rangle = \left\langle \sum_{i=1}^{\sigma} x_i, \sum_{j=1}^{\sigma} x_j \right\rangle = \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} \langle x_i, x_j \rangle = \sum_{i=1}^{\sigma} \langle x_i, x_i \rangle$$

$$\Rightarrow \boxed{\langle x, x \rangle = \sum_{i=1}^{\sigma} \langle x_i, x_i \rangle}$$

$$\langle x, Ax \rangle = \left\langle \sum_{i=1}^{\sigma} x_i, A \sum_{j=1}^{\sigma} x_j \right\rangle = \left\langle \sum_{i=1}^{\sigma} x_i, \sum_{j=1}^{\sigma} \underbrace{Ax_j}_{\lambda_j x_j} \right\rangle = \left\langle \sum_{i=1}^{\sigma} x_i, \sum_{j=1}^{\sigma} \lambda_j x_j \right\rangle$$

$$= \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} \lambda_j \langle x_i, x_j \rangle = \sum_{i=1}^{\sigma} \lambda_i \langle x_i, x_i \rangle$$

$$\Rightarrow \boxed{\langle x, Ax \rangle = \sum_{i=1}^{\sigma} \lambda_i \langle x_i, x_i \rangle}$$

$$\langle x, Ax\rangle = \sum_{i=1}^{\sigma} \lambda_i \langle x_i, x_i\rangle \leq \sum_{i=1}^{\sigma} \lambda_{max} \langle x_i, x_i\rangle = \lambda_{max} \sum_{i=1}^{\sigma} \langle x_i, x_i\rangle = \lambda_{max}\langle x, x\rangle$$

$$\langle x, Ax\rangle = \sum_{i=1}^{\sigma} \lambda_i \langle x_i, x_i\rangle \geq \sum_{i=1}^{\sigma} \lambda_{min} \langle x_i, x_i\rangle = \lambda_{min} \sum_{i=1}^{\sigma} \langle x_i, x_i\rangle = \lambda_{min}\langle x, x\rangle$$

$$\Rightarrow \boxed{\lambda_{min}\langle x, x\rangle \leq \langle x, Ax\rangle \leq \lambda_{max}\langle x, x\rangle}$$

If $x \neq 0$

$$\lambda_{min}(A) \leq \frac{\langle x, Ax\rangle}{\langle x, x\rangle} \leq \lambda_{max}(A)$$

If A is not hermitian, this type quantity is not applicable because all λ_{max}, λ_{min}, $\langle x, Ax\rangle$ can be complex. & we cannot order them like that.

Application: Remember the definition of an induced matrix norm

$$\|A\| \triangleq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Let $\|\cdot\|$ be the induced norm for the inner product $\langle x, y\rangle$.

$$\frac{\|Ax\|^2}{\|x\|^2} \triangleq \frac{\langle Ax, Ax\rangle}{\langle x, x\rangle} = \frac{\langle x, A^H A x\rangle}{\langle x, x\rangle} \leq \lambda_{max}(A^H A)$$

$$(A^H A)^H = A^H A$$

$$\Rightarrow \|A\|^2 = \lambda_{max}(A^H A)$$

$$(A^H A)^H = A^H A$$

$$\Rightarrow \|A\| = \sqrt{\lambda_{max}(A^H A)}$$

DEFINITION: A Hermition matrix A is said to be positive definite if

$$\langle x, Ax \rangle > 0 \quad \forall\, x \neq 0$$

(positive semi-definite when $\langle x, Ax \rangle \geq 0$)

$$\langle x, A^H A x \rangle = \langle Ax, Ax \rangle \geq 0$$

$A^H A$ is at least positive semi definite.

Theorem: If A is positive definite, then all eigenvalues of A are positive.

λ_i, e_i is an arbitrary eigenvalue-eigenvector pair for A.

$$\langle e_i, Ae_i \rangle > 0$$

$$\langle e_i, \lambda_i e_i \rangle = \lambda_i \langle e_i, e_i \rangle > 0$$

$$\Rightarrow \lambda_i > 0$$

FUNCTIONS OF A MATRIX

MOTIVATION: $\dot{x} = Ax$

$\dot{x} \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$, $x \in \mathbb{C}^n$

$$x(0) = x_0 \in \mathbb{C}^n$$

$$x(t) = \sum_{i=0}^{\infty} \alpha_i t^i$$

$\alpha_i \in \mathbb{C}^n$

$$x(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots$$

$$x_0 = \alpha_0 \Rightarrow \alpha_0 = x_0$$

$$\dot{x} = Ax(t) = \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2 + \ldots$$

$$Ax_0 = \alpha_1 \Rightarrow \alpha_1 = Ax_0$$

$$\ddot{x} = A\dot{x} = A^2 x(t) = 2\alpha_2 + 6\alpha_3 t + \dots$$

$$A^2 x_0 = 2\alpha_2 \Rightarrow \alpha_2 = \frac{A^2 x_0}{2}$$

$$\alpha_3 = \frac{A^3 x_0}{6}$$

$$\alpha_i = \frac{A^i x_0}{i!}$$

$$\Rightarrow x(t) = \sum_{i=0}^{\infty} \frac{A^i}{i!} x_0 t^i$$

$$= \sum_{i=0}^{\infty} \frac{(At)^i}{i!} x_0 = \underbrace{\left(\sum_{i=0}^{\infty} \frac{(At)^i}{i!} \right)}_{e^{At}} x_0 \qquad \left(e^t = \sum_{i=0}^{\infty} \frac{t^i}{i!} \right)$$

$$\Rightarrow \boxed{x(t) = e^{At} x_0}$$

$$e^A \triangleq \sum_{i=0}^{\infty} \frac{A^i}{i!} = c_0 I + c_1 A + c_2 A^2 + \dots + c_{n-1} A^{n-1}$$

$$e^A = c_0 I + c_1 A + c_2 A^2 + \dots + c_{n-1} A^{n-1}$$

$$p(s) = c_0 + c_1 s + c_2 s^2 + \dots + c_{n-1} s^{n-1}$$

$$e^A = p(A)$$

How can we calculate $e^A = ?$

SPECIAL CASE

Case that eigenvalues of A are distinct.

$$d(s) = (s-\lambda_1)(s-\lambda_2) \cdots\cdots (s-\lambda_n)$$
$$m(s) = (s-\lambda_1)(s-\lambda_2) \cdots\cdots (s-\lambda_n)$$

$$J = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Solution-1: Consider eigenvalue λ_j $\quad J = 1, \dots, n$

eigenvector e_J

$$e^A = p(A)$$

$$Ae_J = \lambda_J e_J$$
$$A^2 e_J = \lambda_J^2 e_J$$
$$\vdots$$
$$\boxed{A^i e_J = \lambda_J^i e_J}$$

$$e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!}$$

$$e^A e_J = \sum_{i=0}^{\infty} \frac{A^i}{i!} e_J = \sum_{i=0}^{\infty} \frac{\lambda_J^i}{i!} e_J = \left(\sum_{i=0}^{\infty} \frac{\lambda_J^i}{i!}\right) e_J = e^{\lambda_J} e_J$$

$$\boxed{e^A e_J = e^{\lambda_J} e_J}$$

$$p(s) = c_0 + c_1 s + c_2 s^2 + \cdots\cdots + c_{n-1} s^{n-1}$$
$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots\cdots + c_{n-1} A^{n-1}$$
$$p(A) e_J = c_0 e_J + c_1 \lambda_J e_J + c_2 \lambda_J^2 e_J + \cdots\cdots + c_{n-1} \lambda_J^{n-1} e_J$$
$$= \underbrace{(c_0 + c_1 \lambda_J + c_2 \lambda_J^2 + \cdots\cdots + c_{n-1} \lambda_J^{n-1})}_{p(\lambda_J)} e_J$$

$$= p(\lambda_J) e_J$$

$$\boxed{p(A) e_J = p(\lambda_J) e_J}$$

$$e^{\lambda_J} e_J = p(\lambda_J)\, e_J$$

$$\boxed{p(\lambda_J) = e^{\lambda_J}} \qquad J = 1, \dots, n$$

$$\left.\begin{aligned} p(\lambda_1) &= e^{\lambda_1} \\ p(\lambda_2) &= e^{\lambda_2} \\ &\vdots \\ p(\lambda_n) &= e^{\lambda_n} \end{aligned}\right\} \qquad \begin{aligned} c_0 + c_1\lambda_1 + c_2\lambda_1^2 + \dots &= e^{\lambda_1} \\ c_0 + c_1\lambda_2 + c_2\lambda_2^2 + \dots &= e^{\lambda_2} \\ &\vdots \\ c_0 + c_1\lambda_n + c_2\lambda_n^2 + \dots &= e^{\lambda_n} \end{aligned}$$

$$\underbrace{\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & & & & \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{n-1} \end{bmatrix}}_{\mathbb{C}} \underbrace{\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}}_{C} = \underbrace{\begin{bmatrix} e^{\lambda_1} \\ e^{\lambda_2} \\ \vdots \\ e^{\lambda_n} \end{bmatrix}}_{E} \begin{matrix} \to \sin\lambda_1 & \log\lambda_1 \\ \to \sin\lambda_2 & \log\lambda_2 \\ \vdots & \vdots \\ \sin\lambda_n & \log\lambda_n \end{matrix}$$

$$C = \mathbb{C}^{-1} E$$

In the general case when we try to find $f(A)$, we will use the equations

$$p(\lambda_J) = f(\lambda_J) \qquad J = 1, 2, \dots, n$$

EXAMPLE:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \begin{aligned} \lambda_1 &= 3 & e_1 &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \lambda_2 &= 1 & e_2 &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} e^{\lambda_1} \\ e^{\lambda_2} \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} e^3 \\ e \end{bmatrix}$$

$$* \; e^A = p(A) = c_0 I + c_1 A$$

$$\Rightarrow \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} e^3 \\ e \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} e^3 - 3e \\ -e^3 + e \end{bmatrix}$$

$$\Rightarrow e^A = -\frac{1}{2}(e^3 - 3e) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{e^3 - e}{2} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

* calculate $\log(A)$

$$\begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \log 3 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \log 3 \\ 0 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \log 3$$

$$\Rightarrow \log A = p(A) = c_0 I + c_1 A = -\frac{\log 3}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2} \log 3 \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

FIRST SOLUTION

08.01.2015

Friday

$A: \mathbb{C}^{n\times n}$

$f(A) = p(A)$

$= c_0 I + c_1 A + \dots + c_{n-1} A^{n-1}$ } Case of distinct eigen values

↳ if eigenvalues are NOT distinct, power of $m(s)$ can change, then power of $p(A)$ can also change.

Distinct Eigenvalue Case

$d(s) = (s-\lambda_1)(s-\lambda_2)\dots\dots(s-\lambda_n)$

$m(s) = (s-\lambda_1)\dots\dots(s-\lambda_n)$

$p(\lambda_i) = f(\lambda_i) \quad i = 1, \dots, n$

$p(s) = c_0 + c_1 s + \dots + c_{n-1} s^{n-1}$

$f(\cdot)$ has to be defined at λ_i [for example $f(s) = \frac{1}{s}$, if you've eigenvalue $\lambda_i = 0$, $f(s)$ is not defined.]

$$\underbrace{\begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ | & | & & | \\ | & | & & | \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{bmatrix}}_{\text{Vander Monde Matrix}} \begin{bmatrix} c_0 \\ c_1 \\ | \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} f(\lambda_1) \\ | \\ | \\ f(\lambda_n) \end{bmatrix}$$

SECOND SOLUTION

(This solution is valid not only when A has distinct eigenvalues but also when A is diagonalizable. (J is diagonal))

$$J = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ & 0 & & \lambda_n \end{bmatrix}$$

$J = P^{-1} A P$ ↳ we know how to find P

$\boxed{A = PJP^{-1}}$

$$e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!}$$

$A^2 = PJP^{-1}PJP^{-1} = PJ^2P^{-1}$

$A^3 = PJ^3P^{-1}$

|

$A^i = PJ^iP^{-1}$

$$e^A = \sum_{i=0}^{\infty} \frac{PJ^iP^{-1}}{i!} = P\underbrace{\left[\sum_{i=0}^{\infty} \frac{J^i}{i!}\right]}_{e^J} P^{-1}$$

$= Pe^J P^{-1}$

In general, $\boxed{f(A) = P f(J) P^{-1}}$

$$J^i = \begin{bmatrix} \lambda_1^i & & 0 \\ & \lambda_2^i & \\ 0 & & \ddots & \\ & & & \lambda_n^i \end{bmatrix}$$

$$\Rightarrow e^A = P \sum_{i=0}^{\infty} \frac{1}{i!} \begin{bmatrix} \lambda_1^i & & 0 \\ & \lambda_2^i & \\ 0 & & \ddots & \\ & & & \lambda_n^i \end{bmatrix} P^{-1}$$

$$= P \begin{bmatrix} \sum_{i=0}^{\infty} \frac{\lambda_1^i}{i!} & & & 0 \\ & \sum_{i=0}^{\infty} \frac{\lambda_2^i}{i!} & & \\ & & \ddots & \\ 0 & & & \sum_{i=0}^{\infty} \frac{\lambda_n^i}{i!} \end{bmatrix} P^{-1}$$

$$= P \underbrace{\begin{bmatrix} e^{\lambda_1} & & 0 \\ & e^{\lambda_2} & \\ & & \ddots & \\ 0 & & & e^{\lambda_n} \end{bmatrix}}_{e^J} P^{-1}$$

General Case

$$f(A) = P \begin{bmatrix} f(\lambda_1) & & 0 \\ & f(\lambda_2) & \\ 0 & & \ddots & \\ & & & f(\lambda_n) \end{bmatrix} P^{-1}$$

★ eigenvalues of $f(A)$ are $f(\lambda_i)$ $i=1,\dots,n$!

EXAMPLE:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$\lambda_1 = 3$ $\quad e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = 1$ $\quad e_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$P = [e_1 \,\vdots\, e_2]$

$J = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

$A = P \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} P^{-1}$

$$e^A = P \begin{bmatrix} e^3 & 0 \\ 0 & e \end{bmatrix} P^{-1}$$

→ the result is the same with the result of first solution.

Even if the roots of J is repeated, we can apply this solution method

EXAMPLE: $f(s) = (1-s)^{-1}$

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{matrix} \lambda_1 = 0 \\ \lambda_2 = 1 \end{matrix}$$

$$\cancel{f(A) = P \begin{bmatrix} f(0) & 0 \\ 0 & f(1) \end{bmatrix} P^{-1}}$$

→ $\lambda = 1$ is problematic one.

→ $f(A)$ does NOT exist

↳ $f(\cdot)$ is not analytic at $\lambda = 1$

EXAMPLE:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Calculate $A^{1/2}$ $(\sqrt{A})$

$$\sqrt{A}\sqrt{A} = A$$

$$\sqrt{A} = P\sqrt{J}\,P^{-1} = P\sqrt{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}\,P^{-1}$$

$$= P\begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{1} \end{bmatrix} P^{-1} \quad \text{where} \quad P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

→ positive definite sqrt of A because its roots are positive

$$\sqrt{A} = P\begin{bmatrix} -\sqrt{3} & 0 \\ 0 & -\sqrt{1} \end{bmatrix} P^{-1}$$

⇒ $\sqrt{A}$ is not be UNIQUE! [2^n tane ver]

GENERAL CASE

A does not have distinct eigenvalues.

$$f(A) = \sum_{i=0}^{\infty} \beta_i A^i$$

$$d(s) = (s-\lambda_1)^{r_1}(s-\lambda_2)^{r_2} \ldots (s-\lambda_\sigma)^{r_\sigma}$$

$$m(s) = (s-\lambda_1)^{m_1}(s-\lambda_2)^{m_2} \ldots (s-\lambda_\sigma)^{m_\sigma}$$

$$m(A) = 0 \qquad \underbrace{M}_{\text{order of } m(s)} \triangleq \sum_{i=1}^{\sigma} m_i$$

A^M = linear combination of smaller powers of A.

$$f(A) = c_0 I + c_1 A + \ldots + c_{M-1} A^{M-1}$$

SOLUTION-1

$f(A) = p(A)$

Apply the same analysis

M unknowns

$$\left.\begin{aligned} p(\lambda_1) &= f(\lambda_1) \\ p(\lambda_2) &= f(\lambda_2) \\ &\vdots \\ p(\lambda_\sigma) &= f(\lambda_\sigma) \end{aligned}\right\} \sigma \text{ different equations}$$

! σ is less than M, we have more unknowns than equations.

If $\sigma = M = n$ (distinct eigenvalues)

EXAMPLE:

$d(s) = (s-3)^2 (s-4)^5$

$m(s) = (s-3)^2 (s-4)^3$

$M = 5$

$p(s) = c_0 + c_1 s + c_2 s^2 + c_3 s^3 + c_4 s^4$

→ # of unknowns $= 5$

$\lambda_1 = 3$

$$\left.\begin{aligned} p(3) &= f(3) \\ p'(3) &= f'(3) \end{aligned}\right\} \text{\# of equations } 2$$

$\lambda_2 = 4$

$$\underbrace{\begin{aligned} p(4) &= f(4) \\ p'(4) &= f'(4) \\ p''(4) &= f''(4) \end{aligned}}_{\text{Now, we have 5 equations!}}$$

DEFINITION: Let f be an analytic complex valued function defined in an open set containing $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_\sigma\}$ (spectrum of A) Let p be any polynomial satisfying

$$p^{(k)}(\lambda_i) = f^{(k)}(\lambda_i) \qquad \begin{aligned} i &= 1, \dots, \sigma \\ 0 &\leq k \leq m_i - 1 \end{aligned}$$

Then, we define

$$f(A) = p(A)$$

EXAMPLE:

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \qquad \begin{aligned} d(s) &= (s-1)^2(s-2) \\ m(s) &= (s-1)^2(s-2) \\ \mu &= 3 \end{aligned}$$

Find $\sin(\pi A)$

SOLUTION-1

$f(s) = \sin(\pi s)$

$p(s) = c_0 + c_1 s + c_2 s^2$

$$\left.\begin{aligned} p(1) = f(1) &\Rightarrow c_0 + c_1 + c_2 = 0 \\ p'(1) = f'(1) &\Rightarrow c_1 + 2c_2 = -\pi \\ p(2) = f(2) &\Rightarrow c_0 + 2c_1 + 4c_2 = 0 \end{aligned}\right\} \quad \begin{aligned} c_0 &= 2\pi \\ c_1 &= -3\pi \\ c_2 &= \pi \end{aligned}$$

$$f(A) = p(A) = c_0 I + c_1 A + c_2 A^2$$

$$= 2\pi \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 3\pi \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix} + \pi \begin{bmatrix} -2 & -2 & -6 \\ 0 & 1 & 0 \\ 3 & 4 & 7 \end{bmatrix}$$

SOLUTION -2

$J = P^{-1}AP$ ↳ we know how to find P

$A = PJP^{-1}$
↳ J doesnot have to be diagonal

$$J = \begin{bmatrix} J_1 & & 0 \\ & J_2 & \\ 0 & & J_L \end{bmatrix}$$

Jordan blocks

$$J = \begin{bmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \qquad J_1 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad J_2 = 2 \quad J_3 = 1 \quad J_4 = 1$$

$$f(s) = \sum_{i=0}^{\infty} \beta_i s^i \longrightarrow f(A) = \sum_{i=0}^{\infty} \beta_i A^i \quad \& \quad A^i = PJ^iP^{-1}$$

$$\Rightarrow \quad f(A) = \sum_{i=0}^{\infty} \beta_i P J^i P^{-1} = P\left[\sum_{i=0}^{\infty} \beta_i J^i\right] P^{-1} = P f(J) P^{-1}$$

$$J^i = \begin{bmatrix} J_1^i & & & 0 \\ & J_2^i & & \\ & & \ddots & \\ 0 & & & J_L^i \end{bmatrix}$$

$$\Rightarrow \quad f(A) = P\left[\sum_{i=0}^{\infty} \beta_i \begin{bmatrix} J_1^i & & & 0 \\ & J_2^i & & \\ & & \ddots & \\ 0 & & & J_L^i \end{bmatrix}\right] P^{-1}$$

$$= P \begin{bmatrix} \sum_{i=0}^{\infty} \beta_i J_1^i & & & 0 \\ & \sum_{i=0}^{\infty} \beta_i J_2^i & & \\ & & \ddots & \\ 0 & & & \sum_{i=0}^{\infty} \beta_i J_L^i \end{bmatrix} P^{-1}$$

$$= P \begin{bmatrix} f(J_1) & & & 0 \\ & f(J_2) & & \\ & & \ddots & \\ 0 & & & f(J_L) \end{bmatrix} P^{-1}$$

$$f(J_l) = ?$$

$$J_l = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \qquad l = 1, \dots, L$$

$$f(s) = \sum_{i=0}^{\infty} \beta_i s^i$$ → it is around zero.

$$f(s) = \sum_{i=0}^{\infty} \frac{f^{(i)}(\lambda)}{i!} (s-\lambda)^i$$ → it is arond λ

$$= f(\lambda) + f'(\lambda)(s-\lambda) + \frac{f''(\lambda)}{2}(s-\lambda)^2 + \dots$$

$$f(J_l) = \sum_{i=0}^{\infty} \frac{f^{(i)}(\lambda)}{i!} (J_l - \lambda I)^i$$

$$= f(\lambda) I + f'(\lambda)(J_l - \lambda I) + \frac{f''(\lambda)}{2}(J_l - \lambda I)^2 + \dots$$

$$J_l - \lambda I = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix} \longrightarrow (J_l - \lambda I)^2 = \begin{bmatrix} 0 & 0 & 1 & & \\ & 0 & 0 & 1 & \\ & & \ddots & \ddots & 1 \\ & & & 0 & 0 \\ 0 & & & & 0 \end{bmatrix} \longrightarrow (J_l - \lambda I)^k = 0$$

$$f(J_l) = \sum_{0}^{k-1} \frac{f^{(i)}(\lambda)}{i!} (J_l - \lambda I)^i$$

$$f(J_l) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \cdots & \frac{f^{(k-1)}(\lambda)}{k-1} \\ & f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \vdots \\ & & \ddots & \ddots & \frac{f''(\lambda)}{2} \\ & 0 & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix}$$

EXAMPLE

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

$f(s) = \sin(\pi s)$

$f(A) = ?$

$J = P^{-1} A P$

$d(s) = (s-1)^2 (s-2)$

$m(s) = d(s)$

$$J = \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$$

$$A = PJP^{-1}$$

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

$$= P \left[\begin{array}{c|c} f\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) & 0 \\ \hline 0 & f(2) \end{array}\right] P^{-1} = P \left[\begin{array}{cc|c} \sin\pi & \pi\cos\pi & \\ 0 & \sin\pi & 0 \\ \hline 0 & 0 & \sin 2\pi \end{array}\right] P^{-1}$$

$$= P \begin{bmatrix} 0 & -\pi & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^{-1}$$