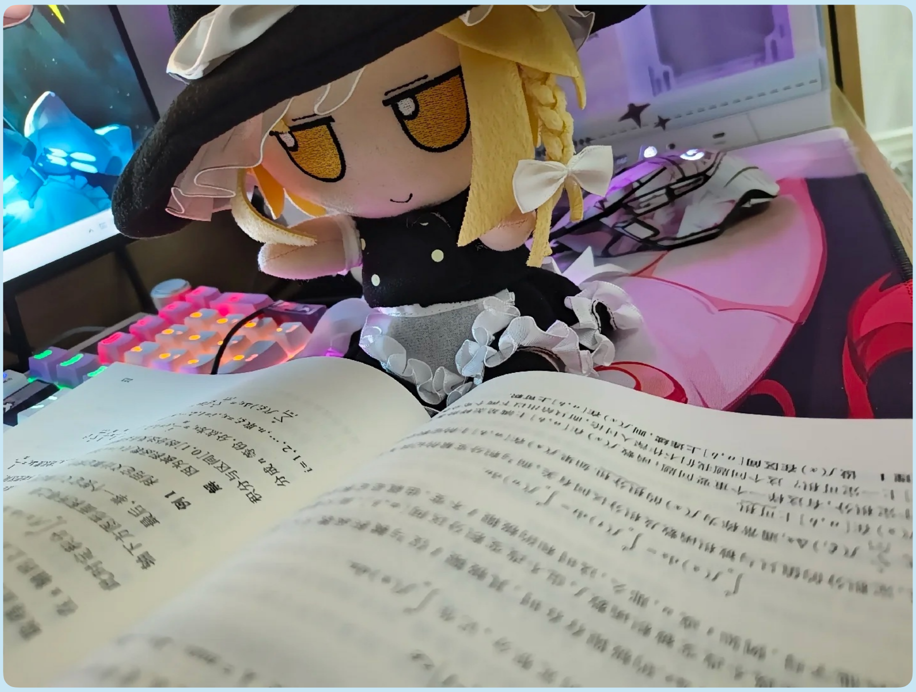
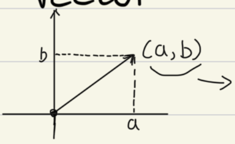
Linear Algebra

せんけいだいすう

Линейная алгебра



Vector



$\boxed{\begin{array}{l} f: \\ 1 \to a \\ 2 \to b \end{array}}$

Function

$f: \mathbb{R} \to \mathbb{R}$

$f: 1 \to a_1$
$\quad 2 \to a_2$
$\quad \cdots$
$\quad n \to a_n$

In the Linear function point of view, a function can be considered as a vector

$(a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) = (a_1+b_1, a_2+b_2, \cdots, a_n+b_n)$ $\quad \mathbb{R}^n$: n dimention

$(f_1+f_2)(x) = f_1(x) + f_2(x)$

Also we have $\underline{\underline{C}} \in \mathbb{R}$, C: scalor $\in F$

$C(a,b) = (Ca, C.b)$ and $(Cf)(x) = Cf(x)$

$F = \mathbb{R}$ or $\mathbb{C}$ complex numbers #s

Def: A vector space V over a field F ($\mathbb{R}$ or $\mathbb{C}$)

consist of a set on which on which two operations, addition, & scalan multiplication

$+: V \times V \to V \qquad +(x, y) = x+y$

$\cdot: F \cdot V \to V \qquad \cdot(a, x) = ax$

s.t. the following condition holds

(VS1) $x+y = y+x \quad \forall x, y \in V$

(VS2) $(x+y)+z = x+(y+z) \quad \forall x, y, z \in V$

(VS3) $\exists$ an element $0 \in V$, s.t. $0+x = x$, $\forall x \in V$

(VS4) $\forall x \in V$, $\exists y \in V$, s.t. $x+y=0$

(VS5) $\forall x \in V$, $1 \cdot x = x$

(VS6) $\forall x \in V, \forall a, b \in F \quad (ab)x = a(bx)$

(VS7) $\forall x, y \in V$, $a \in F \quad a(x+y) = ax+ay$

(VS8) $\forall x \in V \quad \forall a, b \in F$, $(a+b)x = ax+bx$

Elements in V: vectors while F: scalans

Examples: $\mathbb{R}^2$, $\mathbb{R}^n$, $\mathcal{F}(\mathbb{R})^2$ $\mathcal{C}(\mathbb{R})^2$

$P_n(\mathbb{R})$: the set off all polynomials with real coefficient of degree $\leq n$

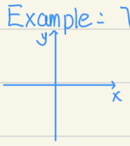
Subspace

Def: A Substract $W \in V$ is a subspace

W is a vector space with the same operation on V

$\mathbb{R}^2 = (a_1, a_2, 0, 0, \cdots, 0)$

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



$W_1 = \{(x, 0) \mid x \in \mathbb{R}\}$

$W_2 = \{(0, y) \mid y \in \mathbb{R}\}$

Prep: (1) Any intusection of subspaces in a subspace

(2) The urim of subspaces may NOT be a subspace

Proof: W_1. W_2 are two subspaces of V

Claim $W_1 \cap W_2$ is a subspace of V

$x, y \in W_1 \cap W_2 \Rightarrow \begin{cases} x, y \in W_1 \Rightarrow x+y \in W_1 \\ x, y \in W_2 \Rightarrow x+y \in W_2 \end{cases} \Rightarrow x+y \in W_1 \cap W_2$

Matrix

2x2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $(a,b,c,d) \sim \mathbb{R}^4$

m x n matrix $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$

Rk The set of all 2x2 matrices in a vector space with the "usual" addition and scalor multiplication

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

$$c\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{pmatrix}$$ entry

V : a vector space over F ($\mathbb{R}$ or $\mathbb{C}$)

Def: Let $v, u_1, u_2, \cdots, u_k \in V$ $\exists\, a_1, a_2, \cdots, a_k \in F$, s.t. $v = a_1u_1 + a_2u_2 + \cdots + a_ku_k$

Then we say that v is a linear combination of $u_1, u_2, \cdots, u_k$

$A \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}_{2\times3}$ $\left.\begin{matrix} A = m\times n \\ B : n\times p \end{matrix}\right\} AB = m\times p$

$B \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}_{3\times2}$

$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1\times1+2\times3+3\times5 & 1\times2+2\times4+3\times6 \\ 4\times1+5\times3+6\times5 & 4\times2+5\times4+6\times6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix}$$

Only when the column number of the first matrix equals to the row number that the product operation can be conducted

$$BA = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1\times1+2\times4 & 1\times2+2\times5 \\ 3\times1+4\times2 & 3\times4+4\times5 \end{pmatrix} = \begin{pmatrix} 9 & 12 \\ 11 & 32 \end{pmatrix}$$

$$\begin{cases} a_1 - 2a_2 \quad\quad + 2a_4 - 3a_5 = 2 & R_1 \\ 2a_1 - 4a_2 + 2a_3 \quad\quad + 8a_5 = 6 & R_2 \\ a_1 - 2a_2 + 3a_3 - 3a_4 + 16a_5 = 8 & R_3 \end{cases} \qquad \begin{pmatrix} 1 & -2 & 0 & 2 & -3 \\ 2 & -4 & 2 & 0 & 8 \\ 1 & -2 & 3 & -3 & 16 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 8 \end{pmatrix}$$

$$\xrightarrow[-R_1+R_3]{-2R_1+R_2} \begin{cases} a_1 - 2a_2 + 0a_3 + 2a_4 - 3a_5 = 2 \\ 0a_1 + 0a_2 + 2a_3 - 4a_4 + 14a_5 = 2 \\ 0a_1 + 0a_2 + 3a_3 - 5a_4 - 3a_5 = 6 \end{cases} \xrightarrow[-\frac{3}{2}R_2+R_3]{\frac{1}{2}R_2} \begin{cases} a_1 - 2a_2 + 0a_3 + 2a_4 - 3a_5 = 2 \\ 0a_1 + 0a_2 + a_3 - 2a_4 + 7a_5 = 1 \\ 0a_1 + 0a_2 + 0a_3 + a_4 - 2a_5 = 3 \end{cases} \xrightarrow[R_2+2R_3]{R_1-2R_3} \begin{cases} a_1 - 2a_2 + 0a_3 + 0a_4 + a_5 = -4 \\ 0a_1 + 0a_2 + a_3 + 0a_4 + 3a_5 = 7 \\ 0a_1 + 0a_2 + 0a_3 + a_4 - 2a_5 = 3 \end{cases}$$

Elementry row operations "Gaussian elimination"

$$\left(\begin{array}{ccccc|c} 1 & -2 & 0 & 2 & -3 & 2 \\ 2 & -4 & 2 & 0 & 8 & 6 \\ 1 & -2 & 3 & -3 & 16 & 8 \end{array}\right) \xrightarrow[-\frac{3}{2}R_2+R_3]{\frac{1}{2}R_2} \left(\begin{array}{ccccc|c} 1 & -2 & 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & -2 & 7 & 1 \\ 0 & 0 & 0 & 1 & -2 & 3 \end{array}\right) \xrightarrow[R_2+2R_3]{R_1-2R_3} \left(\begin{array}{ccccc|c} 1 & -2 & 0 & 0 & 1 & -4 \\ 0 & 0 & 1 & 0 & 3 & 7 \\ 0 & 0 & 0 & 1 & -2 & 3 \end{array}\right)$$

Def. Let $S \subseteq V$, Then we define

$\mathrm{Span}(S) = \{$ all linear combinations of the vectors in $S\}$

(Rk: if $S = \emptyset$, $\mathrm{sign}(S) = \{0\}$

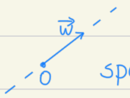
Prop: $\mathrm{Span}(S)$ is the smallest subspace of V that contains S

(i) Any subspace $\supseteq S$ must contain $\mathrm{span}(S)$

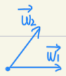
(ii) $\mathrm{Span}(S)$ is a subspace

Def: If $\mathrm{Span}(S) = V$, then we say that S spans V

Prop: $\mathrm{span}(S) =$ the smallest subspace of V that contains S



$\mathrm{span}\{\vec{w}\} = \{c\vec{w} \mid c \in F\}$



$\mathrm{span}(\{\vec{w_1}, \vec{w_2}\})$

Example: $\mathrm{Span}\{x^2+3x-2, 2x^2+5x-3, -x^2-4x+4\} = P_2(\mathbb{R})$?

Need to find $\alpha, \beta, \gamma \in \mathbb{R}$. s.t. $\alpha(x^2+3x-2) + \beta(2x^2+5x-3) + \gamma(-x^2-4x+4) = b_0 + b_1 x + b_2 x^2$

$(b_2, b_1, b_0) = \alpha(1,3,-2) + \beta(2,5,-3) + \gamma(-1,-4,4) = (\alpha+2\beta-\gamma, 3\alpha+5\beta-4\gamma, -2\alpha-3\beta+4\gamma)$

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & b_2 \\ 3 & 5 & -4 & b_1 \\ -2 & -3 & 4 & b_0 \end{array}\right) \xrightarrow[2R_1+R_3]{-3R_1+R_2} \left(\begin{array}{ccc|c} 1 & 2 & -1 & b_2 \\ 0 & -1 & -1 & -3b_2+b_1 \\ 0 & 1 & 2 & 2b_2+b_0 \end{array}\right) \xrightarrow[\substack{2R_2+R_1 \\ R_2+R_3}]{-R_2} \left(\begin{array}{ccc|c} 1 & 0 & -3 & -5b_2+b_1 \\ 0 & 1 & 1 & 3b_2-b_1 \\ 0 & 0 & 1 & -b_2+b_1+b_0 \end{array}\right) \xrightarrow[-R_3+R_2]{3R_3+R_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -8b_2+5b_1+3b_0 \\ 0 & 1 & 0 & 4b_2-2b_1-b_0 \\ 0 & 0 & 1 & -b_2+b_1+b_0 \end{array}\right)$$

so $\alpha = -8b_2+5b_1$ $\quad$ $\beta = 4b_2-2b_1-b_0$ $\quad$ $\gamma = -b_2+b_1+b_0$

Def: A set of (limitedly many) vectors

$\{u_1, u_2, \cdots, u_n\}$ is called linear dependent if $\exists\, a_1, \cdots, a_n \in F$, NOT ALL 0, s.t.

$a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = 0$

Otherwise $\{u_1, u_2, \cdots, u_n\}$ is linear independent

Prop(1) $S_1 \subseteq S_2 \subseteq V$. Then S_1 ℓ dep $\Rightarrow$ S_2 ℓ dep

S_2 ℓ indep $\Rightarrow$ S_1 ℓ indep

(2) Suppose $S \subseteq V$ is ℓ.indep. Then

$S \cup \{\vec{v}\}$ is ℓ.dep $\Leftrightarrow$ $\vec{v} \in \mathrm{span}(S)$

Def: A basis β for a vector space V is a ℓ.indep subset. s.t. $\mathrm{span}(\beta) = V$

A set $S \subseteq V$ is ℓ.dep if it contains a ℓ.dep finite subset

Example: $P =$ the set of all polynomial

$\{1, x, x^2, \cdots, x^n, \cdots\}$ and this is a basis for P

$\mathscr{C}(\mathbb{R}) =$ the set of all continuous function

Theorem

(1) $\beta = \{u_1, \cdots, u_n\} \subseteq V$ is a basis for V

$\Leftrightarrow$ every $v \in V$ can be uniquely expressed as a linear combination of vectors in β

where $v = a_1 u_1 + \cdots + a_n u_n$ for a unique set of scalar $a_1, \cdots, a_n$

(2) $V = \mathrm{span}(S)$ where S is a limited set $\Rightarrow$ some subset of S is a basis of V

proof: Since β spans V $\quad \Leftarrow \mathrm{Span}\,\beta = V$, $0 \in V \Rightarrow \exists a_1, a_2, \cdots, a_n$ s.t. $0 = 0 \cdot u_1 + 0 \cdot u_2 + \cdots + 0 \cdot u_n$

for $\forall v \in V$, $\exists a_1, \cdots, a_n \in F$, s.t. $v = a_1 u_1 + \cdots + a_n u_n$ $\quad 0 = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$

Suppose $\exists b_1, b_2, \cdots, b_n \in F$, s.t. $v = b_1 u_1 + \cdots + b_n u_n$

$0 = (a_1 - b_1) u_1 + \cdots + (a_n - b_n) u_n$, then $a_1 = b_1$, $a_2 = b_2$, $\cdots$, $a_n = b_n$

(3) $S = \{u_1, \cdots, u_n\}$ $\quad \mathrm{span}(S) = V$

Suppose S is ℓ indep. Then we've done

Otherwise S is ℓ.dep i.e. $\exists a_1, \cdots, a_n$. NOT all 0, s.t. $a_1 u_1 + \cdots + a_n u_n = 0$

Suppose $a_1 \neq 0$, $u_1 = -\frac{a_2}{a_1} u_2 - \cdots - \frac{a_n}{a_1} u_n \Rightarrow u_1 \in \mathrm{span}(\{u_2, \cdots, u_n\})$

$\Rightarrow \mathrm{span}\{u_1, \cdots, u_n\} \subseteq \mathrm{span}\{u_2, \cdots, u_n\}$

The Replacement Theorem

Let $G \subseteq V$ be a set of n vectors & $V = \mathrm{span}(G)$

Suppose $L \subseteq V$ is a ℓ.indep subset of m vectors

$m \leq n$ & $\exists$ a subset $H \subseteq G$ containing $(n-m)$ vectors & $\mathrm{span}(L \cup H) = V$

proof: Assume $G \in \{u_1, u_2, \cdots, u_n\}$ $\quad L = \{v_1, \cdots, v_m\}$

* if $L = \{v_1, \cdots, v_m\}$ is ℓ.indep & $m \leq n$

then $\exists$ a sub set $H \subseteq G$ containing $(n-m)$ vectors s.t. $V = \mathrm{span}(L \cup H)$

For $(*)$, we use induction

$m = 1$, $L = \{v_1\}$. L is ℓ.indep $\Rightarrow v_1 \neq 0$

& $v_1 = a_1 u_1 + \cdots + a_n u_n$, so $a_1, \cdots, a_n$ must NOT be all 0

Suppose WLOG $a_1 \neq 0 \Rightarrow a_1 u_1 = v_1 - a_2 u_2 - \cdots - a_n u_n$

$u_1 = \frac{1}{a_1} v_1 - \frac{a_2}{a_1} u_2 - \cdots - \frac{a_n}{a_1} u_n$, $u_1 \in \mathrm{span}(\{v_1, u_2, \cdots, u_n\})$, where $L = \{v_1\}$, $H = \{u_2, \cdots, u_n\}$

Induction hypothesis Suppose $(*)$ holds $m = k \leq n-1$

we claim $(*)$ holds for $m = k+1$

Let $L = \{v_1, \cdots, v_k, v_{k+1}\}$ ℓ.indep $\Rightarrow L' = \{v_1, \cdots, v_k\}$ ℓ.indep

Induction hypothesis $\Rightarrow$ $(*)$ holds for, i.e. $\exists (n-k)$ vectors

in $G = \{u_{k+1}, \cdots, u_n\}$, s.t. $V = \mathrm{span}(\{v_1, \cdots, v_k, u_{k+1}, \cdots, u_n\})$

in particular, $v_{k+1} = b_1 v_1 + \cdots + b_k v_k + a_{k+1} u_{k+1} + \cdots + a_n u_n$

as $a_{k+1}, \cdots, a_n$ cannot be all 0, we may assume $a_{k+1} \neq 0$

Then $u_{k+1} = \frac{1}{a_{k+1}} v_{k+1} - \frac{b_1}{a_{k+1}} v_1 - \cdots - \frac{b_k}{a_{k+1}} v_k - \frac{a_{k+2}}{a_{k+1}} u_{k+2} - \cdots - \frac{a_n}{a_{k+1}} u_n$

so $u_{k+1} \in \mathrm{span}\{v_1, \cdots, v_{k+1}, u_{k+2}, \cdots, u_n\}$

$G = \{u_1, \cdots, u_n\}$, $\mathrm{span}(G) = V$

$L = \{v_1, \cdots, v_m\}$, Need to show $m \leq n$

Suppose $m > n$, $L = \{v_1, \cdots, v_n, v_{n+1}, \cdots, v_m\}$

$\hat{L} = \{v_1, \cdots, v_n\}$

$(*) \Rightarrow \hat{L}$ spans V $\Rightarrow \{v_{n+1}, \cdots, v_m\} \subseteq \mathrm{span}\,\hat{L}$

so L cannot be a ℓ.indep subset of V

Cor. If V has a finite basis, Then all bases of V must have exactly the same # of vectors

Def. That # is the dimention of , $\dim V = n$

Generating set: if $\operatorname{span}(\chi) = V$, then χ is a generating set

Corollarios

(i) Suppose V has a finite basis. Then all bases have exactly same # of vectors

(ii) a. If $\dim(V) = n$, then any generating set must contain at least n vectors and any generating set of n vectors must be a basis of V

b. Any l.indep set contains at most n vectors and a l.indep set contains n vectors must be a basis

c. Any l.indep set can be extended to a basis

(iii) Let V be a finite-dim' l. vector space & $W \subseteq V$ is a subspace, then $\dim W \leq \dim V$

(a) $\dim W = \dim V \Rightarrow W = V$ (b) any basis of W can be extended to a basis of V

Def. Let S be a subset of a vector space V

A maximal l.indep subset of S in a subset $B \subseteq S$

(1) B is l.indep (2) The only l.indep subset of S that contains B is B

Prop: A max l.indep subset of V is a basis

Let S be a l.indep subset of V (any vector space), then $\exists$ max l. indep subset β of V that contains S

(proof) Let F = the family of all l. indep subsets of V that contains S

Def. Let F be a family of sets. A member $M \in F$ is called maximal (with respect to set inclusion)

if M is contained in no other members of F

Example F = the family of all subsets of the sets S & T. where $S \cap T = \emptyset$ S, T

Def. A family of sets C is called a chain if for each pair of sets A, B in C, we have either $A \subseteq B$ or $B \subseteq A$

Zorn's Lemma

Let F be a family of sets, If for each chain $C \subseteq F$, $\exists F \subseteq F_1$

that contain every member in C, then F_1 contains a maximal member

Claim F contains a maximal member. To apply Lorn's Lemma, we let $C \subseteq F$ be a chain

Let U = the union of all members in C

need to show that $U \in F$; i.e. U is a l.indep set

Suppose $u_1, u_2, \cdots, u_n \in U$ $a_1, \cdots, a_n \in \mathbb{R}$

s.t. $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$ for each u_i, $\exists C_i \in C$, s.t. $u_i \in C_i$, $i \in \{1, 2, \cdots, n\}$

Then $\exists C_k$, for some k in $1, \cdots, n$, s.t. $u_i \in C_k$ for $\forall i$

Some C_k is l.indep, we conclude: $a_i = 0$ $\forall i$

i.e. $\{u_1, \cdots, u_n\}$ l. indep and thus U is l.indep i.e. $U \in F$

Def: T is a linear transformation if $T: V \to W$ over F

$T(x+y) = T(x) + T(y)$

$T(cx) = c\,T(x)$

Def: $T: V \to W$ is linear

(i) The kernel of T is $N(T) = \{v \in V \mid T(v) = 0\}$
null space

(ii) The range of T is $R(T) = \{T(v) \mid v \in V\}$
images

Rk Both $N(T)$ & $R(T)$ are subspaces (check)

(iii) If both $N(T)$ & $R(T)$ are finite dimention

then $\text{nullity}(T) = \dim N(T)$

$\text{rank}(T) = \dim R(T)$

Dimention Theorem: If V is finite dimentional, then $\dim V = \text{nullity}(T) + \text{rank}(T)$

proof: $\dim V = n$. Since $N(T)$ is a subspace of V, we have $n \geq \dim N(T) = k$

Suppose $N(T)$ has a basis $\{v_1, v_2, \cdots, v_k\}$

Replacement Theorem $\Rightarrow \exists\, v_{k+1}, \cdots, v_n$ s.t.

$\{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ spans V, $\dim \{v_1, \cdots, v_n\} = n$, so it's a basis for V

Claim: $\{T(v_{k+1}), \cdots, T(v_n)\}$ is a basis for $R(T)$, $x \in V$

$T(x) = T(a_1 v_1 + \cdots + a_k v_k + \cdots + a_n v_n) = a_1 T(v_1) + a_2 T(v_2) + \cdots + a_k T(v_k) + a_{k+1} T(v_{k+1}) + \cdots + a_n T(v_n)$

$T(v_i) = 0,\ i = 1, 2, \cdots, k$, so $R(T) = \text{span}\{T(v_{k+1}), \cdots, T(v_n)\}$

Suppose $0 = \sum_{j=k+1}^{n} b_j T(v_j) = T\left(\sum_{j=k+1}^{n} b_j v_j\right) \Rightarrow \sum_{j=k+1}^{n} b_j v_j \in N(T)$

$\Rightarrow \sum_{j=k+1}^{n} b_j v_j \in N(T) = \text{span}\{v_1, \cdots, v_k\} \Rightarrow b_j = 0,\ j = k+1, \cdots, n$

Collary

(i) T is 1-1 $\Leftrightarrow N(T) = \{0\}$

(ii) If $\dim V = \dim W$, then the followings are equivalent

(a) T is 1-1

(b) T is onto

(c) $\text{rank}(T) = \dim V$

Prop. let $\beta = \{v_1, v_2, \cdots, v_n\}$ be a basis for V and $\{w_1, w_2, \cdots, w_n\}$ be an arbitary subset of n vectors in W. Then there exists a unique linear transformation $T: V \to W$ s.t.

$T(v_j) = w_j,\ j = 1, 2, \cdots, n$

$\mathbb{R}^n \longrightarrow \mathbb{R}^m$

$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \to$ j th component

$x = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_j e_j + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{j=1}^{n} x_j e_j$

$T(x) = \sum_{j=1}^{m} x_j T(e_j) = x_1 T(e_1) + \cdots + x_n T(e_n) = \underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{m\times n \text{ matrix}} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{n} a_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{pmatrix} = \begin{pmatrix} a_{11} x_1 + a_{12} x_1 + \cdots + a_{1n} x_n \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n \end{pmatrix}$

$T(x) = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$

Def. An ordered basis for a vector space V is a basis for V in a specific order

e.g. $\{e_1, e_2, e_3\}$ & $\{e_3, e_2, e_1\}$ are the same bases for V but different as ordered basis

Def. Let $\beta = \{u_1, u_2, \cdots, u_n\}$ be an ordered basis for V

Then for $v \in V$, $\exists\, b_1, \cdots, b_n \in F$ such that

$v = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n$ & we define the coordinate vector of x relative to β as

$[x]_\beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$

Let $\beta = \{v_1, \cdots, v_n\}$ be an ordered basis for V & $\gamma = \{w_1, \cdots, w_n\}$ be an ordered basis for W

& $T: V \to W$ is a linear transformation s.t. $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$, $j = 1, 2, \cdots, n$

then $[T(v_j)]_\gamma = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{mj} \end{pmatrix}$

Def. The $m\times n$ matrix $A(a_{ij})_{m\times n}$ is the matrix representation of T in the ordered basis β & γ, we write $A = [T]_\beta^\gamma$

when $V = W$, $\beta = \gamma$, $[T]_\beta^\gamma = [T]_\beta$

Therefore $[T(x)]_\gamma = [T]_\beta^\gamma [x]_\beta$, $\forall x \in V$

$[T(x)]_\gamma = [T(b_1 v_1) + T(b_2 v_2) + \cdots + T(b_n v_n)]_\gamma = [b_1 T(v_1)]_\gamma + [b_2 T(v_2)]_\gamma + \cdots + [b_n T(v_n)]_\gamma$

$= \begin{pmatrix} b_1 a_{11} + b_2 a_{12} + \cdots + b_n a_{1n} \\ b_1 a_{21} + b_2 a_{22} + \cdots + b_n a_{2n} \\ \cdots \\ b_1 a_{m1} + b_2 a_{m2} + \cdots + b_n a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = [T]_\beta^\gamma$

$[T]_\beta^\gamma = \begin{pmatrix} | & | & & | \\ [T(x_1)]_\gamma & [T(x_2)]_\gamma & \cdots & [T(x_n)]_\gamma \\ | & | & & | \end{pmatrix}$

$A \cdot B = C \to C_{ij} = \sum A_i B_j$

Example $T: P_2 \longrightarrow M_{2\times 2}(\mathbb{R})$

$\beta = \{1, x, x^2\}$ $\quad \gamma = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$

$T(f(x)) = \begin{pmatrix} f'(0) & 2f(1) \\ 0 & f''(3) \end{pmatrix}$

$$[T]_\beta^\gamma = \begin{pmatrix} | & | & | \\ [T(1)]_\gamma & [T(x)]_\gamma & [T(x^2)]_\gamma \\ | & | & | \end{pmatrix}$$

$T(1) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. $[T(1)]_\gamma = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ $\quad [T(x)]_\gamma = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ $\quad T(x^2) = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix}$

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$V \to W$ $\quad L(V, W) =$ the set of all linear transformation from V to W

Observations:

① $L(V, W)$ is a vector space

$(T_1 + T_2)(v) = T_1(v) + T_2(v) \quad (cT)(v) = cT(v)$

② If $\dim V = n$ $\quad \dim W = m \Rightarrow \dim L(V, W) = mn$

$T \longleftrightarrow [T]_\beta^\gamma : m \times n$

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{\beta, n}"'] & W \arrow[d, "{\gamma, m}"] \\ {[v]_\beta} \arrow[r, "{[T]_\beta^\gamma}"] & {[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta} \end{tikzcd}$$

$L(V, W) \sim M_{m\times n}(\mathbb{R})$

Composition of linear transformation / matrix multiplication

$V \xrightarrow{\quad T \quad} W \xrightarrow{\quad U \quad} Z \quad \Rightarrow \quad \underset{\alpha}{V} \xrightarrow{\ UT\ } \underset{\gamma}{Z}$

$\alpha = \{v_1, \cdots, v_n\}$ $\quad \beta = \{w_1, \cdots, w_m\}$ $\quad \gamma = \{z_1, \cdots, z_p\}$ $\qquad [UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$

$B = [T]_\alpha^\beta$ $m\times n$ $\quad A = [U]_\beta^\gamma$ $p \times m$ $\qquad p\times n \quad p \times m \quad m \times n$

(pf) $(UT)(v_j) = U(T(v_j)) = U\left(\sum_{i=1}^{m} b_{ij} w_i\right) = \sum_{i=1}^{m} b_{ij} U(w_i) = \sum_{i=1}^{m} b_{ij} \sum_{k=1}^{p} a_{ki} z_k = \sum_{i=1}^{m}\sum_{k=1}^{p} b_{ij} a_{ki} z_k = \sum_{k=1}^{p}\left[\sum_{k=1}^{m} (b_{ij} a_{ki})\right] z_k$

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

Def. Let $T: V \to W$ be a linear transformation

(1) A map $U: W \to V$ is an inverse of T if $TU = I_W$, $UT = I_V$

(2) If T has an inverse, we say T is invertible & denote its inverse as T^{-1}

Example $T: P_2 \to P_3$ $\quad T(f(x)) = \int_0^x f(t)\,dt$ $\qquad \alpha = \{1, x, x^2\}$ $\quad \beta = \{1, x, x^2, x^3\}$

$U: P_3 \to P_2$ $\quad T(f(x)) = f'(x)$

$T(1) = \int_0^x dt = x$ $\qquad T(x) = \int_0^x t\,dt = \frac{1}{2}x^2$ $\qquad T(x^2) = \int_0^x t^2 dt = \frac{1}{3}x^3$

$[T(1)]_\beta = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ $\quad T(x) = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}$ $\quad T(x^2) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix}$

$TU: P_3 \to P_3$ $\quad T(U(f(x))) = T(f'(x)) = \int_0^x f'(t)\,dt = f(x) - f(0)$

$UT: P_2 \to P_2$ $\quad U(T(f(x))) = U(\int_0^x f(t)dt) = f(x) = I_{P_2}(f(x))$

so $[T]_\alpha^\beta = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$ $\qquad [U]_\beta^\alpha = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

$[UT]_\alpha = [U]_\beta^\alpha \cdot [T]_\alpha^\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_{P_2}$

$[TU]_\beta = [T]_\alpha^\beta \cdot [U]_\beta^\alpha = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \neq I_{P_3}$

Def. $A: n \times n$, A is invertible if $\exists B$ $n \times n$ matrix s.t. $AB = BA = I$

Rk. Only need one

Prop $T: V \to W$ a linear transformation

(1) T is invertible $\iff$ T is one-to-one and onto

(2) If T is invertible, T^{-1} is also invertible

(3) If $T_1: V \to W$ and $T_2: W \to Z$ are both invertible linear transformations.
Then $T_2 T_1: V \to Z$ is also invertible and $(T_2 T_1)^{-1} = T_1^{-1} T_2^{-1}$

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

(5) If V or W are finite dimentional and $T: V \to W$ is invertible,
then $\dim V = \dim W$

Theorem Suppose V and W are finite dimentional and $T: V \to W$ is a linear transformation
Then T is invertible $\iff$ $[T]_\beta^\gamma$ is invertible where β is an ordered basis for V
and γ is an ordered basis for W, $([T]_\beta^\gamma)^{-1} = [T^{-1}]_\gamma^\beta$

Proof: $TT^{-1} = I_W$

$[TT^{-1}]_\gamma = [I_W]_\gamma = [T]_\beta^\gamma [T^{-1}]_\gamma^\beta$

$T^{-1}T = I_V \Rightarrow [T^{-1}T]_\beta = [I_V]_\beta = [T^{-1}]_\gamma^\beta \cdot [T]_\beta^\gamma = I_n$

so $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$ $\qquad (n = \dim V)$

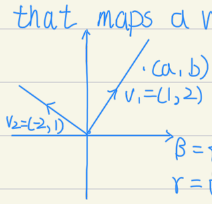
(<=) $A=[T]_\beta^\gamma$ is invertible (as a matrix) for some $\beta=\{v_1,\cdots,v_n\}$, $\gamma=\{w_1,\cdots,w_n\}$

i.e. $\exists$ B, a matrix, s.t. $AB=BA=I_n$ $\quad B=(b_{ij})_{n\times n}=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{bmatrix}$

Define $U: W\to V$ as follow, $W(w_j)=b_{1j}v_1+\cdots+b_{nj}v_n=\sum_{i=1}^{n} b_{ij}v_i$

$UT=I: V\to V$, $[UT]_\beta=[U]_\gamma^\beta[T]_\beta^\gamma=BA=I_n$

Example Let T be the linear transformation that maps a vector in $\mathbb{R}^2$ to its reflection



$\beta=\{(1,0),(0,1)\}$
$\gamma=\{(1,2),(-2,1)\}$

$\binom{a}{b}=[v]_\beta$

$[I(v)]_\gamma=[I]_\beta^\gamma\cdot[v]_\beta=[I]_\beta^\gamma\cdot\begin{bmatrix}a\\b\end{bmatrix}$

$[I]_\beta^\gamma=([I]_\gamma^\beta)^{-1}=\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}^{-1}$

Theorem. Let V, W be two finite dimentional vector spaces. Then $\exists$ an one-to-one onto linear transformation $T: V\to W$ <=> $\dim V=\dim W$

proof: ① $\dim V=\dim N(T)+\dim R(T)$

one-to-one => $N(T)=\{0\}$ onto => $R(T)=W$. so $\dim V=\dim W$

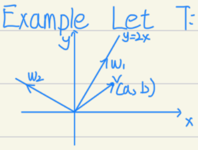
② let $\beta=\{v_1,v_2,\cdots,v_n\}$ be a basis for V, $\gamma=\{w_1,w_2,\cdots,w_m\}$ be a basis for W

Define $T: V\to W$ by $T(v_j)=w_j$, $j=1,2,\cdots,n$ and T is linear

Clearly T is one-to-one and onto

Def: We say that V and W are isomorphic if $\exists$ a linear transformation $T: V\to W$ which is one-to-one and onto

Example Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be the linear transformation which maps a vector to its reflection w.r.t. the line $y=2x$, find $T(a,b)$



$\beta=\{e_1,e_2\}$
$\gamma=\{(1,2),(-2,1)\}$ $\quad [T(v)]_\beta=[T]_\beta[v]_\beta$

$[T(v)]_\gamma=[T]_\gamma[v]_\gamma$

$[v]_\beta=[I(v)]_\gamma^\beta[v]_\gamma=\begin{bmatrix}1 & -2\\ 2 & 1\end{bmatrix}[v]_\beta$

$\boxed{\begin{matrix}[T]_\gamma^\beta=[TI]_\gamma^\beta=[T]_\beta[I]_\gamma^\beta\\ [T]_\gamma^\beta=[IT]_\gamma^\beta=[I]_\gamma^\beta[T]_\gamma\end{matrix}}$ $\Rightarrow$ $[T]_\gamma=([I]_\gamma^\beta)^{-1}[T]_\beta[I]_\gamma^\beta$ $\qquad [T]_\gamma=\begin{bmatrix}[T(w_1)]_\gamma & [T(w_2)]_\gamma\end{bmatrix}=\begin{bmatrix}[w_1]_\gamma & [-w_2]_\gamma\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$

so $[T]_\gamma=\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}^{-1}[T]_\beta\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}$

$[T]_\beta=\begin{bmatrix}1 & -2\\ 2 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & -2\\ 2 & 1\end{bmatrix}^{-1}=\begin{bmatrix}1 & -2\\ 2 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 2\\ -2 & 1\end{bmatrix}\frac{1}{5}$

the change of coordinate

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = (ad-bc)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \det(A) \cdot I$$

Theorem

(i) let V be a finite dimentional vector space with two ordered basis β. γ. Then

(i) $Q = [I]_\gamma^\beta$ is invertible

(ii) $[v]_\beta = Q[v]_\gamma = [I]_\gamma^\beta [v]_\gamma = [I(v)]_\beta$ (Change of coordinates)

Ougmented Matrix

$$\overset{A \qquad I}{\left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right)} \xrightarrow[R_2-2R_1]{R_1+2R_2} \left(\begin{array}{cc|cc} 5 & 0 & 1 & 2 \\ 0 & 5 & -2 & 1 \end{array}\right) \xrightarrow[\frac{1}{5}R_2]{\frac{1}{5}R_1} \overset{I \qquad A^{-1}}{\left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{5} & \frac{2}{5} \\ 0 & 1 & -\frac{2}{5} & \frac{1}{5} \end{array}\right)}$$

Elementary matrix operation for A

Def. Elementary row operations are

1. interchanging two rows of A
2. multiplying any row with a nonzero number
3. adding a scalar multiple of a row of A

$$\overset{I}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \xrightarrow{\text{type1}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_1 \Leftrightarrow R_2} \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[C \neq 0 \; CR_3]{\text{type2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & C \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{CR_3} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7C & 8C & 9C \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_2+CR_3]{\text{type3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & C \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_2+CR_3} \begin{bmatrix} 1 & 2 & 3 \\ 4+7C & 5+8C & 6+9C \\ 7 & 8 & 9 \end{bmatrix}$$

elementary matrices

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & C \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7C & 8C & 9C \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & C \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4+7C & 5+8C & 6+9C \\ 7 & 8 & 9 \end{bmatrix}$$

As shown, for any invertible matrices, the matrix operation is a multiplication with a elementary matrix so for any invertible matrix A, it can be transformed into I by finite matrix multiplication

$E_1E_2\cdots E_kA = I$. $E_1E_2\cdots E_k = A^{-1}$

so $E_1E_2\cdots E_k[A|I] = [I|A^{-1}]$

Theorem Let $A \in M_{m\times n}$. Suppose that B is obtained from A by an elementary row/column operation. Then $\exists$ $m\times m$/$n\times n$ elementary matrix E s.t. $B = EA/AE$

Def The rank of $A \in M_{m\times n}(F)$ is the rank of $L_A: F^n \to F^m$

$T: \overset{\beta}{V} \to \overset{\gamma}{W}$

$\mathrm{rank}(T) = \mathrm{rank}([T]_\beta^\gamma)$

Prop (1) $A \in M_{m\times n}(F)$, P, Q are invertible, where $P \in M_{m\times m}$ $Q \in M_{n\times n}$

$\Rightarrow \mathrm{rank}(PA) = \mathrm{rank}(A)$; $\mathrm{rank}(AQ) = \mathrm{rank}(A)$

(In particular, elementary row/column operations are rank-preserving)

(2) rank(A) is the maximal number of linearly independent column vectors

(proof) $\mathrm{rank}(A) = \mathrm{rank}(L_A) = \dim[L_A(F^n)]$

Let β = the standard ordered basis for F^n

$R(L_A) = \mathrm{span}\{L_A(e_1), L_A(e_2), \cdots, L_A(e_n)\} = \mathrm{span}\{Ae_1, Ae_2, \cdots, Ae_n\}$

Theorem: Let $A \in M_{m\times n}(F)$ of rank r

Then $r \le m$ and $r \le n$. Moreover by a finite number of elementary operations (row/column)

A can be transformed into $D = \left[\begin{array}{c|c} I_r & O_1 \\ \hline O_2 & O_3 \end{array}\right]_{m\times n}$

proof: If $A = O$, ✓ $A \ne O \Rightarrow r > 0$

To use induction on m

$m=1$, $A = (a_1, \cdots, a_n) \longrightarrow (1, a_2', \cdots, a_n') \longrightarrow (1, 0, 0, \cdots, 0)$

Induction Hypothesis. The statement holds for any matrix A with $(m-1)$ vectors

Now, let $A \in M_{m\times n}(F) \ne O$, $A \to \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ 0 & & B & \\ \vdots & & & \\ 0 & & & \end{array}\right]$ $\quad B = M_{(m-1)\times(n-1)}(F)$

$\mathrm{rank}(A) = \mathrm{rank}(B) + 1 \Rightarrow \mathrm{rank}\,B = r-1$

so $B = \left[\begin{array}{c|c} I_{r-1} & O \\ \hline O & O \end{array}\right]$

Corollary 1. let $A \in M_{m\times n}(F)$ of rank r

Then $\exists$ an invertible matrix $B \in M_{m\times m}(F)$ and an invertible matrix $C \in M_{n\times n}(F)$ s.t. $BAC = D$

2. $\mathrm{rank}(A) = \mathrm{rant}(A^t)$

proof: $D = BAC \Rightarrow D^t = (BAC)^t = C^t A^t B^t$

B, C invertible $\Rightarrow B^t, C^t$ invertible

$D^t = \left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right]$, so $\mathrm{rank}(D^t) = \mathrm{rank}(C^t A^t B^t) = \mathrm{rank}(A^t) = \mathrm{rank}(A)$

3. rank(A) = the maximal number of linearly independent row vectors

4. Every invertible matrix is a product of elementry matrices

$I = D = E_p \cdots E_2 E_1 A F_1 F_2 \cdots F_q$, $(E_p \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_p^{-1}$

$(E_p \cdots E_1)^{-1} I (F_1 \cdots F_q)^{-1} = A$

Theorem (1) $V \xrightarrow{T} W \xrightarrow{U} Z$

V, W, Z are all finite-dimentional vector spaces, T, U are linear transformations

then $rank(UT) \leq \min\{rank(U), rank(T)\}$

(2) A, B two matrices s.t. AB is defined

Then $rank(AB) \leq \min\{rank(A), rank(B)\}$

proof: $rank(UT) = \dim R(UT) = \dim UT(V)$, where $T(V) \subseteq W$.

so $rank(UT) \leq \dim U(W) = \dim(U)$

$rank(AB) = rank(L_{AB}) = rank(L_A L_B) \leq rank(L_A) = rank(A)$

$rank(AB) = rank(AB)^t = rank(B^t A^t) \leq rank(B^t) = rank(B)$

Rigorous to show that $rank(UT) \leq rank(T)$: $\overset{\alpha}{V} \xrightarrow{T} \overset{\beta}{W} \xrightarrow{U} \overset{\gamma}{Z}$

$A' = [U]_\beta^\gamma$, $B' = [T]_\alpha^\beta$

$A'B' = [UT]_\alpha^\gamma$, $rank(UT) = rank(A'B') \leq rank(B') = rank(T)$

Example: Prove directly that $rank(UT) \leq rank(T)$

$rank(UT) = \dim R(UT) = \dim R(U(R(T))) + \dim R(U(N(T)))$

Theorem $A \in M_{n\times n}(F)$ is invertible

$$A^{-1}[A \mid I] \qquad [I \mid B]$$
$$\| \qquad\qquad \|$$
$$[A^{-1}A \mid A^{-1}I] = [I \mid A^{-1}]$$

Systems of linear equasions

Def $Ax=b$, magnation in n unknowns

(1) the system is homogeneous if $b=0$

(2) nonhomogeneous otherwise

Prop. The set of all sets of $Ax=0$. $K_h = N(L_A)$, $\dim K_h = n - \operatorname{rank}(A)$

$L_A : F^n \longrightarrow F^m$. by dimention theorem, $\dim(F^n) = \dim N(L_A) + \dim R(L_A)$

Cov. $m \times n \Rightarrow Ax=0$ has a nonzero set

let $K=$ the set of all sets to $Ax=b$. Then $K = K_h + \{s\}$. where $s \in K$ is any set to $Ax=b$

Def A matrix A is said to be in "reduced row echelon form" if the following holds

(1) Any row containing a nonzero entry precades any row in which all entries are 0

(2) The first nonezero entry in each row is the only nonzero entry in its column

(3) The first nonzero entry in each row is 1 and it occurs in a column to the right of the first nonzero entry in the preceding row.

Gaussian elimination

The procedure of reducing an augmented matrix to its reduced row echelon form.

Theorem Let A be a $m \times n$ matrix of rank r and B be a RREF of A

(1) The number of nonzero rows in B is r

(2) For each $i = 1, \cdots, r$, $\exists$ a column b_{j_i} of B s.t. $b_{j_i} = e_i$

(3) Let a_{j_i} be the corresbonding column of $A \Rightarrow \{a_{j_1}, \cdots a_{j_r}\}$ is linearly independent

(4) If the column of B, $b_k = d_1 e_1 + \cdots + d_r e_r$, then the corresbonding column k of A, $a_k = d_1 a_{j_1} + \cdots + d_r a_{j_r}$

(5) B is unique (Picks up the first r linearly independent column vectors of A)

Key : $\exists$ an invertible matrix C s.t. $CA=B$ $(\Rightarrow C\underline{a_i} = \underline{b_i})$ column vectors

$\forall i = 1, 2, \cdots, n$

(pf.) (1) $CA=B$ & C is invertible

$\Rightarrow \operatorname{rank}(A) = \operatorname{rank}(CA) = \operatorname{rank}(B) = r$

(2) By the definition

(3), (4) $C(d_1 a_{j_1} + d_2 a_{j_2} + \cdots + d_r a_{j_r}) = d_1 C a_{j_1} + \cdots + d_r C a_{j_r} = d_1 b_{j_1} + \cdots + d_r b_{j_r}$

Example Let $V = \text{span}(S)$. $S = \{2+x+2x^2+3x^3, 4+2x+4x^2+6x^3, 6+3x+8x^2+7x^3, 2+x+5x^3, 4+x+9x^3\}$

Find a basis for V which is a subset of S

$P_3(\mathbb{R})$ $\{1, x, x^2, x^3\}$

$$\begin{bmatrix} 2 & 4 & 6 & 2 & 4 \\ 1 & 2 & 3 & 1 & 1 \\ 2 & 4 & 8 & 0 & 0 \\ 3 & 6 & 7 & 5 & 9 \end{bmatrix}$$

Example Let $V = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid x_1+7x_2+5x_3-4x_4+2x_5=0\}$

(1) Verify: V is a vector space

(2) Show: $S = \{(2,0,0,1,1), (1,1,-2,-1,-1), (-5,1,0,1,1)\} \subseteq V$ and linearly independent

(3) Extend S to a basis for V

$x_1 = -7x_2 - 5x_3 + 4x_4 - 2x_5$

$(-7, 1, 0, 0, 0)$

$(-5, 0, 1, 0, 0)$

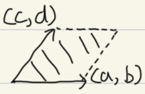
$(4, 0, 0, 1, 0)$

$(-2, 0, 0, 0, 1)$

$$\begin{pmatrix} 2 & 1 & -7 & 5 & 4 & -2 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Determinant

$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad-bc$



$= |ad-bc|$

Determinant of $A \in M_{n\times n}$

$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ $\quad |A| = \det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$

(Eliminate the i-th row and the j-th column, then you get A_{ij})

Prop (1) ☆ $\det(A)$ is a linear function of each row when the remaining rows are held fixed

(2) If A has a row consisting entirely of zeros $\Rightarrow \det(A)=0$

(3) $\det(A)$ can be evaluated by cofactor expansion along any row

(4) $\det(A)=0$ if A has two identical rows

(5) ☆ If B is obtained from A by any interchange of two rows, then $\det(B) = -\det(A)$

(6) ☆ If B is adding a multiple of one row of A to another, then $\det(B)=\det(A)$

(7) $\operatorname{rank}(A) < n \Rightarrow \det(A)=0$

(8) $\det(AB) = \det(A)\cdot\det(B)$

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

(10) $\det(A) = \det(A^t)$

(1) $A = \begin{bmatrix} -a_1- \\ -a_2- \\ \cdots \\ -u+kv- \\ -a_n- \end{bmatrix}$ ($a_i \leftarrow$) $\quad B = \begin{bmatrix} -a_1- \\ -a_2- \\ \cdots \\ -u- \\ \cdots \\ -a_n- \end{bmatrix} \quad C = \begin{bmatrix} -a_1- \\ -a_2- \\ \cdots \\ -v- \\ -a_n- \end{bmatrix}$

$\det(A) = \det(B) + k\det(C)$

proof: $i=1$, obviously true

$i>1$: $|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| = \sum_{j=1}^{n} (-1)^{i+j} |A_{ij}| (b_{ij} + k c_{ij})$

$= \sum_{j=1}^{n} (-1)^{i+j} b_{ij} |B_{ij}| + k \sum_{j=1}^{n} (-1)^{i+j} c_{ij} |C_{ij}| = |B| + k|C|$

(2) $A = \begin{bmatrix} -a_1- \\ \cdots \\ -0- \\ \cdots \\ -a_n- \end{bmatrix} = \begin{bmatrix} -a_1- \\ \cdots \\ -0+0- \\ \cdots \\ -a_n- \end{bmatrix}$

so $|A| = \begin{vmatrix} -a_1- \\ \cdots \\ -0- \\ \cdots \\ -a_n- \end{vmatrix} + \begin{vmatrix} -a_1- \\ \cdots \\ -0- \\ \cdots \\ -a_n- \end{vmatrix} = |A| + |A|$, so $|A| = 0$

(4) when $n=2$, true

when $n>2$, $\exists\ 1 \le i \le n$, $a_{ij} = 0$,

then the determinant is zero from i-th row

so $|A| = 0+0+\cdots+0 = 0$

(5) $\underset{A}{\begin{vmatrix} -a_1- \\ \cdots \\ -a_r- \\ \cdots \\ -a_s- \\ \cdots \\ -a_n- \end{vmatrix}} + \underset{B}{\begin{vmatrix} -a_1- \\ \cdots \\ -a_s- \\ \cdots \\ -a_r- \\ \cdots \\ -a_n- \end{vmatrix}} = \begin{vmatrix} -a_1- \\ \cdots \\ -a_r+a_s- \\ \cdots \\ -a_r+a_s- \\ \cdots \\ -a_n- \end{vmatrix} = 0$

so $|A| = -|B|$

(6) $\begin{vmatrix} -a_1- \\ \cdots \\ -a_i+ta_j- \\ \cdots \\ -a_n- \end{vmatrix} = \begin{vmatrix} -a_1- \\ \cdots \\ -a_i- \\ \cdots \\ -a_n- \end{vmatrix} + \overset{\nearrow 0}{\begin{vmatrix} -a_1- \\ \cdots \\ -a_j- \\ \cdots \\ -a_n- \end{vmatrix}} = \begin{vmatrix} -a_1- \\ \cdots \\ -a_i- \\ \cdots \\ -a_n- \end{vmatrix}$

so $|A| = |B|$

(7) $\operatorname{rank}(A) < n \Rightarrow c_1a_1 + \cdots + c_na_n = 0$ Not All c_i's are 0

WLOG. $c_n \neq 0$, $a_n = -\frac{c_1}{c_n}a_1 - \frac{c_2}{c_n}a_2 - \cdots - \frac{c_{n-1}}{c_n}a_{n-1}$

so $|A| = \begin{vmatrix} -a_1- \\ \vdots \\ -a_{n-1}- \\ -\frac{c_1}{c_n}a_1 - \cdots - \frac{c_{n-1}}{c_n}a_{n-1} \end{vmatrix} = 0$

(8) $\operatorname{rank}(A) < n \Rightarrow \operatorname{rank}(AB) < n \Rightarrow |AB| = |A|\cdot|B| = 0$

$\operatorname{rank}(A) = n \Rightarrow A = E_mE_{m-1}\cdots E_1$

$\det(EB) = \det(E)\cdot\det(B)$, true

so $|AB| = |E_m \cdots E_1 B| = |E_m||E_{m-1}|\cdots|E_1||B| = |A||B|$

(9) $AA^{-1} = I$, so $\det(AA^{-1}) = 1 = \det(A)\cdot\det(A^{-1})$

(10) $\operatorname{rank}(A) = n \Rightarrow A = E_m \cdots E_1 \Rightarrow A^t = (E_m \cdots E_1)^t$

$A^t = E_1^t E_2^t \cdots E_m^t$, $\det(A^t) = |E_1^t||E_2^t|\cdots|E_m^t|$

$= |E_1||E_2|\cdots|E_m| = |E_m \cdots E_2E_1| = |A|$, so $|A^t| = |A|$

Cofactor expansion along row $i>1$

$$|A| = \begin{vmatrix} - & a_1 & - \\ & \cdots & \\ a_{i1}e_1 + & \cdots & + a_{in}e_n \\ & \cdots & \\ - & a_n & - \end{vmatrix} = a_{i1}\begin{vmatrix} - a_1 - \\ \cdots \\ - e_1 - \\ \cdots \\ - a_n - \end{vmatrix} + \cdots + a_{in}\begin{vmatrix} - a_1 - \\ \cdots \\ - e_n - \\ \cdots \\ - a_n - \end{vmatrix}$$

$$B = \begin{bmatrix} - a_1 - \\ \cdots \\ - e_k - \\ \cdots \\ - a_n - \end{bmatrix} \searrow \text{the } i\text{-th row} \qquad \underset{(n-1)\times(n-1)}{B_{ik}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \cdots & & \cdots & 0 \\ \vdots & & & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & \cdots & a_{nn} \end{bmatrix}$$

Let C_{ij} be the $(n-2)\times(n-2)$ matrix obtained from B by eliminating rows $1 \& i$, column $j \& k$

$$\underset{(n-2)\times(n-2)}{C_{ij}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & & \cdots & a_{2n} \\ \vdots & \vdots & & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & & \cdots & a_{nn} \end{bmatrix}$$

$$|B| = \sum_{j=1}^{n} (-1)^{1+j} a_{ij} |B_{ij}| \qquad (i-1)\text{-th row in } B_{ij} \text{ is } \begin{cases} e_{k-1}, & j<k \\ 0, & j=k \\ e_k, & j>k \end{cases}$$

By induction hyposis $|B_{ij}| = \begin{cases} (-1)^{(i-1)+(k-1)} |C_{ij}|, & j<k \\ 0, & j=k \\ (-1)^{(i-1)+k} |C_{ij}|, & j>k \end{cases}$

Prop. The determinant of an uppertriangular matrix is the product of digonal entries

proof: $n=2$ $\begin{vmatrix} a & b \\ 0 & c \end{vmatrix} = ac$ ✓

Induction Hypothesis. let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$, $\det(A) = \prod_{i=1}^{n} a_{ii}$

let $B = \left[\begin{array}{c|c} A & \begin{matrix} b_1 \\ \vdots \\ b_n \end{matrix} \\ \hline 0 & b_{n+1} \end{array}\right]$, so $\det(B) = (-1)^{2n+2} b_{n+1} \det(A) = a_{11}a_{22}\cdots a_{nn}b_{n+1}$

Def. A linear transformation $T: V \to V$ where V is a finite dimentional vector space, is diagonalizable if $\exists$ ordered basis β for V s.t. T is a diagonal matrix

$\underset{n\times n}{A}$ $L_A: \mathbb{R}^n \to \mathbb{R}^n$ $A = Q^{-1}DQ$

Def. If $T(v) = \lambda v$ (i.e. $Av = \lambda v$), then we say that v is an <u>eigenvector</u> (proper vector) and λ an <u>eigenvalue</u> (proper value)

If T (or A) is diagonalizable, then V has an ordered basis β which consists of n eigenvectors of T (or A)

$\beta = \{v_1, \cdots, v_n\}$, $T(v_1) = \lambda_1 v_1, \cdots, T(v_n) = \lambda_n v_n$, so $[T]_\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$

<u>Theorem</u>. $\lambda \in F$ is an eigenvalue of $T \iff \det(A - \lambda I) = 0$, where $A = [T]_\beta$

proof: $\lambda \in F$ is an eigenvalue of $T \iff \exists\, v \neq 0$, s.t. $T(v) = \lambda v$

i.e. $(T - \lambda I)(v) = 0$, which means $N(T - \lambda I) \neq \{0\}$

$\iff R(T - \lambda I)$ has a dimention less than n, so $\operatorname{rank}(A - \lambda I) < n$

so $\det(A - \lambda I) = 0$

<u>Def.</u> $|A - \lambda I| = 0$ is called the characteristic polynomial of A ($|[T]_\beta - \lambda I| = 0$ is called the ~ of T)

$0 = |A - \lambda I| = |Q^{-1}BQ - \lambda I| = |Q^{-1}BQ - Q^{-1}\lambda I Q| = |Q^{-1}(B - \lambda I)Q| = |Q^{-1}||B - \lambda I||Q| = |B - \lambda I|$

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

$0 = |A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 0 & 2-\lambda \end{vmatrix} = (2-\lambda)^2$, $\lambda = 2$

$0 = (A - \lambda I)\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$

so $b = 0$, $\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. A has only one eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

so A is not digonalizable

Theorem. Let $A \in m_{n\times n}$. If A has n distinct eigenvalue, then A is digonalizable

proof: If λ_1 is an eigenvalue of $A \Rightarrow \exists$ at least one linearly independent eigenvectors of A corresponding to λ_1

$\lambda_1, \lambda_2, \cdots, \lambda_n$ all distinct $\Rightarrow v_1, v_2, \cdots, v_n$ are linearly independent

$T: V \to V$, $\lambda \in F$ $(\mathbb{R}, \mathbb{C})$ is an eigenvalue if $\exists\, 0 \neq v \in V$ s.t. $T(v) = \lambda v$

If $\exists$ ordered basis $\beta = \{v_1, v_2, \cdots, v_n\}$ all eigenvectors for T, then T is digonalizable, $[T]_\beta$ is a diagonal matrix

Prop. If $\lambda_1, \cdots, \lambda_k$ are eigenvalues for T with $v_1, \cdots, v_k$ being the corresponding eigenvectors then $\{v_1, v_2, \cdots, v_k\}$ is linearly independent

pf. Suppose $a_1 v_1 + \cdots + a_k v_k = 0$

Induction, when $k=1$, $a_1 v_1 = 0$, so $a_1 = 0$

Hypothesis: Assume the conclusion holds for $k=m$

$a_1 v_1 + \cdots + a_m v_m + a_{m+1} v_{m+1} = 0$, $T(a_1 v_1 + \cdots + a_{m+1} v_{m+1}) = 0$

so $a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_m \lambda_m v_m + a_{m+1} \lambda_{m+1} v_m = 0$

$a_1 \lambda_{m+1} v_1 + \cdots + a_m \lambda_{m+1} v_m + a_{m+1} \lambda_{m+1} v_{m+1} = 0$

$a_1 (\lambda_1 - \lambda_{m+1}) v_1 + \cdots + a_m (\lambda_m - \lambda_{m+1}) v_m = 0$

Since $\lambda_1, \lambda_2, \cdots, \lambda_{m+1}$ are distinct, then $a_1 = a_2 = \cdots = a_m = 0$, also $a_{m+1} = 0$

so $\{v_1, v_2, \cdots, v_n\}$ is linearly independent for any $n \in \mathbb{N}^+$

If A is diagonalizable i.e. $\exists$ ordered basis of eigenvectors

thus $\exists Q$ s.t. $A = Q^{-1} D Q$, where D is a diagonal matrix

$|A - \lambda I| = |Q^{-1} D Q - Q^{-1}(\lambda I) Q| = |Q^{-1}|\,|D - \lambda I|\,|Q| = |D - \lambda I|$

$$0 = |A - \lambda I| = |D - \lambda I| = \begin{vmatrix} d_1 - \lambda & & O \\ & \ddots & \\ O & & d_n - \lambda \end{vmatrix} = (d_1 - \lambda)(d_2 - \lambda)\cdots(d_n - \lambda)$$

the eigen-polynomial of A can be factored in to a product of linear factors

i.e. $|A - \lambda I|$ splits, So this is a necessary condition

Def. The ( ) multiplicity of an eigenvalue λ of T (or A) is the largest positive integer k s.t. $(t-\lambda)^k$ is a factor of the characteristic polynomial $|A - \lambda I| = 0$

Def. If λ is an eigenvalue of T (or A), then $E_\lambda = \{v \in V \mid T(v) = \lambda v\}$ is called the eigenspace corresponding to λ

$(= N(T - \lambda I))$

Prop. $\dim E_\lambda \leq m =$ the multiplicity of λ

(proof) Let $p = \dim E_\lambda$. Let $\{v_1, \cdots, v_p\}$ be an ordered basis for E_λ

Then the replacement theorem allows us to extend it to a basis for V

Then $A = [T]_\beta = \left[\begin{array}{c|c} \begin{matrix} \lambda & & O \\ & \ddots & \\ O & & \lambda \end{matrix} & * \\ \hline O & C \end{array}\right]$, $0 = |A - tI| = \left|\begin{array}{c|c} \begin{matrix} \lambda-t & & O \\ & \ddots & \\ O & & \lambda-t \end{matrix} & * \\ \hline O & C-tI \end{array}\right| = (\lambda-t)^p \left|\begin{array}{c|c} I & * \\ \hline O & C-tI \end{array}\right| = (\lambda-t)^p |C - tI|$

A necessary and sufficient condition for diagonalizability

Theorem T is diagonalizable $\Leftrightarrow$

(i) $|T - \lambda I|$ splits

(ii) For every eigenvalue λ of T, then we have multiplicity of $\lambda = \dim E_\lambda = n - \operatorname{rank}(T - \lambda I)$

(proof) $\dim E_k + \operatorname{rank}(T - \lambda I) = n$

T is diagonalizable $\Rightarrow |T - \lambda I|$ splits, we say with roots $\lambda_1, \cdots, \lambda_k$ with $m_1, \cdots, m_k$ as the corresponding

Let β_i be a ordered basis for E_{λ_i}. T is diagonalizable

$\Leftrightarrow \dim E_{\lambda_i} = m_i$, $\forall i = 1, 2, \cdots, k$, $\beta_1 \cup \beta_2 \cdots \cup \beta_k = \beta$ is an ordered basis for V

Cayley-Hamilton Theorem

Let $T: V \to V$ is a linear transformation and V is a finite-dimentional vector space

If $f(x)$ is the characteristic polynomial of T, then $f(T) = 0$

Def. Let $T: V \to V$ be a linear operation

A subspace $W \subseteq V$ is called T-invariant if $T(W) \subseteq W$

Def. Let $v \neq 0$ in V and $T: V \to V$ be a linear operation

Then $W = \text{span}\{v, T(v), T^2(v), \cdots, T^k(v), \cdots\}$ is called the T-cyclic subspace of V generated by V

Prop. Let $T: V \to V$ be a linear operation and V is finite-dimentional

Suppose W is a T-invariant subspace of V. Then the characteristic polynomial of T restricted to W, denoted as T_W, devides that of T

(proof) Let $\beta = \{v_1, \cdots, v_k\}$ be a ordered basis for W $\qquad [T(v_i)]_\gamma = [T_W(v_i)]_\gamma$

Extend it to a basis $\gamma = \{v_1, \cdots, v_k, \cdots, v_n\}$ for V

$[T]_\gamma = \left[\begin{array}{c|c} [T_W]_\beta & * \\ \hline O & C \end{array}\right]$ $\quad |[T]_\beta - \lambda I| = |[T_W]_\beta - \lambda I| \cdot |C - \lambda I|$, where $|[T_W]_\beta - \lambda I|$ is the characteristic polynomial

Prop. Let $T: V \to V$ be a linear transformation and V is finite-dimentional

and W be a T-cyclic subspace of V generated by $v \neq 0$ in V

Suppose $\dim W = k$, then $\exists v \in V$, s.t.

(1) $\{v, T(v), \cdots, T^{k-1}(v)\}$ is a basis for W

(2) If $a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$, then

the characteristic polynomial for T_W is $g(\lambda) = (-1)^k (a_0 + a_1 \lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k)$

(proof) Let j be the largest integer that

$\beta = \{v, T(v), \cdots, T^{j-1}(v)\}$ is linearly independent

for $j \geq 1$. let $Z = \text{span}(\beta)$ (i.e. $\dim Z = j$)

$T^j(v) \in Z$, $T(a_0 v + a_1 T(v) + \cdots + a_{j-1} T^{j-1}(v)) = a_0 T(v) + \cdots + a_{j-1} T^j(v) \in Z$

so $Z = W$, $j = k$

$\exists\, a_0, a_1, \cdots, a_{k-1}$ NOTALL zero, s.t. $a_0 v + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$

$$[T_W]_\beta = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix}$$

$g(\lambda) = |[T_W]_\beta - \lambda I| = (-1)^k (a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k)$

Proof of Cayley-Hamilton Theorem

$v \neq 0$, let W be the T-cyclic subspace generated by v

Suppose $\dim W = k$, then $\exists\, a_0, a_1, \cdots, a_{k-1}$ s.t.

$a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + a_k T^k(v) = 0$ and $g(\lambda) = (-1)^k (a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k) = 0$

$g(T)(v) = 0$ $\quad$ $f(\lambda) = P(\lambda)\, g(\lambda) = P(T)\, g(T) \Rightarrow f(T)(v) = P(T)\, g(T(v)) = 0$

<u>Def.</u> A sequence $A_k = (a_{ij}^{(k)})_{m\times n} \in M_{m\times n}(\mathbb{C})$, $k = 1, 2, \cdots$, is said to converge to $L = (l_{ij})$ in $M_{m\times n}(\mathbb{C})$ (denoted as $\lim_{k\to\infty} A_k = L$

<u>Prop.</u> (1) $A_k \to L \Rightarrow PA_k \to PL$ as $k \to \infty$ and $A_k Q \to LQ$

(2) $A^m \to L \Rightarrow (QAQ^{-1})^m \longrightarrow QLQ^{-1}$

(Another proof which is easier to comprehen)

Since T is diagonalizable, we assume the characteristic polynomial is

$f(t) = (-1)^n (t-\lambda_1)^{m_1} (t-\lambda_2)^{m_2} \cdots (t-\lambda_k)^{m_k}$, where $\sum_{i=1}^{k} m_i = n$

We choose eigenvectors from $E_{\lambda_1}, E_{\lambda_2}, \cdots, E_{\lambda_k}$, and denote them as $v_1, \cdots, v_k$

so any vector in V, $v = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$

$f(T)(v) = f(T)(a_1 v_1 + \cdots + a_k v_k) = a_1 f(T)(v_1) + a_2 f(T)(v_2) + \cdots + a_k f(T)(v_k) = a_1 f(\lambda_1) v_1 + a_2 f(\lambda_2) v_2 + \cdots + a_k f(\lambda_k) v_k = 0$

<u>Def.</u> Let V be a vector space over F ($\mathbb{R}$ or $\mathbb{C}$)
An inner product is a function $\langle \cdot, \cdot \rangle : V \times V \to F$
s.t. $\forall x, y, z \in V$, $c \in F$, we have
(1) $\langle x+z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$
(2) $\langle cx, y \rangle = c\langle x, y \rangle$
(3) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ (for real x, y, $\langle x, y \rangle = \langle y, x \rangle$)
(4) $\langle x, x \rangle \geq 0$, only when $x=0$, $\langle x, x \rangle = 0$
Example: (1) $\mathbb{R}^n$, $x=(x_1, x_2, \cdots, x_n)$ $y=(y_1, y_2, \cdots, y_n)$, $\langle x, y \rangle = x_1y_1 + \cdots + x_ny_n$
$\mathbb{C}^n$: $\langle x, y \rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n$
(2) $C([0,1])$ = the set of all complex-valued continuous functions on $[0,1]$
$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\,dt$

<u>Def.</u> A vector space equipped with an inner product is called inner-product space
<u>Prop.</u> Let V be a inner-product space. Then $\forall x, y, z \in V$, $c \in F$
(1) $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$
(2) $\langle x, cy \rangle = \bar{c}\langle x, y \rangle$
(3) $\langle x, 0 \rangle = 0 = \langle 0, x \rangle$
(4) $\langle x, y \rangle = \langle x, z \rangle$ for $\forall v \in V \Rightarrow y = z$
(proof): $\langle x, y+z \rangle = \overline{\langle y+z, x \rangle} = \overline{\langle y, x \rangle + \langle z, x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle$
$\langle x, cy \rangle = \overline{\langle cy, x \rangle} = \overline{c\langle y, x \rangle} = \bar{c}\,\overline{\langle y, x \rangle} = \bar{c}\langle x, y \rangle$

<u>Def.</u> Let V be an inner-product space over F and $v \in V$
Then $\|x\| = \sqrt{\langle x, x \rangle}$ is called the norm of x (or the length of x)
Example. In $\mathbb{R}^n/\mathbb{C}^n$, $x = (x_1, x_2, \cdots, x_n)$, $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$
On $C([0,1])$, $\|f\| = \left(\int_0^1 |f(t)|^2 dt\right)^{\frac{1}{2}}$
<u>Prop.</u> (1) $\|cx\| = c\|x\|$
(2) $\|x\| \geq 0$
(3) $|\langle x, y \rangle| \leq \|x\|\|y\|$ (Cauchy-Schwarz inequality)
(4) $\|x+y\| \leq \|x\| + \|y\|$ (Triangular inequality)
(Equality holds when x, y are colinear)
proof: (3) when $y=0$, it holds
Otherwise, set $c = \frac{\langle x, y \rangle}{\|y\|^2}$, $0 \leq \|x - cy\|^2 = \langle x-cy, x-cy \rangle$
$= \langle x, x \rangle - \langle cy, x \rangle - \langle x, cy \rangle\ \langle cy, cy \rangle = \|x\|^2 - \bar{c}\overline{\langle x, y \rangle} - c\langle x, y \rangle + c\bar{c}\|y\|^2$
$= \|x\|^2 - \frac{\overline{\langle x, y \rangle}^2}{\|y\|^2} - \frac{\langle x, y \rangle^2}{\|y\|^2}$

Def. Let $x, y \in V$: inner product space

(1) We say that $x \perp y$ (x, y are orthogonal) if $\langle x, y\rangle = 0$

(2) A subset $S \subseteq V$ is orthogonal if any two distinct vectors in S are orthogonal

(3) A subset $S \subseteq V$ is orthonormal if S is orthogonal and consists of entirely unit vectors (i.e. vectors of norm 1)

Example.

1. $\mathbb{R}^n$, $S = \{e_1, e_2, \cdots, e_n\}$ is orthonormal

2. $\mathbb{C}^n$, $S = \sim$

3. H = complex valued continuous function on $[0, 2\pi]$

with $\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$

$S = \{\cos nt + i\sin nt \mid t \in [0, 2\pi], n \in \mathbb{N}^+\}$, a orthonormal

$\frac{1}{2\pi}\int_0^{2\pi}(\cos nt + i\sin nt)(\cos mt + i\sin mt)dt = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases}$

Def. An orthonormal basis for an inner product space is an ordered basis that is orthonormal

Prop. Let $S = \{v_1, v_2, \cdots, v_n\} \subseteq V$. where V is an inner product space

(1) If S is orthagonal and $y \in \operatorname{span}(S)$, then $y = \sum_{j=1}^{n} \langle y, \frac{v_j}{\|v_j\|}\rangle \frac{v_j}{\|v_j\|}$

(2) If S in orthanormal and $y \in \operatorname{span}(S)$, then $y = \sum_{j=1}^{n} \langle y, v_j\rangle v_j$ (Since $\|v_j\| = 1$)

(3) If S if orthagonal and $v_j \neq 0$, the S is linearly independent

(proof)(1) Let $y \in \operatorname{span}(S) \Rightarrow y = a_1v_1 + a_2v_2 + \cdots + a_nv_n$

$\langle y, v_j\rangle = a_1\langle v_1, v_j\rangle + a_2\langle v_2, v_j\rangle + \cdots + a_n\langle v_n, v_j\rangle = a_j\|v_j\|^2$, so $a_j = \frac{\langle y, v_j\rangle}{\|v_j\|^2}$

(3) $a_1v_1 + \cdots + a_nv_n = 0$, $\langle 0, v_j\rangle = a_j\|v_j\|^2 = 0$, Since $\|v_j\| \neq 0$, then $a_j = 0$

Theorem (Gram-Schmidt Orthogonalization)

Let $S = \{w_1, w_2, \cdots, w_n\}$ be a linearly independent subset in an inner product space V

Then $\{v_1, v_2, \cdots, v_n\}$ is an orthogonal set where

$v_1 = w_1, \cdots, v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\|v_j\|^2} v_j$, $\forall k = 1, 2, \cdots, n$

(Proof:) $v_2 = w_2 - \frac{\langle w_2, w_1\rangle}{\|w_1\|^2} w_1$, $\langle v_2, v_1\rangle = \langle w_2 - \frac{\langle w_2, w_1\rangle}{\|w_1\|^2} w_1, w_1\rangle = \langle w_2, w_1\rangle - \frac{\langle w_2, w_1\rangle}{\|w_1\|^2}\langle w_1, w_1\rangle = 0$

$\langle v_3, v_2\rangle = \langle w_3 - \frac{\langle w_3, v_2\rangle}{\|v_2\|^2} v_2 - \frac{\langle w_3, v_1\rangle}{\|v_1\|^2} v_1, v_2\rangle = \langle w_3, w_2\rangle - \langle w_3, w_2\rangle - \frac{\langle w_3, v_1\rangle}{\|v_1\|^2}\langle v_1, v_2\rangle = 0$

by induction, $\{v_1, v_2, \cdots, v_n\}$ is orthogonal

Collary.

1. Any finite dimentional inner product space V has an orthanormal basis for V

2. Let $\beta = \{v_1, v_2, \cdots, v_n\}$ be an orthanormal basis for V and $T: V \to V$ be a linear transformation on V

Then $[T]_\beta = (a_{ij})_{n \times n}$, where $a_{ij} = \langle T(v_j), v_i\rangle$

Def. Let $S \neq \phi$ in V. Define the orthogonal complement of S as:
$S^{\perp} = \{x \in V \mid, \langle x, y \rangle = 0, \forall y \in S\}$

Theorem. Let W be a finite dimentional subspace of an inner product space V
For any $y \in V$, $\exists!\ u \in W$ and $z \in W^{\perp}$. s.t. $y = u + z$
Moreover, if $\beta = \{w_1, w_2, \cdots, w_k\}$ is an orthanormal basis for W, then $u = \sum_{j=1}^{k} \langle y, w_j \rangle w_j$ and $\|y\|^2 = \|u\|^2 + \|z\|^2$
(proof) Set $u = \sum_{j=1}^{k} \langle y, w_j \rangle w_j$ and $z = y - u$
$\langle z, w_i \rangle = \langle y - u, w_i \rangle = \langle y, w_i \rangle - \langle u, w_i \rangle = \langle y, w_i \rangle - \|w_i\|^2 \langle y, w_i \rangle = 0$
$u_1 \in W$ s.t. $y = u_1 + z_1$, $z_1 \in W^{\perp}$
$u_2 \in W$ s.t. $y = u_2 + z_2$, $z_2 \in W^{\perp}$
$0 = (u_1 - u_2) + (z_1 - z_2)$, where $(u_1 - u_1) \in W$, $(z_1 - z_2) \in W^{\perp}$
$\|y\|^2 = \langle y, y \rangle = \langle u+z, u+z \rangle = \langle u, u \rangle + \langle u, z \rangle + \langle z, u \rangle + \langle z, z \rangle = \|u\|^2 + \|z\|^2$

Collary. $\forall x \in W$, we have $\|y - u\| \leq \|y - x\|$
(i.e. u is the vector in W that is "closest" to y)
$\|y - x\|^2 = \langle u + z - x, u + z - x \rangle = \langle u - x, u - x \rangle + \langle u - x, z \rangle + \langle z, u - x \rangle + \langle z, z \rangle$
$= \|u - x\|^2 + \|z\|^2 = \|u - x\|^2 + \|y - u\|^2 \geq \|y - u\|^2$

The Adjoint of a linear operation

$T: V \to F$, V a finite dimentional inner product space

For a fixed $y \in V$, $f(x) = \langle x, y \rangle$, $\forall x \in V$

Eazy to see that $f(x)$ is a linear function

Theorem. Given a linear functional $g: V \to F$ where V is finite dimentional inner product space
$\exists!\ y \in V$, s.t. $g(x) = \langle x, y \rangle$

Theorem. Let $T: V \to V$ be a linear operation and V is a finite dimentional product space
Then $\exists!$ linear operation $T^*: V \to V$ s.t. $\langle T(x), y \rangle = \langle x, T^* y \rangle$ $\forall x, y \in V$
(proof) For every $y \in V$, define $g_y: V \to \mathbb{F}$ by $g_y(x) = \langle T(x), y \rangle$
$\forall x \in V$, we see that $g_y: V \to \mathbb{F}$ is linear, then the previous theorem $\Rightarrow \exists!\ y^* \in V$, s.t. $g_y(x) = \langle x, y^* \rangle$
Define $T^*: V \to V$ by $T^*(y) = y^*$, $\langle T(x), y \rangle = g_y(x) = \langle x, y^* \rangle = \langle x, T^*(y) \rangle$, where T^* is linear

Rk * V being finite dimentional is necessary for the existence of T^*
Rk For an $n \times n$ matrix $A = (a_{ij})_{n \times n}$ the adjoint matrix of A, $A^* = (a^*_{ij})_{n \times n}$
where $a^*_{ij} = \overline{a_{ji}}$

Prop. $T, U: V \to V$ linear operator, where V is a finite dimentional inner product space

(1) $(T+U)^* = T^* + U^*$
(2) $(cT)^* = \bar{c} T^*$, $c \in F$
(3) $(TU)^* = U^* T^*$
(4) $T^{**} = (T^*)^* = T$
(5) $I^* = I$

$\langle x, (cT)^*(y) \rangle = \langle cT(x), y \rangle = c \langle x, T^*(y) \rangle = \langle x, \bar{c} T^*(y) \rangle$
$\langle x, (TU)^*(y) \rangle = \langle TU(x), y \rangle = \langle U(x), T^*(y) \rangle = \langle x, U^* T^*(y) \rangle$

Let $T: V \to V$ be a linear operation and V is a finite dimentional inner product space

Lemma. If λ is an eigenvalue of T, then $\bar{\lambda}$ is an eigenvalue of T^*

(proof): $T(v)=\lambda v$, $0=\langle (T-\lambda I)(v), y\rangle = \langle v, (T-\lambda I)^*(y)\rangle = \langle v, (T^*-\bar{\lambda} I)(y)\rangle$. so $v \perp R(T^*-\bar{\lambda} I)$

Schur's Lemma Suppose the characteristic polynomial of T splits, then $\exists$ an orthonormal basis β s.t. $[T]_\beta$ is upper triangular

(proof) By induction, $n=1$, it holds. Now suppose the conclusion holds for $\forall$ inner product space with $\dim \leq n-1$

Suppose $\dim V = n$

Since the characteristic polynomial of T splits, we have one eigenvalue λ

By the prievious lemma, $\bar{\lambda}$ is a eigenvalue of T^*

Suppose z is a corresponding eigenvector

Let $W = \operatorname{span}\{z\}$. Claim: $\forall x \in W^\perp$, $T(x) \in W^\perp$

Suppose $y \in W^\perp$, $\langle T(y), z\rangle = \langle y, T^*(z)\rangle = \langle y, \bar{\lambda} z\rangle = 0$

Since the characteristic polynomial of $T|_{W^\perp}$ devides that of T and thus the characteristic polynomial of $T|_{W^\perp}$ also splits. As $\dim W^\perp = n-1$, induction hypothesis implies that $\exists$ an orthonormal basis β' of $W^\perp$ s.t. $[T|_{W^\perp}]_{\beta'}$ is upper triangular

Now let $\beta = \beta' \cup \left\{\frac{z}{\|z\|}\right\}$, Clearly β is orthonormal and

$$[T]_\beta = \left(\begin{array}{c|c} \begin{matrix} \ddots & [T|_{W^\perp}]_{\beta'} \\ O & \ddots \end{matrix} & O \\ \hline O & [T(z)]_\beta \end{array}\right)$$

Def. $T: V \to V$ be a linear transformation where V is a finite dimentional inner product space. Then we say that

(1) T is normal if $TT^* = T^*T$

(2) T is self-adjoint if $T = T^*$

Def. $A \in M_{n\times n}(F)$, $(F = \mathbb{R}/\mathbb{C})$

(3) A is normal if $AA^* = A^*A$

(4) A is self-adjoint if $A = A^*$

Prop. (1) T is normal $\Leftrightarrow [T]_\beta$ is normal

T is self-adjoint $\Leftrightarrow [T]_\beta$ is self adjoint $\}$ β is orthonormal

Prop. Let T be a normal operator, then

(1) $\|T(x)\| = \|T^*(x)\|$ $\forall x \in V$

(2) $T - \lambda I$ is normal, $\forall \lambda \in F$

(3) $T(x) = \lambda x \Rightarrow T^*(x) = \bar{\lambda} x$

(4) If $\lambda_1 \neq \lambda_2$ are two eigenvalues of T

where x_1, x_2 are the corresponding eigenvectors, then $x_1 \perp x_2$

(Proof) (1) $\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*T(x)\rangle = \langle x, TT^*(x)\rangle = \langle x, (T^*)^*T^*(x)\rangle = \langle T^*(x), T^*(x)\rangle = \|T^*(x)\|^2$

(2) $(T-\lambda I)(T-\lambda I)^* = (T-\lambda I)(T^* - \bar{\lambda} I) = TT^* - \bar{\lambda} TI - \lambda T^*I + |\lambda|^2 I$

$(T-\lambda I)^*(T-\lambda I) = (T^* - \bar{\lambda} I)(T - \lambda I) = T^*T - \bar{\lambda} TI - \lambda T^*I + |\lambda|^2 I$

so $(T-\lambda I)$ is normal

(3) $(T-\lambda I)(x) = (T^* - \bar{\lambda} I)(x) = 0$

(4) $\lambda_1\langle x_1, x_2\rangle = \langle \lambda_1 x_1, x_2\rangle = \langle T(x_1), x_2\rangle = \langle x_1, T^*(x_2)\rangle = \langle x_1, \bar{\lambda}_2 x_2\rangle = \lambda_2\langle x_1, x_2\rangle$

Since $\lambda_1, \lambda_2 \neq 0$, $\lambda_1 \neq \lambda_2$, then $\langle x_1, x_2\rangle = 0$

Theorem. Let T be a linear operator on a finite-dimentional inner product space V

Then T is normal $\iff \exists$ an orthonormal basis for V consisting of eigenvectors of T

(proof) $[T]_\beta$ is diagonal $\Rightarrow$ $[T]_\beta^*$ also diagonal

$[TT^*]_\beta = [T]_\beta([T]_\beta)^* = ([T]_\beta)^*[T]_\beta = [T^*T]_\beta$, so $TT^* = T^*T$

By the fundamental theorem of algebra, every polynomial splits over $\mathbb{C}$

Then Schur's Lemma guarantees that $\exists$ an orthonormal basis β for V s.t. $[T]_\beta$ is upper triangular

Let $\beta = \{v_1, v_2, \cdots, v_n\}$, Claim that $[T]_\beta$ is upper triangular

Since $[T]_\beta$ is upper-triangular, then $T(v_1) = a_{11}v_1$ $\qquad [T]_\beta = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$

By induction, assume $v_1, v_2, \cdots, v_{k-1}$ are eigenvectors for V

Need to show that v_k is an eigenvector for T

$T(v_k) = a_{1k}v_1 + a_{2k}v_2 + \cdots + a_{kk}v_k$ (Since $[T]_\beta$ is upper-triangular)

For $j = 1, 2, \cdots, k-1$, $\langle T(v_k), v_j \rangle = \langle v_k, T^*(v_j) \rangle$

Since T is normal, then $T^*(v_j) = \bar{\lambda}_j v_j$

Therefore $a_{1k}\|v_1\| = \cdots = a_{k-1,k-1}\|v_{k-1}\| = 0$, since $\|v_j\| = 1$, then $a_{j,j} = 0$

so $T(v_k) = a_{kk}\lambda_k v_k$, and we see that v_k is also an eigenvector

Examples. $\mathbb{H}$ = Complex valued continuous functions on $[0, 2\pi]$

$S = \{f_n = \cos(nt) + i\sin(nt) \mid n \in \mathbb{Z}\}$

$\langle f, g \rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$

$V = \operatorname{span}(S)$ $\quad T, U: V \to V$

$T(f_n) = f_{n+1}$

$U(f_n) = f_{n-1}$

$\langle T(f_k), f_l \rangle = \langle e^{i(k+1)t}, e^{ilt} \rangle = \begin{cases} 1, & \text{if } k+1 = l \\ 0, & \text{otherwise} \end{cases}$

Lemma. Let $T: V \to V$ be a self-adjoint linear operation and V a inner product space

Then (1) every eigenvalue of T is real

(2) Suppose V is real. then the characteristic polynomial of T splits over $\mathbb{R}$

(proof) (1). $T(v) = \lambda v$, $T^*(v) = \bar{\lambda} v$, since T is self-adjoint, then $T(v) = T^*(v)$, $\lambda = \bar{\lambda}$. $\Leftrightarrow \lambda$ is real

(2) Suppose $\dim V = n$, let β be an orthonormal basis for V and $A = [T]_\beta$ and A is self-adjoint

Let γ be the standard ordered basis, then $[L_A]_\gamma = A$

by the fundamental theorem of algebra, the characteristic polynomial of A splits

since A is self-adjoint, then it splits over $\mathbb{R}$

Theorem. S Let T be a linear operation on a real inner product space. then T is self-adjoint

$\Leftrightarrow \exists$ orthonormal basis for V consisting of eigenvalues of T

(proof) Lemma (2) $\Rightarrow$ the characteristic polynomial of T splits

Then Schur's Lemma $\Rightarrow \exists$ an orthonormal basis β for V s.t. $[T]_\beta$ is upper-triangular

$A = Q^{-1} \underset{\substack{\uparrow \\ \text{diagonal}}}{D} Q$. If columns of Q forms an orthonormal basis, then $Q^t Q = I$

Def. A matrix $A \in M_{n\times n}(F)$ is called

(1) an orthogonal matrix if $A^t A = A A^t = I$ $(F = \mathbb{R})$

(2) a unitary matrix if $A^* A = A A^* = I$ $(F = \mathbb{C})$

Rk: A matrix is orthogonal / unitary $\Leftrightarrow$ its row/columns forms an orthonormal basis

Theorem. $(F = \mathbb{C})$ A matrix A is normal $\Leftrightarrow$ A is unitarily equivalent to a diagonal matrix

$(F = \mathbb{R})$ A matrix A is self-adjoint $\Leftrightarrow$ A is orthogonally equivalent to a diagonal matrix

Def. Let V be a linear operation on a finite-dimentional inner product space V over F

(F = $\mathbb{C}$) if $\|T(x)\| = \|x\|$ for $\forall x \in V$, then T is unitary

(F = $\mathbb{R}$) if $\|T(x)\| = \|x\|$ for $\forall x \in V$, then T is orthogonal

Prop. Let $T: V \to V$ be a linear operation on a finite-dimentional inner product space V

then the following statements are equivalent

(1) $TT^* = T^*T = I$

(2) $\langle T(x), T(y) \rangle = \langle x, y \rangle$, $\forall x, y \in V$

(3) If β is an orthonormal basis for V, then AD is $T(\beta)$

(4) $\exists$ an orthonormal basis β s.t. $T(\beta)$ is also orthonormal

(5) $\|T(x)\| = \|x\|$ for $\forall x \in V$

(1) => (2): $\langle T(x), T(y) \rangle = \langle x, T^*T(y) \rangle = \langle x, y \rangle$

(2) => (3): $\beta = \{v_1, \cdots, v_n\}$, $T(\beta) = \{T(v_1), T(v_2), \cdots, T(v_n)\}$

$\langle T(v_i), T(v_j) \rangle = \langle v_i, v_j \rangle$

(3) => (4): √

(4) => (5)

Collary $T: V \to V$ linear operator and V is a finite-dim'l inner product sp. over F

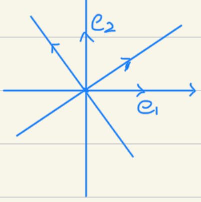
(1) (F = $\mathbb{R}$) T is self-adjoint and orthogonal <=> $\exists$ an orthonormal basis β of eigenvectors

of with all corresponding eigenvalues having absolute value 1 $\quad \|T(v_i)\| = \|\lambda v_i\| = \|v_i\|$, $|\lambda_i| = 1$

(2) (F = $\mathbb{C}$) T is unitary <=> $\exists$ an orthonormal basis β of eigenvectors

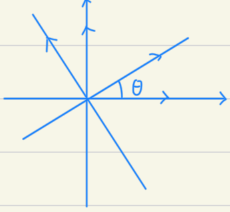
with corresponding eigenvalues having absolute value 1

Example. (1) Rotation by angle θ



$$[T]_\beta = \begin{pmatrix} [T(e_1)]_\beta & [T(e_2)]_\beta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(2) Reflection



$$[T]_\gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$[I]_\gamma^\beta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$[T]_\beta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{-1}$$

Theorem. Let T be an orthogonal op. on $\mathbb{R}^2$

& $A=[T]_\beta$ where $\beta=\{e_1,e_2\}$ is the standard ordered basis, then

(1) T is a rotation if $|A|=1$

(2) T is a reflection if $|A|=-1$

(proof)

(1) $A=[T]_\beta=\begin{pmatrix} [T(e_1)]_\beta & [T(e_2)]_\beta \end{pmatrix}=\begin{pmatrix} \cos\theta & [T(e_2)]_\beta \\ \sin\theta & \end{pmatrix}$

$T(e_2)\perp [T(e_1)]_\beta$, then $T(e_2)=\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}$ or $\begin{pmatrix}\sin\theta \\ -\cos\theta\end{pmatrix}$

$A=\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix} \Leftrightarrow |A|=1$ rotation by θ

or $=\begin{pmatrix}\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta\end{pmatrix} \Leftrightarrow |A|=-1$ reflection w.r.t. line $\frac{\theta}{2}$

Def. Let W_1, W_2 be two subspaces of a vector space V with $W_1 \oplus W_2 = V$

(i.e. $V=W_1+W_2$ and $W_1\cap W_2=\{0\}$)

A function $T:V\to V$ is called a projection on W_1 along W_2 if $T(x)=x_1$ where $x=x_1+x_2$, $x_1\in W_1$, $x_2\in W_2$

Prop. T is a projection $\Leftrightarrow T=T^2$

(proof) $(\Rightarrow)$ $x=x_1+x_2$, then $T(x)=T^2(x)=x_1$

$(\Leftarrow)$ Set $W_1=R(T)$, $W_2=N(T)$

Claim $x=T(x)+(x-T(x))$, let $x\in R(T)\cap N(T)$

then $\exists y\in V$ s.t. $T(y)=x$, $T^2(y)=T(x)=0$

$V=N(T)\oplus R(T)$

Def. Let V be a projection on an inner product space V

Then we say that T is an orthogonal projection if $R(T)^\perp=N(T)$ and $R(T)=N(T)^T$

Remark. ① If V is finite dimentional, then we need only one of the above

② If W is a finite-dimentional subspace of $V \Rightarrow \exists |$ orthogonal projection on W

$\beta=\{w_1,w_2,\cdots,w_n\}$ orthonormal basis for W

$x=\sum_{j=1}^{n}\langle x,w_j\rangle w_j$

Prop. Let T be a linear operator on an inner product space V

Then T is an orthogonal projection $\Leftrightarrow T^*$ exists and $T^2=T=T^*$

Theorem Let $T: V \to V$ be a linear operation and V be an inner product space
Then T is an orthogonal projection $\Leftrightarrow T^2 = T = T^*$
(proof) Since T is a projection, then $T = T^2$, $V = N(T) \oplus R(T)$
$\langle x, T(y) \rangle = \langle x_1 + x_2, T(y_1 + y_2) \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x_1, y_1 \rangle$
Similarly $\langle T(x), y \rangle = \langle x_1, y_1 + y_2 \rangle = \langle x_1, y_1 \rangle$, so $\langle x, T(y) \rangle = \langle T(x), y \rangle$, $T = T^*$
If $T = T^2$, then T is a projection , so $V = N(T) \oplus R(T)$
Let $x \in R(T)$, $y \in N(T)$. $\langle x, y \rangle = \langle T(x), y \rangle = \langle x, T^*(y) \rangle = \langle x, T(y) \rangle = 0$, so $R(T) \subseteq N(T)^\perp$
Reversely . $N(T)^\perp \subseteq R(T)$, so $N(T)^\perp = R(T)$

The Spectual Theorem
Let $T: V \to V$ be a linear operation and V is a finite dimentional inner product space over F
Suppose (i) T is normal if $F = \mathbb{C}$ (ii) T is self-adjoint if $F = \mathbb{R}$
Let $\lambda_1, \cdots, \lambda_k$ be the distinct eigenvalues of T and E_λ be the corresponding eigenspaces
and T_i be the projection from V to E_{λ_i} . then
(1) $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$
(2) $E_i^\perp = E_{\lambda_1} \oplus \cdots \oplus E_{i-1} \oplus E_{i+1} \oplus \cdots \oplus E_{\lambda_k}$
(3) $T_i T_j = \delta_{ij} T_i$, $\forall 1 \le i, j \le k$, $\delta_{ij} = \begin{cases} 0, & i \ne j \\ 1, & i = j \end{cases}$
(4) $I = T_1 + T_2 + \cdots + T_k$
(5) $T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k$

Collary
(1) T is normal $\Leftrightarrow T^* = P(T)$ for some polynomial T
(2) T is unitary $\Leftrightarrow$ T is normal and $|\lambda| = 1$ $\forall$ eigenvalues λ of T
(3) T is normal , then T is self-adjoint $\Leftrightarrow$ all eigenvalues are real

Collary Each T_j is a polynomial of T, $\forall j$, $T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k$
Choose g_j s.t. $g_j(\lambda_i) = \begin{cases} 0, & i \ne j \\ 1, & i = j \end{cases}$

$T: V \to V$ is a linear operation and V is a finite-dimentional vector space with characteristic polynomial splits, i.e it takes form

$(t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\cdots(t-\lambda_k)^{m_k}$, $m_1+m_2+\cdots+m_k = n = \dim(V)$, $\lambda_1, \lambda_2, \cdots, \lambda_k$ are all distinct

Say λ is an eigenvalue with multiplicity m

E_λ = the eigenspace $\{v \in V \mid T(v) = \lambda v\} = \{v \in V \mid (T-\lambda I)(v) = 0\}$

Def. $K_\lambda = \{v \in V \mid (T-\lambda I)^p(v) = 0$ for some integer $p\}$ is called the generalized eigenspace corresponding to λ and $v \in K_\lambda$ is called a generalized eigenvector

Prop (1) K_λ is T-invariant

(2) $(T-\mu I)|_{K_\lambda}$ is one-to-one, $\forall \mu \neq \lambda$

Theorem (1) Suppose λ is an eigenvalue with multiplicity m then $\dim(K_\lambda) = m$ and $K_\lambda = \ker(T-\lambda I)^m$

(2) Suppose $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of T, then

$V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots \oplus K_{\lambda_k}$

(proof)(1) Set $W = K_\lambda$ and $h(t)$ is the characteristic polynomial of $T|_W$

Then $h(t)$ devides the characteristic polynomial of T

Suppose $\dim W = d$, $\Rightarrow h(t) = (-1)^d(t-\lambda)^d$ (Since $T|_W$ has only one eigenvalue λ) and $d \le m$

Hamilton-Cayley theorem says that $h(T|_W) = 0$. $(T-\lambda I)^d(x) = 0$ for $\forall x \in W$

Since $d \le m$, then $(T-\lambda I)^m(x) = 0$ for $\forall x \in W$ (which means $x \in K_\lambda \Rightarrow x \in N((T-\lambda I)^m)$)

Also if $x \in N((T-\lambda I)^m)$, then $x \in K_\lambda$, so $K_\lambda = N((T-\lambda I)^m)$. $\dim(K_\lambda) = m$

(2) Notice that $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$ (E_λ is T-invariant)

When $k=1$, it holds by the Hamilton-Cayley theorem (*)

Induction Hypothesis (*) holds for any T with eigenvalues fewer than k

Now T has k distinct eigenvalues. Set $W = \mathrm{Im}((T-\lambda_k I)^{m_k})$ where m_k = the multiplicity of λ_k

$(T-\lambda_k I): K_{\lambda_j} \to K_{\lambda_j}$, which is one-to-one and onto $\forall j < k$

$\Rightarrow (T-\lambda_k I)^{m_k}: K_{\lambda_j} \to K_{\lambda_j}$ is one-to-one and onto $\left((T-\lambda_k I)^{m_k}(v_j) \neq 0 \text{ for } \forall m_k, v_j \in K_{\lambda_j}\right)$

i.e. $K_{\lambda_j} \subseteq W$, i.e. $\lambda_1, \cdots, \lambda_{k-1}$ are distinct eigenvalues

Induction hypothesis to $T|_W$, then $W = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_{k-1}}$

$\forall x \in V$, $(T-\lambda_k I)^{m_k}(x) = x_1 + x_2 + \cdots + x_{k-1}$ for some $x_i \in E_{\lambda_i}$

Since $(T-\lambda_k I)^{m_k}$ is one-to-one onto

$\Rightarrow \exists y_i \in K_{\lambda_i}$ s.t. $x_i = (T-\lambda_k I)^{m_k}(y_i)$

so $(T-\lambda_k I)^{m_k}(x) = (T-\lambda_k I)^{m_k}(y_1+y_2+\cdots+y_{k-1})$

so $(T-\lambda_k I)^m(x-(y_1+y_2+\cdots+y_{k-1})) = 0$, since $E_{\lambda_k} = \ker(T-\lambda_k I)^{m_k}$

then $\exists y_k \in E_{\lambda_k}$, $x-(y_1+y_2+\cdots+y_{k-1}) = y_k$, $x = y_1+y_2+\cdots+y_k$

so $V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots \oplus K_{\lambda_k}$

Theorem $\dim V = m_1 + m_2 + \cdots + m_k$

$$V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots \oplus K_{\lambda_k}$$
$$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$$

K_{λ_j} is T-invariant $T: K_{\lambda_j} \to K_{\lambda_j}$

$$[T]_\beta = \left(\begin{array}{c|c|c} [T]_{\beta_1} & O & O \\ \hline O & [T]_{\beta_2} & O \\ \hline O & O & \ddots \end{array}\right) \longrightarrow K_\lambda = \left(\begin{array}{c|c|c} \begin{matrix}\lambda & 1 \\ & \ddots \\ & & \lambda\end{matrix} & O & O \\ \hline O & \begin{matrix}\lambda & 1 \\ & \ddots \\ & & \lambda\end{matrix} & O \\ \hline O & O & \ddots \end{array}\right) = [T]_\beta$$

Def. Let $x \neq 0 \in K_\lambda$, let $p \in \mathbb{N}$ be the smallest integer that $(T-\lambda I)^p(x) = 0$

In other word $\{(T-\lambda I)^{p-1}(x), \cdots, (T-\lambda I)(x), x\}$ contains no zero vector

and we call it a cycle of generated eigenvectors

Let $\begin{cases} (T-\lambda I)(v_p) = v_{p-1} \\ (T-\lambda I)(v_{p-1}) = v_{p-2} \\ \cdots \\ (T-\lambda I)(v_1) = 0 \end{cases}$ (we allow this since K_λ is $(T-\lambda I)$-invariant

so $T(v_1) = \lambda v_1$, $T(v_2) = v_2 + \lambda v_1$, $\cdots$, $T(v_k) = v_k + \lambda T v_{k-1}$

let $\gamma = \{v_1, v_2, \cdots, v_p\}$

then $[T]_\gamma = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \ddots & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$, which is the Jordan Block

Remark $\{v_1, v_2, \cdots, v_p\}$ is linearly independent if $x = (T-\lambda I)(v)$, $v \in K_\lambda$, then $x \in K_\lambda$ ✓

<u>Theorem</u>. K_λ has an orderded basis consisting of distinct cycles of generated eigenvectors corresponding to λ

(proof) Induction on $\dim(K_\lambda) = n$. when $n=1$, it holds

Suppose the Theorem holds when $\dim(K_\lambda) \leq n-1$, Let $U = (T-\lambda I)|_{K_\lambda}$

Observe that $\mathrm{im}(U) \subsetneq K_\lambda$. $\dim K_\lambda = \mathrm{rank}(U) + \mathrm{null}(U)$

so $\dim(R(U)) \leq n-1$ and therefore the induction hypothesis applies to $R(U)$

i.e. $R(U)$ has a basis consisting of distinct cycles $\gamma_1, \gamma_2, \cdots, \gamma_q$

γ_i: let x_i be its end vector and $U(v_i) = x_i$ $\quad (T-\lambda I)^{p}(v_i) = w_i$

Let $\tilde{\gamma}_i = \gamma_i \cup \{v_i\}$ be a cycle in K_λ (and v_i is a generator)

Let w_j be the initial vector of γ_j (Here w_j is an eigenvector)

$\{w_1, \cdots, w_q\}$ is linearly independent and we can extend it to a basis for E_λ: $\{w_1, \cdots, w_q, u_1, \cdots, u_s\}$

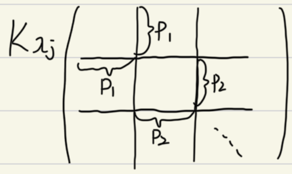
Claim: $\hat{\gamma} = \hat{\gamma}_1 \cup \cdots \cup \hat{\gamma}_q \cup \{u_1, \cdots, u_s\}$ is a basis for K_λ

(pf) $\hat{\gamma}$ is linearly independent. Let $W = \mathrm{span}(\hat{\gamma})$ and $k = \#\hat{\gamma}$

$k \geq \dim W = \dim(R(U)) + \dim(N(U)) = \#(\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_q) + q + s = \#(\hat{\gamma}_1 \cup \hat{\gamma}_2 \cup \cdots \cup \hat{\gamma}_q) + s = k$

so $\hat{\gamma}$ is linearly independent

$$\begin{pmatrix} K_{\lambda_1} & & & \\ & K_{\lambda_2} & & \\ & & \ddots & \\ & & & K_{\lambda_k} \end{pmatrix}$$



r_1: the number of eigenvectors

$\{(T-\lambda I)^{p-1}(x), (T-\lambda I)^{p-2}(x), \cdots, x\}$

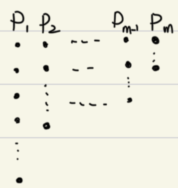
initial vector $\qquad$ end vector

$r_1 = \dim K_\lambda - \mathrm{rank}(T|_{K_\lambda} - \lambda I)$

$r_2 = \mathrm{rank}(T|_{K_\lambda} - \lambda I) - \mathrm{rank}(T|_{K_\lambda} - \lambda I)^2$

...

$r_j = \mathrm{rank}(T|_{K_\lambda} - \lambda I)^{j-1} - \mathrm{rank}(T|_{K_\lambda} - \lambda I)^j$



$r_1 \geq r_2 \geq r_3$

Example V = the space of all polynomial in x, y of power ≤ 2 , $T(f(x,y)) = \frac{\partial f}{\partial x}$

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

$$[T]_\beta = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & -\lambda & 0 & 2 & 0 & 0 \\ 0 & 0 & -\lambda & 0 & 1 & 0 \\ 0 & 0 & 0 & -\lambda & 0 & 0 \\ 0 & 0 & 0 & & -\lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda \end{vmatrix} = \lambda^6 = 0$$

$\lambda = 0, 0, 0, 0, 0, 0$

$V = K_\lambda$

$r_1 = 6 - \operatorname{rank} A = 3$

$r_2 = \operatorname{rank} A - \operatorname{rank} A^2 = 2$

$$J = \left(\begin{array}{ccc|cc|c} 0 & 1 & 0 & & & \\ & 0 & 1 & & 0 & 0 \\ & & 0 & & & \\ \hline & & & 0 & 1 & \\ & 0 & & & 0 & 0 \\ \hline & 0 & & & & 0 \end{array}\right)$$

$$0 = (A - \lambda I)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} x_2 \\ 2x_4 \\ x_5 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_2 = x_4 = x_5 = 0$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$Aw_1 = v_1 \Rightarrow w_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad Au_1 = w_1 \Rightarrow u_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{2} \\ 0 \\ 0 \end{pmatrix}$$

$\gamma = \{v_1, w_1, u_1, v_2, w_2, u_2, v_3, w_3, u_3\}$