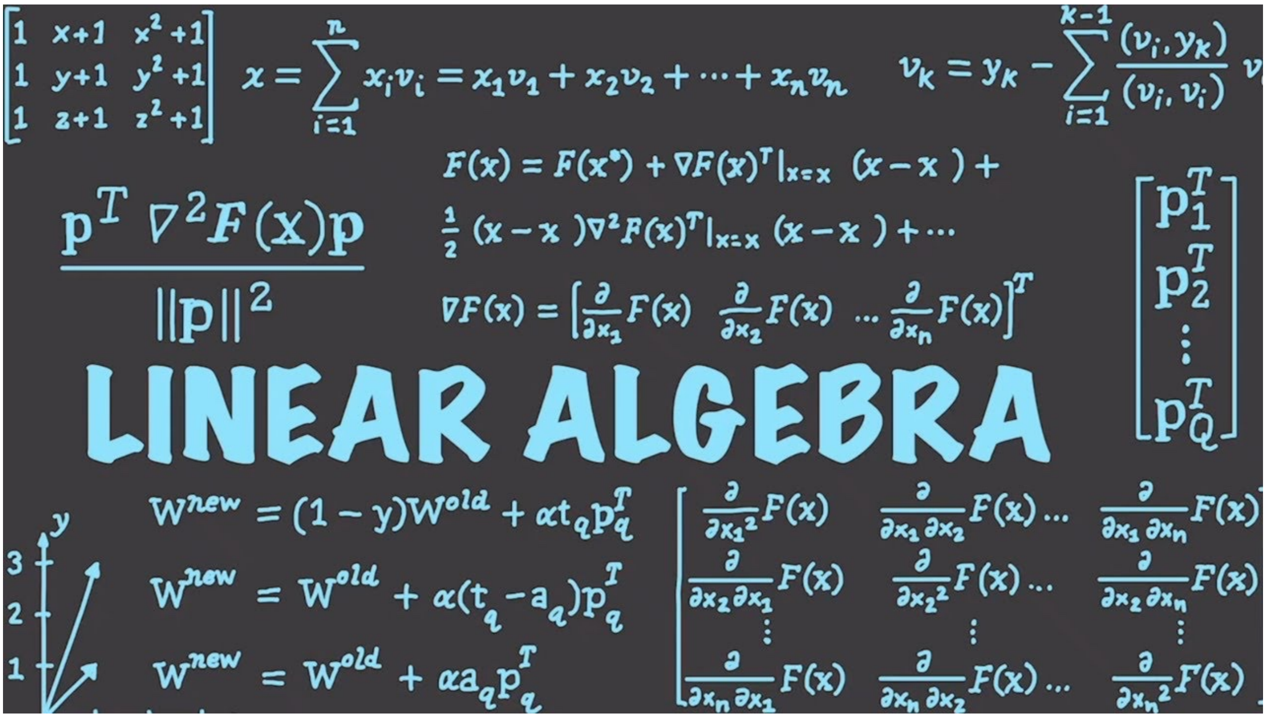


MIT open course & MAT2040 Notebook

Youthy WANG

王俣哲(WANG Yuzhe)

Linear Algebra & applications

(MIT open course, MAT2040)

• Linear Equations & matrices

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases} \Leftrightarrow \underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{b}$$

A → coefficient matrix

1) row pictures of $Ax = b$.

Show $2x - y = 0$, $-x + 2y = 3$ in the same xy-plane, then finding crossing point.

2) column pictures of $Ax = b$.

write as $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$. The linear combination of columns $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ & $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$

How to multiply vectors by matrices?

(when the matrix is $m \times n$, meaning m rows, n columns, when $n=1$, it expresses a vector with m parameters.)

⇒ Use linear combination. eg. $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$

terminology $\begin{cases} m > n: \text{overdetermined system} \\ m < n: \text{underdetermined system} \end{cases}$

linear systems ⇒ $\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$

• Matrix operations

1) Addition: If A & B are both $m \times n$ matrices, then $A+B$, formed by addition of elements in row i, column j both for A & B.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & \cdots & \cdots & b_{2n} \\ \vdots & & & \vdots \\ b_{m1} & \cdots & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & & & \vdots \\ a_{m1}+b_{m1} & \cdots & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

2) Product: $A_{m\times n} B_{n\times p} = C_{m \times p}$ (A with n columns, B with n rows)

Way I: row i — $\underbrace{\begin{bmatrix} \cdots a_{ik} \cdots \end{bmatrix}}_{A_{m\times n}} \underbrace{\begin{bmatrix} \vdots \\ b_{kj} \\ \vdots \end{bmatrix}}_{B_{n\times p}}$ (column j) $= \underbrace{\begin{bmatrix} \cdot \end{bmatrix}}_{C=AB}$ (c_{ij})

consider c_{ij} = sum of dot product $= \sum_{k=1}^{n} a_{ik}b_{kj}$

$\underline{(\text{row}_i \text{ of } A) \times (\text{column}_j \text{ of } B)}$

Way II: column ways $A\cdot colk$

$$\begin{bmatrix} & & \\ & & \end{bmatrix}_{A_{m\times n}} \begin{bmatrix} & \overset{colk}{|} & \\ & | & \end{bmatrix}_{B_{n\times p}} = \begin{bmatrix} & | & \\ & | & \end{bmatrix}_{C=AB}$$

Recall $A\cdot colk$ = new column

Columns of C are combinations of columns of A.

(coefficients are columns of B)

Way III: row ways

$$\begin{bmatrix} \text{rowk} \\ \end{bmatrix}_{A_{m\times n}} \begin{bmatrix} \\ \end{bmatrix}_{B_{n\times p}} = \begin{bmatrix} \text{rowk}\cdot B \\ \end{bmatrix}_{C=AB}$$

Recall rowk $\cdot B$ = new row

Rows of C are combinations of rows of B.

(coefficients are rows of A)

Way IV: recall $A_{m\times 1} B_{1\times p} = [\,][\quad] = C_{m\times p}$

$C = AB =$ sum of (columns of A) $\times$ (rows of B) $= \sum_{i=1}^{n}$ (col i of A)(row i of B)

Symbol of diagonal matrix

$D_{n\times n} \triangleq diag(d_{11}, d_{22} \cdots, d_{nn})$

• Inverses, Transpose & Permutation

1) Only square matrices can have their inverses, satisfying $\boxed{A^{-1}A = AA^{-1} = I}$

I – identity matrix $\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$ (only 1's on the left diagnal.)

2) The matrix A is invertible if there is a matrix A^{-1}, s.t. $AA^{-1} = A^{-1}A = I$ (square m)

on the opposite, A is said to be singular.

3) The equation test for invertibility: if there is a non-zero vector x, s.t. $Ax=0$, then A is singular.

(Reason: suppose A^{-1} exists. $A^{-1}Ax = 0$, then $A^{-1}Ax = Ix = x = 0$, contradictory)

4) Find inverses by elimination:

Gauss-Jordan Idea & augmented matrices

$[A|I] \xrightarrow{\text{elimination}} [I|A^{-1}]$ (Reason: $[EA|EI] \overset{E=A^{-1}}{=\!=} [I|A^{-1}]$, just find that E)

augmented matrices ($[A|b]$ in linear systems)

5) Operations: $(AB)^{-1}=B^{-1}A^{-1}$, $(ABC)^{-1}=C^{-1}B^{-1}A^{-1}$ (reverse order) generally not equal

A^{-1}, B^{-1} exists $\not\Leftrightarrow (A+B)^{-1}$ exists ; $(A+B)^{-1}$ "not necessarily" $\neq A^{-1}+B^{-1}$

6) Transpose: for a given A, columns of A^T are the rows of A.

eg. $A=\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix} \Rightarrow A^T=\begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 4 \end{bmatrix}$ $((A^T)_{ij}=A_{ji})$

7) Operations: $(AB)^T=B^TA^T$, $(ABC)^T=C^TB^TA^T$, $(A^{-1})^T=(A^T)^{-1}$ (reverse order)

$(A+B)^T=A^T+B^T$

8) A symmetric matrix has $S^T=S$, S^TS also always symmetric.

9) Permutation: family of matrices. (eg by 3×3)

Permutation P has the same row as I, totaly $n!$ different orders/forms.

$P_{12}=\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $P_{23}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $P_{31}=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ properties:

$P^{-1}=P^T$ (orthogonal matrix)

P for others $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ & $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $P^TP=P^{-1}P=I$

Additions:

all products of P's get another P in the family.

☆ Skew-symmetric / antisymmetric matrices: (have) $A^T=-A \Rightarrow a_{ij}=\begin{cases} -a_{ji}, & i\neq j \\ 0, & i=j \end{cases}$

Partition $A=\underbrace{\frac{A+A^T}{2}}_{\text{symmetric}}+\underbrace{\frac{A-A^T}{2}}_{\text{anti-symmetric}}$

☆ Uniqueness of inverse (proof): assume B, C are both inverse of A (invertible)

$B=BI=(BA)C=IC=C \Rightarrow$ unique A^{-1}

- Other computing tech:

$A(BC)=(AB)C$ parentheses not needed

$AB \neq BA$ (not necessarily)

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

☆ Block Multiplication

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & \cdots \\ \cdots & \cdots \end{bmatrix}$$ same as basic rules

all matrices

Same # rows — A_{11}. A_{12}; same # cols: A_{11}, A_{21}

rows, determined by the partition of n;
columns — randomly partition!

- The Idea of Elimination.

☆ The goal of elimination — get an upper triangle system

eg. $x+2y+z=2$
$3x+8y+z=12$
$4y+z=0$

$$\xRightarrow{\text{transfer}} \begin{bmatrix} \boxed{1} & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}_{A} \xrightarrow{\text{by } E_{21}} \begin{bmatrix} \boxed{1} & 2 & 1 \\ 0 & \boxed{2} & -2 \\ 0 & 4 & 1 \end{bmatrix} \xrightarrow{\text{by } E_{32}} \begin{bmatrix} \boxed{1} & 2 & 1 \\ 0 & \boxed{2} & -2 \\ 0 & 0 & \boxed{5} \end{bmatrix}_{U}$$

first pivot; second pivot; 3rd pivot

$x+2y+z=2$
$2y-2z=6$
$5z=-10$

(when 0 is in the pivot & no 0s below it, exchange row!)

$Ax=b \Rightarrow Ux=c$

- Relationship between A & U

$E_{32}E_{21}A=U$, by inversity $A=(E_{21}^{-1}E_{32}^{-1})U=LU$ → upper triangle system

E → elementary matrices (Page. 36)

Lower triangle system

$$E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \quad E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

• The reason why $A = LU$.

By definition of elimination. Row 3 of U = Row 3 of A − l_{31}(Row 1 of U) − l_{32}(Row 2 of U)

$A = \hat{L} D \hat{U}$ → diagonalized

unit (with diagonal all = 1)

Row 3 of $A = \underbrace{[\,l_{31}\ \ l_{32}\ \ 1\,]}_{\text{Row 3 of } L} \begin{bmatrix} \text{Row 1 of } U \\ \text{Row 2 of } U \\ \text{Row 3 of } U \end{bmatrix}$

How many times needed to calculate triangular systems & elimination?

1) Triangular system: $\boxed{\frac{n^2}{2}}$ multiply-subtracts. $(\underset{\text{last}}{1} + 2 + \cdots + (n-1) + \underset{\text{first}}{n}) = \frac{n(n-1)}{2} \sim \frac{n^2}{2}$)

2) Elimination ($n \times n$): (G-J)

$\begin{bmatrix} - & - & - \\ - & - & - \\ - & - & - \end{bmatrix} \xrightarrow[n^2]{\text{first}} \begin{bmatrix} 0 & - & - \\ 0 & - & - \\ 0 & - & - \end{bmatrix} \xrightarrow{(n-1)^2}$ stay the same — totally $\boxed{n^3/3}$ (because of $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \sim n^3/3$)

→ (existence & uniqueness ⇒ needs futher proof)

• Reduced Row Echleon Form. $R = \text{rref}(A)$

all pivots = 1 & zeros below and above them

e.g. $R = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(solution of $Ax=b$ when I (identity))

all pivot columns $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}$

all free columns $\begin{bmatrix} F \\ 0 \end{bmatrix}$

☆ ↳ get ⇒ Gauss-Jordan Elimination

⇔ { Gassian Elimination (forward elimination) + Back-substitution

☆ "rref" for augmented matrix (the last column may be pivot column → no solution!)

• Vector Space

☆ The vector space is closed under linear combinations. (i.e. additions & multiples)

e.g. if $\vec{u}$ & $\vec{v}$ in vector space, $c\vec{u} + d\vec{v}$ in it for any $c, d \in \mathbb{R}$

$\mathbb{R}^n$ consists of all column vectors $\vec{v}$ with n components.

★ A subspace of $\mathbb{R}^n$ is a vector space inside $\mathbb{R}^n$ (e.g. $y = x$ line in $\mathbb{R}^2$)

(check $\underset{\text{not } \Phi}{0}$ & closed under linear combintion for subspaces)

All subspaces of R^2: 1) R^2; 2) Lines L through $\begin{bmatrix}0\\0\end{bmatrix}$; 3) $Z=\begin{bmatrix}0\\0\end{bmatrix}$

2 subspaces $P \cap L$ is also a subspace.
2 subspaces $P \cup L$ is always not a subspace.
Additions more precisely: usually not a subspace.
☆ Span (defn) $\mathcal{U}=\{u_1, u_2, \dots u_m\} \subseteq R^n$. $\boxed{span(\mathcal{U})} := \{k_1u_1 + \dots + k_mu_m \mid (k_1, \dots, k_m) \in R^n\}$
(a definition)

Other form of spaces. (generally linear combination / independence)
M: the vector space of all real 2×2 matrices
F: the vector space of all real functions
can add & multiply
considered as vector
eg. all 3×3 matrices space
subspaces: (eg) all upper triangles; all symmetric matrices; all diagonal matrices

☆ spanning set (defn): $\mathcal{U}=\{u_1, u_2, \dots u_m\} \subseteq R^n$. $\mathcal{U}$ is called the/a spanning set of R^n
if $span(\mathcal{U}) = R^n$.

• Rank & solving $Ax=b$

free columns
pivots

eg. $A=\begin{bmatrix}1&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix} \xrightarrow[\text{elimination}]{\text{do}}$ "$A \Rightarrow R$" $\begin{bmatrix}\boxed{1}&2&0&-2\\0&0&\boxed{1}&2\\0&0&0&0\end{bmatrix}$

pivot columns

★ # pivots / pivot columns = rank of A
(R(A))
‖
non-zero rows

Solve $Ax=b$:
1) $X_{particular}$: set all free varibles (variables corresponding with free columns), to be zero, solve for pivot.
free / independent variables
2) $X_{nullspace}$. (every time one free variable to be 1, others zero)
3) $X_{complete} = X_p$ + any of X_n

Discussion over rank r:

$r \leq m$ & $r \leq n$ (by definition)

1) Full column rank, namely $r=n \leq m$: all pivot columns; nullspace $N(A) = \text{zero}$

It has 0/1 solution, which is the special solution if 1. $R = \begin{bmatrix} I \\ 0 \end{bmatrix}$

2) Full row rank, namely $r=m \leq n$: no zero rows; solutions for every RHS b

It usually has infinite solutions.

3) Others: $r=m=n$, invertible matrix, $R=I$, 1 solution for every b (except for zero)

$r<m$ & $r<n$, not full rank, 0 or ∞ solutions.

☆ $Ax=b$ has solutions $\Leftrightarrow$ $\text{rank}(A) = \text{rank}(A,b)$

$AX=B$ (matrix) has solutions $\Leftrightarrow$ $\text{rank}(A,B) = \text{rank}(A)$

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

• Independence, Basis & Dimension

☆ Linear independence: vectors $\vec{x}_1, \vec{x}_2, \cdots \vec{x}_n$ are independent if no combinations give a zero vector except for zero ones (all $c_i = 0$)

When $v_1, v_2, \cdots v_n$ are columns of A, they are independent if $N(A) = \text{zero vector}$ & rank of $A = n$. (rank of $A < n$ when dependent)

Basis for a vector space is a sequence of vectors $v_1, v_2, \cdots v_d$ (spanning set - space) with two properties: 1) They're all independent. 2) They span a space.

cannot be more / cannot be less

Dimension of a vector space is the number of vectors in every basis.

(Prerequest: # of every basis stays the same.)

proof — argue by contradicts ($m<n \Rightarrow$ dependent n vectors)

Some Notes & Relations.

1) The vectors $v_1, v_2, \dots v_n$ are a basis for $\mathbb{R}^n$ exactly when they are the columns of an $n \times n$ invertible matrix.

2) The pivot columns of A are a basis of C(A)

eg. Dimension & Basis for the whole $n \times n$ matrix space — n^2 (every 1 position be 1, others 0)

Dimension of Symmetric matrices $_{(n\times n)}$ = Dimension of upper triangles $_{(n\times n)}$ = $\frac{1}{2}n^2 + \frac{1}{2}n$

Function spaces: (eg) all solutions to $\frac{d^2y}{dx^2} + y = 0 \rightarrow y = c_1\cos x + c_2 \sin x$

"Unique expression" of a vector based on basis. $\Rightarrow$ basis, dimension = 2

☆ (Defn) — ~~Operator to take coordinate in the linear tranformation~~

Suppose $\mathcal{U} = \{u_1, \dots u_n\}$. $x = c_1u_1 + \dots + c_nu_n$, then $[x]_{\mathcal{U}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ coefficients

coordinates in a general vector space

$[\ ]_{\mathcal{U}}$ is a linear combination

$[\alpha x + \beta y]_{\mathcal{U}} = \alpha[x]_{\mathcal{U}} + \beta[y]_{\mathcal{U}}$

(change bases: Page 32) $\rightarrow$ "$[\ ]_{\mathcal{U}} = V^{-1}U[\ ]_{\mathcal{V}}$"

• Matrices with rank = 1

1) If R(A) (rank of A) = 1, $A \xrightarrow[\text{written as}]{\text{can be}} UV^T$ (a column $\times$ a row) $\rightarrow$ rank one product / outer product

(reason: all lines are times of the previous one)

2) Can all rank one matrices form a space? No, rank may change (same size)

$\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$ (can be proved)

(Use $\text{rank}(A+B) \leq \text{rank}(A, B) \leq \text{rank}(A) + \text{rank}(B)$)

• Generalized Elementary Transformations / Elementary Row Transformations for block matrices

① $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \xrightarrow[\text{exchange}]{\text{row}} \begin{bmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{bmatrix}$ ② $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \xrightarrow[\text{invertible } P]{[i(p)]} \begin{bmatrix} A_{11} & A_{12} \\ PA_{21} & PA_{22} \end{bmatrix}$

③ $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \xrightarrow[\text{invertible } P]{[2+1(p)]} \begin{bmatrix} A_{11} & A_{12} \\ PA_{11}+A_{21} & PA_{12}+A_{22} \end{bmatrix}$

• Four Basic Subspaces. (suppose A: $m\times n$ matrix)

1) Column Space of A: consists of all linear combinations of columns of A

(in $\mathbb{R}^m$) Basis — all pivot columns of A / R; $\to$ no row exchanges

Dimension of $C(A)$ ($\dim C(A)$) = rank(A) (r)

2) Row Space of A / $C(A^T)$: consists of all linear combinations of rows of A / columns of A^T

(in $\mathbb{R}^n$) Basis — all non-zero row of R / all pivot columns of A^T

Dimension of $C(A^T)$ ($\dim C(A^T)$) = rank(A) (r)

3) Null-Space of A / $N(A)$: consists of all solutions to $Ax=0$

(in $\mathbb{R}^n$) Dimension of $N(A)$ = # free variables $= n-r$

[A nullspace is a space $\longrightarrow$ $Ac_1=0, Ac_2=0 \Rightarrow A(c_1+c_2)=0, A(kc_1)=0$]

4) Nullspace of A^T / Left nullspace of A ($A^Ty=0 \Rightarrow y^TA=0^T=0$) / $N(A^T)$: all solutions to $A^Tx=0$

(in $\mathbb{R}^m$) Dimension of $N(A^T) = m-r$

(Note*: By row exchange, $C(A)\neq C(R)$; $C(A^T)=C(R^T)$)

☆

Rank Theorem: # independent rows = # independent columns

Counting Theorem: $\dim C(A) + \dim N(A) = \dim \mathbb{R}^n$ (A — $m\times n$ matrix)

(Proof of the rank theorem)

Do row operations $\Rightarrow$ RREF $\Rightarrow$ (# non zero rows = # pivot columns)

$\therefore$ row rank of A = column rank of A

☆ Nullity (defn): $n(A) = \dim(N(A)) = n - \text{rank}A$.

• A rank Equality: Let $A\in\mathbb{R}^{m\times n}$, then $[\text{rank}(I_m - AA') - \text{rank}(I_n - A'A) = m-n]$

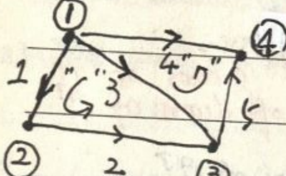
(proof.) construct $\begin{bmatrix} I_m - AA' & 0 \\ 0 & I_n \end{bmatrix}$ & do elementary operations (generalized)

Note that $\text{rank}(A)+\text{rank}(B) = \text{rank}\begin{bmatrix} A & 0 \\ 0 & B\end{bmatrix}$

• Graphs & Networks ; Edge-Node Matrix

★ Application. (Kirchhoff's Laws)

Incidence matrix: every entry of which comes from $\{0, \pm 1\}$



Incidence matrix A (5×4)

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

outward

inward

the i-th row: edge i

the j-th column = node j

Tree: the graph that has no closed loops.

★ Elimination reduces every graph to tree.

(Rows are linearly dependent when edges form a loop.)

$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, are voltages of each node.

1) Then the nullspace of A / $N(A)$ — all voltages are equal. dim = 1

(Reason: $Ax = 0 \Leftrightarrow \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_3 - x_1 \\ x_4 - x_1 \\ x_4 - x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$) $(\dim N(A) = 1)$

Basis $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$. Given $x_i = 0$, → known as "grounded" for $Ax = e$ ← outside voltage

2) The Row space of A / $C(A^T)$, $\dim C(A^T) = \text{rank}(A) = 4 - 1 = 3$

v in $C(A^T)$ if & only if it is perpendicular to $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ in $N(A)$.

3) The column space of A / $C(A)$ & KVL (Kirchhoff's Voltage Law)

Solve $Ax = b$; KVL: components of $Ax = b$ add to zero around every loop.

4) The left-nullspace of A / $N(A^T)$ & KCL (Kirchhoff's Current Law)

$\dim N(A^T) = 5 - \text{rank}(A^T) = 2$

Consider $A^Ty=0$ ← KCL → Flows in = Flows out at each node.

Basis for $N(A^T)$ (find 2 loops) $\begin{bmatrix}1\\1\\-1\\0\\0\end{bmatrix}$ & $\begin{bmatrix}0\\0\\1\\-1\end{bmatrix}$

$$\#\text{ independent loops} = \overset{(\dim N(A^T))}{m-\underset{\dim C(A^T)}{r}} = m-(n-\overset{\dim N(A)}{1}) = m-n+1 = \#\text{edges} - \#\text{nodes} + 1$$

⇒ Euler's Formula: #nodes − #edges + # independent small loops = 1

Big Picture: $\left\{\begin{aligned} &\overset{\text{no source}}{Ax=e} \\ &y=Ce \text{ (OHM'S LAW)} \longrightarrow y=\overset{\text{direction}}{-}CAx \\ &A^Ty=0 \text{ (KCL)} \longrightarrow \overset{\text{with source}}{A^Ty=f} \end{aligned}\right\} \Rightarrow f=A^TCAx$

- Rank Inequalities: (with IDEA of proofs

① $\text{rank}(A+B) \le \text{rank}(A)+\text{rank}(B)$ & $\text{rank}(A-B) \ge \text{rank}(A)-\text{rank}(B)$

② Sylvester's Inequality: Let $A\in\mathbb{R}^{m\times n}$; $B\in\mathbb{R}^{n\times k}$, the we have:

$$\underbrace{\text{rank}(A)+\text{rank}(B)-n}_{\text{lower bound}} \le \text{rank}(AB) \le \underbrace{\min\{\text{rank}(A),\text{rank}(B)\}}_{\text{upper bound}}$$

(proof:) RHS − easy (see page 7), LHS: $\text{rank}(AB)+n = \text{rank}\begin{pmatrix} I_n & 0 \\ 0 & AB\end{pmatrix} \xlongequal[\text{operations generalized}]{\text{elementry row}}$

$\text{rank}\begin{pmatrix} I_n & 0 \\ A & AB\end{pmatrix} \overset{(\ldots)}{=} \text{rank}\begin{pmatrix} I_n & -B \\ A & 0\end{pmatrix} \overset{(\ldots)}{=} \text{rank}\begin{pmatrix} B & I_n \\ 0 & A\end{pmatrix} \ge \text{rank}(A)+\text{rank}(B)$.

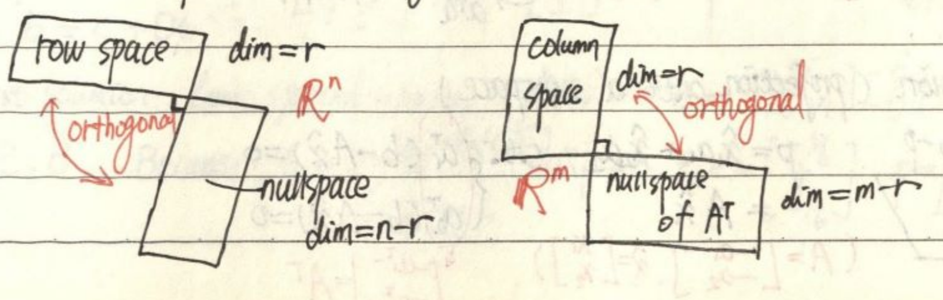
(or more directly, use $\begin{bmatrix} I_n & B \\ A & 0\end{bmatrix} = \begin{bmatrix} I_n & 0 \\ A & I_m\end{bmatrix}\begin{bmatrix} I_n & 0 \\ 0 & AB\end{bmatrix}\begin{bmatrix} I_n & B \\ 0 & -I_k\end{bmatrix}$)

- Orthogonality

☆ Orthogonal vectors have $x^Ty=0$. (perpendicular in other words) denoted as $x\perp y$

Orthogonal subspaces have every vector in S & every vector in T are orthogonal.

Four basic subspaces (the big picture)



(Reason: row space is orthogonal to nullspace

$$Ax=0 \Rightarrow \begin{bmatrix} \text{Row 1 of } A \\ \vdots \\ \text{Row m of } A \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$ (meaning that Row i of $A \cdot x_{null} = 0$)

Then ($x_{\text{in rowspace}}$ = combinations of rows) $\cdot\, x_{null} = 0$

The other pair — change A to A^T.)

☆ Orthogonal Complement: the orthogonal complement of a subspace V contains every vector that is perpendicular to V, denoted by $V^\perp$ ($\dim V + \dim V^\perp = n$ & V's basis $+ V^\perp$'s basis $\rightarrow \mathbb{R}^n$'s ~)

(Row space & Nullspace are partitions of $\mathbb{R}^n$ — orthogonal complement.)

(Defn) — Directed sum / Direct sum: $W = U \oplus V$, if $w \in W$ (any w), $\exists\, u \in U, v \in V$, s.t. $w = u + v$. & $U \cap V = \{\underline{0}\}$.

• Matrix Derivatives

(1) Categories:

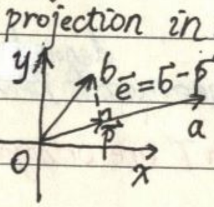
	scalar	vector	matrix
scalar	dy/dx	$\frac{d\vec{y}}{dx} = \left[\frac{\partial y_i}{\partial x}\right]_{n\times1}$	$\frac{dM}{dx} = \left[\frac{\partial y_{ij}}{\partial x}\right]_{m\times n}$ ($M \in \mathbb{R}^{m\times n}$)
vector	$\frac{dy}{d\vec{x}} = \left[\frac{\partial y}{\partial x_j}\right]_{1\times m}$ row vector	$\frac{d\vec{y}}{d\vec{x}} = \left[\frac{\partial y_i}{\partial x_j}\right]_{n\times m}$	($\vec{y} \in \mathbb{R}^n, \vec{x} \in \mathbb{R}^m$)
matrix	$\frac{dy}{dX} = \left[\frac{\partial y}{\partial x_{ji}}\right]_{n\times m}$ ($X \in \mathbb{R}^{m\times n}$)		

same, Jacobian matrix

different from usually defined as $\left[\frac{\partial y}{\partial x_j}\right]_{m\times1}$, etc. Need to do transpose in optimization derivatives.

• Projections

1) projection in $\mathbb{R}^2$ (can be visualized):



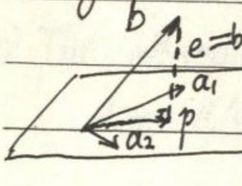
aim — find $\vec{p}$ (smallest $\vec{e}$, shortest distance)

$\vec{p} = \hat{x}\vec{a}$ (find x), then $\vec{a} \cdot (\vec{b} - \hat{x}\vec{a}) = a^T(b - \hat{x}a) = 0$

we can get $p = a\frac{a^Tb}{a^Ta}$, proj $p = \underset{\text{project matrix}}{P}\, b$

$\left\{ \begin{matrix} m\times m \text{ matrix} \\ \text{rank}(P) = 1 \end{matrix} \right. \leftarrow P = \frac{aa^T}{a^Ta}$ ($\hat{x}$) $\rightarrow \frac{a^Tb}{a^Ta}$ ($\hat{x} \neq P$)

2) high-dimension (projection onto a subspace)



$p = \hat{x}_1 a_1 + \hat{x}_2 a_2$, s.t. $\left\{ \begin{matrix} a_1^T(b - A\hat{x}) = 0 \\ a_2^T(b - A\hat{x}) = 0 \end{matrix} \right.$

$= A\hat{x}$

$\left(A = \begin{bmatrix} -a_1- \\ -a_2- \end{bmatrix}, \hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)$ $\quad \begin{bmatrix} -a_1^T- \\ -a_2^T- \end{bmatrix} = A^T$

$$A^T(b-A\hat{x})=0 \Rightarrow \begin{cases} \hat{x}=(A^TA)^{-1}A^Tb \\ p = A\hat{x} = A\underbrace{(A^TA)^{-1}A^T}b \\ P \xrightarrow[\text{matrix}]{} A(A^TA)^{-1}A^T \end{cases}$$

(Note: when A is a square matrix & invertible, we can write P as $A(A^TA)^{-1}A^T = (A\cdot A^{-1})(A^T)^{-1}A^T = I$,
(meaning that the projection of the whole $\mathbb{R}^m/\mathbb{R}^n$ is itself))

3) Properties. $P^2 = P = P^T$ (in algebra & true meaning); idempotent ($P^2=P$)
$e = b - p = (I-P)b$ in $N(A^T)$ (p in $C(A)$)

e.g. $\Pi_\Omega(y)$, with $y\in\mathbb{R}^n$, $\Omega=\{x\in\mathbb{R}^n \mid Ax=b,\ A\in\mathbb{R}^{m\times n},\ b\in\mathbb{R}^m\}$ (projection of y on Ω)

One way: Go back to the defn, using the KKT conditions.
$\boxed{\Pi_\Omega(y) = (I - A^T(AA^T)^{-1}A)y - A^T(AA^T)^{-1}b}$ cannot use a projection matrix to replace. Because Ω is not a space.

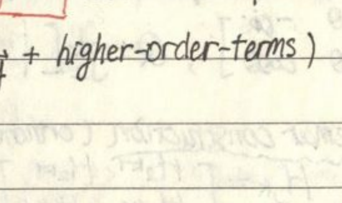
- Matrix Derivatives (cont'd, page 12)

Define the differential $\boxed{dy(x) = y(x+dx) - y(x)}$ even work for x, y not scalars
(e.g. $\vec{y}(\vec{x})$; $d\vec{y}(\vec{x}) = \vec{y}(\vec{x}+d\vec{x}) - y(\vec{x}) = J_{\vec{x}\to\vec{y}}$ + higher-order-terms)
Take linear part!

- Least Sqares Approximations

☆ When $Ax=b$ has no solution, $A^TA\hat{x}=A^Tb$ ($\hat{x}$ - least square solution)
($\|b-Ax\|$ attaches to its minimum)

Least squared fitting by a line.



(eg. as RHS shows) $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$
$b = C + Dt$

best solution - least squared solution $A^TA\hat{x} = A^Tb \Rightarrow \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}\begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$

$\Rightarrow \hat{C}, \hat{D}$. Because $b = p + e$, $e \to N(A^T)$, $p \to C(A)$, $e^Tp = 0$ ($e = \begin{bmatrix} -1/6 \\ 1/3 \\ -1/6 \end{bmatrix}$ $p = \begin{bmatrix} 7/6 \\ 5/3 \\ 13/6 \end{bmatrix}$

(fitting by a parabola: $C+Dt+Et^2=b$, $A^TA\hat{x} = A^Tb$) ($\hat{x}=\begin{bmatrix} C \\ D \\ E \end{bmatrix}$)

☆ Prove $N(A^TA)=N(A)$

suppose $A^TAx=0$, then $x^TA^TAx=0$, $\Rightarrow (Ax)^TAx=0 \xrightarrow{\text{square}} Ax=0$

(from p15. bottom)

4) Cauchy-Schwarz Inequality: $u, v \in V$. $|\langle u,v\rangle| \leq \|u\|\|v\|$

(proof. $\|u-kv\|^2 \geq 0$, with $\Delta \leq 0$)

• Orthonormal Vectors & Bases

1) orthonormal vectors / columns: $q_1, q_2, \cdots q_n$ are orthonormal if $q_i^Tq_j = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$

2) orthogonal / matrix: Q with its columns orthonormal. $Q^TQ=I$

(unitary) (If Q is square matrix, $Q^T=Q^{-1}$) → Q has n orthonormal bases

(eg. $Q=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, $Q=\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $Q=\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Adhemar construction (orthonormal with all ± 1)

$H_{2^k}=\begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}$)

3) Conside projection matrix on $C(Q)$

$P=Q(Q^TQ)^{-1}Q^T=QQ^T=\begin{cases} QQ^T, \text{usually } m>n \text{ (independence)} \\ QQ^T=I, \ m=n \text{ (square)} \end{cases}$

Good property $Q^TQ\hat{x}=Q^Tb$, $\Rightarrow \hat{x}=Q^Tb$ with $\hat{x}_i=q_i^Tb$

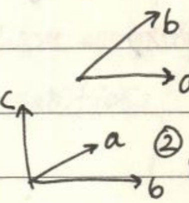
($A\to Q$ make the problem easier)

4) Gram-Schmidt Process: from A to Q ($A=QR$) $\rightarrow$ upper triangle

special case: ① vectors a & b (independent) $\rightarrow$ orthogonal A & B

fix $a=A$, $B=b-p=b-\frac{aa^T}{a^Ta}b=b-\frac{A^Tb}{A^TA}A$

$\frac{A}{\|A\|}$, $\frac{B}{\|B\|}$



② vectors a, b & c (independent) $\rightarrow$ orthonormal A, B & C

fix $a=A$, $B=b-p=b-\frac{A^Tb}{A^TA}A$

$C=c-\frac{A^Tc}{A^TA}A-\frac{B^Tc}{B^TB}B$, $\frac{A}{\|A\|}, \frac{B}{\|B\|}, \frac{C}{\|C\|}$

(Then we know the high-dimension Gram-Schmidt Process)

Why $A=QR$? ($[a_1\ a_2]=[q_1\ q_2]\begin{bmatrix} q_1^Ta_1 & q_1^Ta_2 \\ \underbrace{q_2^Ta_1}_{0} & q_2^Ta_2\end{bmatrix}$) $\longrightarrow$ by Gram-Schmidt

$\Downarrow$

$R=Q^TA$ (because $q_1=\frac{a_1}{\|a_1\|}$)

☆ Inner product spaces: Let V be a vector space, & an inner product is an operation, assigning a real number $\langle x, y\rangle$ for each pair of $x, y \in V$.

$\langle\cdot,\cdot\rangle$ satisfies ① $\langle x, x\rangle \geq 0$ with equality iff $x=0$

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

③ $\langle \alpha x+\beta y, z\rangle=\alpha\langle x, z\rangle+\beta\langle y, z\rangle$, $\forall x, y, z \in V$, $\alpha, \beta \in \mathbb{R}$

If V has an inner product operation on V, V is called the inner product space

e.g. (i) Euclidean Vector space: $\langle x, y\rangle = y^Tx$ or x^Ty

(ii) Matrix "Vector" space $\mathbb{R}^{m\times n}$ (Frobenius inner product)

(iii) Vector space $C[a, b]$ (continuous func on $[a, b]$):

$\langle f, g\rangle=$

1) Length in Inner product spaces: $v\in V$, length of $v := \|v\| = \sqrt{\langle v, v\rangle}$

$\|A\|_F=\sqrt{\langle A, A\rangle}=\sqrt{tr(A^TA)}$ or $\sqrt{tr(AA^T)}$ $\rightarrow$ Frobenius norm

2) Orthogonal in Inner product spaces: $u, v\in V$, $\langle u, v\rangle=0$

$\Rightarrow u, v$ are othogonal in V

3) Pythagorean's Law: $\|u\pm v\|^2=\|u\|^2+\|v\|^2$ when u, v orthogonal in V.

• Properties of Determinants

Determinant: the determinant of A ($\det(A)$ / $|A|$) is a number reflecting the property of matrix A (must be square matrix)

☆ Three Basic Properties.

property I: $\det(I_n) = 1, \forall n \geq 2, n \in \mathbb{N}^+$

property II: The determinant change sign under row exchange. $\Big\} \Rightarrow \det(P) = \pm 1$

property III: The determinant is a linear function for each row seperately.

i.e. $\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}$; $\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$

Corollary for other properties

property IV: 2 equal rows $\Rightarrow \det A = 0$

(use property II, $\det A = -\det A$ with row exchange)

property V: Subtract l row i from row k, DET doesnot change. $\Rightarrow$ $A = LU$, $\det A = \pm \det U$ ($\pm$: row exchange)

(use property III & IV, $\begin{vmatrix} a & b \\ c-la & c-lb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} - l\begin{vmatrix} a & b \\ a & b \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$)

property VI: a matrix with rows of zeros has DET $= 0$

(use property V $\begin{vmatrix} a & b \\ 0 & 0 \end{vmatrix} = \begin{vmatrix} a & b \\ a-a & b-b \end{vmatrix} = \begin{vmatrix} a & b \\ a & b \end{vmatrix} = 0$)

property VII: For trianglular & diagonal matrices U. $\det U =$ Products of pivots $\prod_{i=1}^{n} d_i$ (in diagonal)

(use property III & I)

☆ property VIII: $\det A = 0$ exactly when A is singular.

$\det A \neq 0$ exactly when A is invertible.

(use property VI & VII. singular $A \to U$ has rows of zeros)

☆ property IX: $\det AB = (\det A)(\det B)$ → (proof: Every matrix A → written as multiples of E (elementary) ...... use $\det(EA) = \det E \det A$)

(eg. $\det AA^{-1} = 1$, $\Rightarrow \det A^{-1} = (\det A)^{-1}$; $\det A^2 = (\det A)^2$, $\det 2A = 2^n \det A$)

☆ property X: $\det A^T = \det A$

($\det A^T = \pm\det U^T = \pm\det U = \det A$)

consider $\det L^T = \det L$

Addition From ORTHOGONALITY (P14, 15)

☆ Normed Vector spaces: V as a vector space is called the normed vector space / normed linear space if $\forall v \in V$ is associated with a real number $\|v\| \in \mathbb{R}$ (the norm of v)

$\|\cdot\|$ satisfies: ① $\|v\| \geq 0$ and equality iff $v=0$
② $\|\alpha v\| = |\alpha|\|v\|$, $\forall \alpha \in \mathbb{R}$, $v \in V$
③ $\|v+u\| \leq \|v\| + \|u\|$, $\forall v, u \in V$

(NOT unique)

(Thm) For the inner product space V, $\forall v \in V$ $\|v\| = \sqrt{\langle v, v\rangle}$ defines a norm on V

(Orthogonal set & orthonomal set)

- Formula for det(A), cofactors

1) $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} = ad\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + bc\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = ad - bc$

$2 = 2!$ terms with "±" signs (3×3 also satisfies)

$n \times n$ — $n!$ terms with "±" signs

BIG FORMULA $\det A = \sum_{n!\text{ terms}} \overset{\text{sign}}{\det P}\, a_{1\alpha} a_{2\beta} \cdots a_{n\omega}$ (different $\alpha, \beta, \cdots$) (Leibnitz formula)

$(\alpha, \beta, \gamma, \cdots, \omega) \leftarrow P =$ permutations of $(1, 2, 3, \cdots, n)$

2) cofactor — origin: take out items of $a_{i1}, a_{i2} \cdots a_{in}$, others put in brackets.

↳ Cofactor of $a_{ij} = C_{ij} = \pm \det\begin{pmatrix} n-1 \text{ matrix (square)} \\ \text{with row } i\text{, column } j \\ \text{erased} \end{pmatrix} = (-1)^{i+j}\det\cdots$

determine

So we can derive the cofactor formula

$$\det A = \sum_{k=1}^{n} a_{ik} C_{ik}$$

3) Recall $A^{-1} = \boxed{\frac{1}{ad-bc}}\begin{bmatrix} +d & -b \\ -c & +a \end{bmatrix}$ ← cofactors

(when $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$) $\frac{1}{\det A}$ then any invertible A has

$A^{-1} = \frac{1}{\det A} C^T$ ← transpose of cofactor matrix → (also called $\mathrm{adj}(A)$ adjacent matrix of A.)

(check $C^T A = (\det A) I$ — use matrix multiply)

Orthonormal matrix $QQ^T = I$, $\det Q = \pm 1$

- Matrix Derivatives (cont'd page 13) Following the defn, we have

$d(\alpha X) = \alpha dX$; $d(X+Y) = dX + dY$; $[\, d\,tr(X) = tr(X+dX) - tr(X) = tr(dX) \,]$

$d(XY) = X dY + (dX) Y$; $d(X \otimes Y) = (dX) \otimes Y + X \otimes dY$; $d(X \circ Y) = (dX) \circ Y + X \circ dY$

← Kronecker Product

Hadamard Product

the same product rule

(same page. next)

- Cramer's Rule — a neat idea to solve $Ax = b$

when A is invertible, $x = A^{-1}b = \frac{1}{\det A} C^T b$

consider $|A| x_1 = b_1 C_{11} + b_2 C_{21} + \cdots + b_n C_{n1}$ → means if $B_1 = \begin{bmatrix} b_1 & & & \\ b_2 & A_2 & \cdots & A_n \\ \vdots & & & \\ b_n & & & \end{bmatrix}$ ← b

$|A| x_1 = \det B_1^T$, $x_1 = \frac{\det B_1}{\det A}$

☆ CRAMER'S RULE $x_1 = \frac{\det B_1}{\det A}$, $x_2 = \frac{\det B_2}{\det A}$, ... $x_n = \frac{\det B_n}{\det A}$

(with B_i → ith column of A substitute by b)

(cont'd)

$dX^{-1} = -X^{-1}(dX)X^{-1}$ (comes from $0 = dI = (dX^{-1})X + X^{-1}dX$

$d|X| = |X|\, tr(X^{-1}dX)$; $d \log|X| = tr(X^{-1}dX)$

← determinant

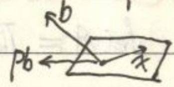
$dX^* = (dX)^*$; $*$: rearrange-ment operator

- Eigenvalues & Eigenvectors

1) Aim – to find vectors x s.t. Ax is parallel to x (i.e. $Ax = \lambda x$)

x – eigenvectors, λ – eigenvalues (multiplying factor)

Some special cases: i) If A is singular, by linear dependence, $\lambda = 0$ is an eigenvalue.

ii) For projection matrix P  $\begin{cases} x \text{ in the plane: } Px = x \ (\lambda = 1) \\ x \text{ perp. to the plane: } Px = \vec{0} \ (\lambda = 0) \end{cases}$

2) How to solve $Ax = \lambda x$?

$\Rightarrow \underbrace{(A - \lambda I)}_{\text{must be singular}} x = 0$ $\quad \det(A - \lambda I) = 0 \rightarrow$ characteristic / eigenvalue equation (polynomial)

$\begin{cases} \text{Trace: sum of } \lambda\text{'s} = \text{sum of the diagonal entries (i.e. } a_{11} + a_{22} + \cdots + a_{nn}) \\ \det(A) = \prod_{i=1}^{n} \lambda_i \text{ (product of } \lambda\text{'s)} \end{cases}$

☆ $n \times n$ square matrix has totally n eigenvalues (either real number or complex number)

If $Ax = \lambda x$, $(A + cI)x = (\lambda + c)x$, $A^2 x = \lambda A x = \lambda^2 x$. $\left(\begin{array}{l} \text{If } \det(A) \neq 0 \\ \lambda - \text{eigenvalue of } A \\ \lambda^n - \text{eigenvalue of } A^n \end{array} \right)$ $(n \in \mathbb{Z})$

(Note: $\underbrace{Ax = \lambda x;\ Bx = \alpha x}_{\text{generally not the same } x\text{ !!!}}$ $\quad (A+B)x \underset{\text{generally}}{\neq} (\lambda + \alpha)x$)

☆ Rotation Matrices (θ) $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = R$ $\quad \left(Rx = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix} \text{ (rotate } \theta) \right)$

consider $\theta = 90°$ case: $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ $\quad \det(R - \lambda I) = \lambda^2 + 1 = 0$, $\underbrace{\lambda_1 = i, \lambda_2 = -i}_{\text{complex conjugate}}$

☆ For symmetric matrix $S^T = S$, all eigenvalues real #.

For orthognal matrix $Q^T = Q^{-1}$, all eigenvalues pure imaginary #.

Degenerate Matrix:

eg. $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ $\quad \lambda_1 = 3, \lambda_2 = 3$; $\quad x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, NO 2^{nd} INDEPENDENT x.

☆ A & B share the same n independent eigenvectors iff $AB = BA$

" $\Longrightarrow$ " $A = S^{-1}\Lambda_1 S^{-1}$, $B = S^{-1}\Lambda_2 S^{-1}$, then $AB = S^{-1}\Lambda_1\Lambda_2 S^{-1} \underset{\text{property of diagonalizable}\sim}{=} S^{-1}\Lambda_2\Lambda_1 S^{-1} = BA$

" $\Longleftarrow$ " (Assume A. B has non-repeated eigenvalues)

Suppose $Ax = \lambda_A x \Rightarrow ABx = BAx = \lambda_A Bx$. Bx. x both eigenvectors ~ $\Rightarrow Bx = \lambda_B x$ $(\exists \lambda_B \in \mathbb{R})$

Proofs of $tr(A) = \sum \lambda_i$ & $det(A) = \prod \lambda_i$: (sketch)

expand it using cofactor formular

Consider characteristic equation: only $\prod(a_{ii}-\lambda)$ has $\geq n-1$ power of λ

(has the first & second highest order term). then $tr(A) = \sum a_{ii} = \sum \lambda_i$

Use another expression $P_A(\lambda) = det(A-\lambda I) = (-1)^n \prod(\lambda-\lambda_i)$. $[(-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots]$

Take $\lambda = 0$. get $\det A = \prod \lambda_i$

- Properties of Trace:

(1) $tr(AB) = tr(BA)$ for square matrices, Generally, $tr(A^TB) = tr(B^TA)$

(2) $tr(A) = tr(A^T)$ $\quad (tr(aa^T) = a^Ta)$

- Diagonalize a Matrix (Diagonalizability $\Longleftrightarrow$ Indep. Eigenvectors)

1) Suppose n independent eigenvectors of A are put in the matrix S.

promise invertible

$$AS = A\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & & | \end{bmatrix} = \underbrace{\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix}}_{S}\underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{\Lambda}$$

So we get $A = S\Lambda S^{-1}$, & $\Lambda = S^{-1}AS$ (diagnolization)

Remark: ☆ A is sure to have n independent vectors if all λ's are different.

(proof.) Suppose $c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0$, multiply by $A \Rightarrow c_1\lambda_1x_1 + c_2\lambda_2x_2 + \cdots + \lambda_nc_nx_n = 0$

multiply by λ_n to eliminate (x_n) - subtract $c_1(\lambda_1-\lambda_n)x_1 + c_2(\lambda_2-\lambda_n)x_2 + \cdots + c_{n-1}(\lambda_{n-1}-\lambda_n)x_n = 0$

Then, $\Rightarrow$ multiply by A & λ_{n-1}; A & λ_{n-2} $\cdots$ A & λ_2 with subtraction.

we have $c_1(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\cdots(\lambda_1-\lambda_n)x_1 = 0$, then $c_1 = 0$

Similarly, $c_1 = c_2 = \cdots = c_n = 0$, independent x_i.

Thus, when any matrix has non-repeated eigenvalues, it can be diagonalized.

For matrices with repeated eigenvalues // may/may not have independent eigenvectors.

2) A way to use diagonalization: (compute powers)

$A^2 = (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda^2S^{-1}$, $A^k = S\Lambda^kS^{-1}$, $\forall k \geq 2, k \in \mathbb{N}^+$

$\lim_{k\to\infty} A^k = 0$ if all $|\lambda_i| < 1$.

3) Difference Equations / One-order system:

$u_{k+1} = A u_k$, (with diagonalizable matrix A)

$\Rightarrow u_k = A^k u_0 = S\Lambda^k S^{-1} u_0$, suppose (write) $u_0 = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n = Sc$

Then $u_k = S\Lambda^k c$

Application example: Fabonacci numbers (tricks)

$u_{k+1} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}}_{= u_k}$

suppose we want F_{100}. Because $\det(A - \lambda I) = (1-\lambda)(-\lambda) - 1 = 0$, Characteristic equation

$\Rightarrow \lambda = \frac{1 \pm \sqrt{5}}{2}$, we have $x_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$

$u_k = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (\frac{1+\sqrt{5}}{2})^k & 0 \\ 0 & (\frac{1-\sqrt{5}}{2})^k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = c_1 \begin{bmatrix} (\frac{1+\sqrt{5}}{2})^{k+1} \\ (\frac{1+\sqrt{5}}{2})^k \end{bmatrix} + c_2 \begin{bmatrix} (\frac{1-\sqrt{5}}{2})^{k+1} \\ (\frac{1-\sqrt{5}}{2})^k \end{bmatrix}$

$\therefore F_k = c_1 \lambda_1^k + c_2 \lambda_2^k$, use $F_1 = 1$, $F_0 = 0$, $u_1 = c_1 x_1 + c_2 x_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ...

☆ Geometric multiplicity: $\gamma_A(\lambda_0) = \dim(\underbrace{N(A - \lambda_0 I)}_{\text{eigen-space of } \lambda_0})$ → nullspace of $A - \lambda_0 I$

Algebric multiplicity: $\mu_A(\lambda_0)$: $\underbrace{(\lambda - \lambda_0)^{\mu_A(\lambda_0)} \,\|\, P_A(\lambda)}$ → the highest order of $(\lambda - \lambda_0)$ in $P_A(\lambda)$

(but $(\lambda - \lambda_0)^{\mu_A(\lambda_0)+1} \nmid P_A(\lambda)$)

☆ $\boxed{1 \leq \gamma_A(\lambda) \leq \mu_A(\lambda) \leq n}$ $(A \in \mathbb{R}^{n \times n})$

(Diagonalizable when "=" holds!)

Explanation: $A \overset{\text{row equivalent}}{\sim} \begin{bmatrix} \lambda_0 & 0 & \cdots & 0 & a_{1,m+1} & \cdots & a_{1,m+r} \\ 0 & \lambda_0 & & \vdots & a_{2,m+1} & \cdots & a_{2,m+r} \\ \vdots & & \ddots & \lambda_0 & \vdots & & \vdots \\ 0 & \cdots & & 0 & a_{n,m+1} & \cdots & a_{n,m+r} \end{bmatrix}$ (totally $\gamma_A(\lambda_0)$)

Then $P_A(\lambda) = \det\left(\begin{bmatrix} \lambda_0 - \lambda & & 0 & a_{m+1} - \lambda & \cdots & a_{m+r} - \lambda \\ & \ddots & & \vdots & & \vdots \\ 0 & & \lambda_0 - \lambda & a_{n,m+1} - \lambda & \cdots & a_{n,m+r} - \lambda \end{bmatrix} \right) = \underbrace{(\lambda - \lambda_0)^{\gamma_A(\lambda_0)}} \cdot (-1)^{\gamma_A(\lambda_0)} \cdot \det(\ )$

$\Rightarrow \gamma_A(\lambda_0) \leq \mu_A(\lambda_0)$ (by defn)

- Systems of Differential Equations

☆ solutions for constant coefficient linear equations

eg. $\left.\begin{aligned} \frac{du_1}{dt} &= -u_1 + 2u_2 \\ \frac{du_2}{dt} &= u_1 - 2u_2 \end{aligned}\right\} \Rightarrow \frac{du}{dt} = Au$ with $A = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}$, $u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

① Eigenvalues of A: $\lambda_{1,2} = 0, -3$ with corresponding eigenvectors $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

② Solutions $u(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2$ (check in the $\frac{du}{dt} = Au$)

$u(0) = c_1 x_1 + c_2 x_2 = Sc \quad (\Rightarrow c_1 = c_2 = \frac{1}{3})$

☆ Three situations:

- Stability: $(u(t) \to 0) \Longrightarrow \mathrm{Re}(\lambda) < 0$, for $\mathrm{Im}(\lambda)$, $e^{\mathrm{Im}(\lambda)i} \to$ rotate in the unit circ DOES NOT act efficiently.
- Steady State: all λ_i's $= 0$ / some λ_i's $= 0$ with others satisfying $\mathrm{Re}(\lambda) < 0$
- Blow up: if any λ s.t. $\mathrm{Re}(\lambda) > 0$

Comment: 2×2 stability $(\lambda_1 < 0, \lambda_2 < 0)$

$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ trace $= a + d < 0$; $\det A = ad - bc > 0$

- Why this solutions? & Exponential of a matrix (cont'd)

$\frac{du}{dt} = Au$, thus $u(t) = e^{At} \cdot u(0)$

Set $A = S \Lambda S^{-1}$ (diagonalize) & $u = Sv$; $\frac{du}{dt} = S \frac{dv}{dt} = S \Lambda S^{-1} S v = S \Lambda v$

Therefore, $\frac{dv}{dt} = \Lambda v$, $v(t) = e^{\Lambda t} \cdot v(0) \Rightarrow u(t) = Sv(t) = S e^{\Lambda t} S^{-1} u(0)$

no coupling anymore (seperate u_i parts)

☆ $\Rightarrow \boxed{e^{At} = S e^{\Lambda t} S^{-1}}$

Matrix Exponential $e^{A} = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^n}{n!} + \cdots$ (convergent items) (always works!)

(another: $(I - A)^{-1} = I + A + A^2 + \cdots + A^n + \cdots$ works when $\lambda(A) < 1$)

$\Rightarrow e^{At} = I + S \Lambda S^{-1} t + \frac{S(\Lambda t)^2 S^{-1}}{2!} + \frac{S(\Lambda t)^3 S^{-1}}{3!} + \cdots + \frac{S(\Lambda t)^n S^{-1}}{n!} = S e^{\Lambda t} S^{-1}$ works when A is diagonalizable.

☆ $e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & & 0 \\ & e^{\lambda_2 t} & \\ 0 & & \ddots \; e^{\lambda_n t} \end{bmatrix}$ (good property!)

For Second order equations $\to$ 2×2 1^{st} order systems

$y'' + by' + ky = 0$; $u = \begin{bmatrix} y' \\ y \end{bmatrix}$, $u' = \begin{bmatrix} y'' \\ y' \end{bmatrix} \Rightarrow u' = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} u$

☆ k^{th} order linear $\to$ $k \times k$ 1^{st} order systems with $A = \begin{bmatrix} - & - & \cdots & - \\ 1 & & & \\ & 1 & & 0 \\ 0 & & \ddots & \\ & & & 1 \end{bmatrix}$

- Kronecker Products:

(1) [Defn] $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes B = \begin{bmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{bmatrix}$

like ordinary matrix product, associative, distributive, NOT commutative.

(2) Properties:

- Markov matrices / Stochastic matrices (applied with MC).

Markov matrices: 1) All Entries ≥ 0 2) all columns add to one

$$\text{eg.} \begin{bmatrix} .1 & .99 & .3 \\ .2 & .01 & .3 \\ .7 & 0 & .4 \end{bmatrix}$$

steady state (for difference equations)

$\begin{cases} \lambda = 1, \text{ corresponding } x_1 \geq 0 \\ \text{all other } \lambda_i\text{'s s.t. } |\lambda_i| < 1 \end{cases}$ (det)

Consider $u_{k+1} = Au_k$, then $u_k = c_1x_1 + c_2x_2\lambda_2^k + \cdots + c_nx_n\lambda_n^k$

(always approaches $u_\infty = c_1x_1$)

☆ Eigenvalues of A & A^T are the same.

$0 = \det(A - \lambda I) = \det(A^T - \lambda I)$ (reason)

$\Rightarrow$ For Markov matrices, $A - I$ must be singular because all rows add to zero.

$(1, 1, 1)$ in $N(A^T) \Rightarrow x_1$ in $N(A)$.

NOT necessary because $A^T\mathbb{1} = \mathbb{1}$

Proof. Equivalent to prove $(A - I)x = 0$, $x \geq 0$ has non-zero solutions

consider minimize$_x$ c^Tx subject to $(A-I)x = 0$, $x \geq 0$ (with $c < 0$). its dual problem maximize$_y$ $0^Ty = 0$ subject to $(A^T - I)y \leq c$

Suppose y^* satisfies $(A^T - I)y^* \leq c$, with y_k be the smallest entry.

$(\{y_i \mid y_i \leq y_j \ \forall j\} = y_k)$ Because $(A^T - I)\cdot m\mathbb{1} = 0$ with $m \in \mathbb{R}^1$ $\mathbb{1} \in \mathbb{R}^{n\times 1}$, we can move y^* to $y^* - y_k\mathbb{1}$ still satisfying the inequality. Consider the k^{th} row of $(A^T - I)(y^* - y_k\mathbb{1})$ it adds up to a number ≥ 0, contradicts! So the dual problem is infeasible.

Because $x = 0$ is a solution of the primal, it must be unbounded (duality theory)

$\Rightarrow$ It has non-zero x s.t. $Ax = x$, $x \geq 0$ (eigenvalue 1)

Applications of Markov matrices

$\begin{bmatrix} u_{cal} \\ u_{mass} \end{bmatrix}_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$ $\quad u_{k+1} = \underbrace{A}_{\text{populatio matrix}} u_k$ $\left\{ \begin{array}{l} \text{four entries – fraction of populations} \\ \geq 0 \text{ – only positive all involved (probability)} \\ \Sigma = 1 \text{ – all people in a state} \end{array} \right.$

consider people movement between

California $\underset{0.2}{\overset{0.1}{\rightleftarrows}}$ Massachusetts

$\begin{bmatrix} u_{cal} \\ u_{mass} \end{bmatrix}_{t=k+1} = \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} \begin{bmatrix} u_{cal} \\ u_{mass} \end{bmatrix}_{t=k} \Rightarrow$ $\begin{array}{ll} x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} & \lambda_1 = 1 \text{ (steady state)} \\ x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} & \lambda_2 = 0.7 \end{array}$

• Kronecker Products: (cont'd. page 23)

$\left\{ \begin{array}{l} (A \otimes B)^T = A^T \otimes B^T \ ; \ (A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \leftarrow \text{different from ordinary.} \\ (A \otimes B)(C \otimes D) = AC \otimes BD \end{array} \right.$

(3) For square A & B. ei-values & ei-vectors of $A \otimes B$ is given by

$$A \otimes B = S_A \Lambda_A S_A^{-1} \otimes S_B \Lambda_B S_B^{-1} = (S_A \otimes S_B)(\Lambda_A \otimes \Lambda_B)(S_A \otimes S_B)^{-1}$$

$$\Rightarrow \left\{ \begin{array}{l} \text{rank}(A \otimes B) = \text{rank}(A)\,\text{rank}(B) \\ tr(A \otimes B) = tr(\Lambda_A \otimes \Lambda_B) = tr(A)\,tr(B) \\ |A \otimes B| = |A|^{\text{rank}(B)} |B|^{\text{rank}(A)} \end{array} \right.$$

• Application: Fourier series (for functions)

1) (Def.) The vector $v = (v_1, v_2 \cdots)$ & the function $f(x)$ are in our infinite-demensional Hilbert spaces iff their length are finite.

$\|v\| = v^T v = \sum_{i=1}^{\infty} v_i^2$; $\|f\| = \langle f, f \rangle = \int_0^{2\pi} |f(x)|^2 dx$

2) How to define inner product in functions?

vectors: $\langle v, w \rangle = v^T w = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$

functions $\langle f, g \rangle = \int_0^{2\pi} f(x) g(x) dx$ (0–2π can be changed to other intervals)

3) Fourier series

$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots$ (periodic functions)

It has orthogonal basis!

Get the Fourier Coefficients. (recall if $v = x_1 q_1 + x_2 q_2 + \cdots + x_n q_n = Qx$, $x = Q^T v$)
orthonormal $x_1 = q_1^T v$

(eg.) $a_1 = \dfrac{\int_0^{2\pi} f(x)\cos x\, dx}{\int_0^{2\pi} \cos x^2\, dx}$ (because of orthogonal)

$\Rightarrow a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\, dx$, $b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin x\, dx$ $(k \geq 1)$

Fourier transform (continuous) $\hat{f}(\alpha) = \int_{-\infty}^{\infty} f(t) e^{i\alpha t} dt$, denoted as $F(f)$

$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\alpha) e^{i\alpha t} d\alpha$, denoted as $F^{-1}(\hat{f})$

Properties: $F(f'(\alpha)) = i\alpha F(f(\alpha))$; $F(f * g) = F(f)F(g)$

Convolution (卷积) $f_1(t) * f_2(t) := \int_{-\infty}^{\infty} f_1(\tau) f_2(t-\tau) d\tau$

- Vec: define $vec\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{bmatrix}$ as a stacked column of a matrix.

$vec(ABC) = (C^T \otimes A)\, vec(B)$

- Symmetric matrix (all real numbers)

1) Properties of symmetric matrix
① Eigenvalues are REAL
② Eigenvectors can be chosen PERPENDICULAR

Proof: ① $Ax = \lambda x \xrightarrow{\text{conjugate}} \bar{A}\bar{x} = \bar{\lambda}\bar{x} \Leftrightarrow A\bar{x} = \bar{\lambda}\bar{x} \xrightarrow{\text{transpose}} \bar{x}^T A^T = \bar{x}^T \bar{\lambda} \Leftrightarrow \bar{x}^T A = \bar{x}^T \bar{\lambda}$

$\Rightarrow \underbrace{\bar{x}^T A x = \lambda \bar{x}^T x \ \&\ \bar{x}^T A x = \bar{\lambda}\bar{x}^T x}_{\text{same}}$ so $\underbrace{\lambda = \bar{\lambda}}_{\text{real number}}$ ($\bar{x}^T x$ not zero) magnitude of x

② Suppose $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$ (assume $\lambda_1 \neq \lambda_2$)

$(\lambda_1 x_1)^T x_2 = (Ax_1)^T x_2 = x_1^T A^T x_2 = x_1^T A x_2 = \lambda_2 x_1^T x_2$, so $x_1^T x_2(\lambda_1 - \lambda_2) = 0$

$\Rightarrow x_1^T x_2 = 0$, orthogonal. $\Rightarrow$ (can choose orthogonomal by normalization)

2) For symmetric matrices $A = Q\Lambda Q^T$ (coming from $A = S\Lambda S^{-1}$)

spectrum theorem / principal axis theorem

$A = Q\Lambda Q^T = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots \Rightarrow$ every symmetric matrix is a combination of perpendicular projection matrices.

3) Pivots & eigenvalues

For matrices which is symmetric $\begin{cases} \text{signs of pivots} = \text{signs of eigenvalues} \\ \text{\# pos pivots} = \text{\# pos eigenvalues} \end{cases}$

• Hardamard Product

(1) [Defn] $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \circ \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$

★ Schur's product thm:

$A \succeq 0,\ B \succeq 0 \Rightarrow \begin{cases} A \circ B \succeq 0 \\ A \otimes B \succeq 0 \end{cases}$ diff. from ordinary "·"

(2) Basic Properties. (use $\mathrm{diag}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}$.)

......

• Complex Matrix & FFT

1) Real vs Complex

(i) Length $\|x\|_2^2 = x^T x,\ x \in \mathbb{R}^n$; $\|z\|_2^2 = \bar{z}^T z = z^H z,\ z \in \mathbb{C}^n$ (z^H - Hermitian)

(ii) Transpose $(A^T)_{ij} = A_{ji}$; conjugate transpose $(A^H)_{ij} = \bar{A}_{ji}$ (A^H - Hermitian)

(iii) Inner Product $\langle x, y\rangle = x^T y,\ x, y \in \mathbb{R}^n$; $\langle x, y\rangle = \bar{x}^T y = x^H y,\ x, y \in \mathbb{C}^n$

(iv) Symmetric: $A^T = A$; Hermitian $A^H = A$

(v) Orthogonality: $Q^T Q = I$; $\bar{U}^T U = U^H U = I$ (unitary matrix)

(vi) Diagonalization: $S = Q \Lambda Q^T$; $S = U \Lambda U^H$ (both real Λ)

2) The Fast Fourier Transform (FFT)

(i) Aim - multiply quickly by F & F^{-1} (the Fourier matrix)

(ii) $n \times n$ Fourier matrix (general form)

$(F_n)_{ij} = w^{ij}$, where $w^n = 1$, $w = e^{\frac{2\pi}{n} i}$ (primitive root)

$i, j = 0, 1, 2, \cdots, n-1$

↑ start from 0^{th} row & column in EE.

$$F_n = \underbrace{\begin{bmatrix} 1 & 1 & & 1 \\ 1 & \omega & & \omega^{n-1} \\ \vdots & \omega^2 & \cdots & \omega^{2(n-1)} \\ 1 & \omega^{n-1} & & \omega^{(n-1)^2} \end{bmatrix}}$$ is symmetric.

orthogonal columns $\Rightarrow F^{-1} = \frac{1}{n} F^H$

$\frac{1}{\sqrt{n}} F_n$ - orthonormal bases

(iii) One step in FFT, (eg. $n=4$)

$$F_4 = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{array}\right]$$ because $\omega_8^2 = (e^{\frac{\pi}{4}i})^2 = (e^{\frac{\pi}{2}i}) = i = \omega_4$.

The idea.

$$F_8 = \begin{bmatrix} I_4 & D_4 \\ I_4 & -D_4 \end{bmatrix} \begin{bmatrix} F_4 & 0 \\ 0 & F_4 \end{bmatrix} \left[\text{Permutation (even-odd)} \right], \text{ with } D_n = \begin{bmatrix} 1 & & & 0 \\ & \omega & & \\ & & \omega^2 & \\ 0 & & & \ddots \ \omega^{n-1} \end{bmatrix}$$

$\omega = \omega_{2n}$

$\begin{bmatrix} IF_4 & DF_4 \\ IF_4 & -DF_4 \end{bmatrix} \longrightarrow$ need column exchange

$\begin{bmatrix} 1 & & & & 0 \\ & & 1 & & \\ & 1 & & & \\ & & & 1 & \end{bmatrix} \longrightarrow$ change even-odd columns

Set $m = \frac{1}{2}n$, the first & last m components of $y = F_n c$, $c \in \mathbb{R}^n$ combines the half size transform of $y' = F_m c'$ & $y'' = F_m c''$ (split c to c' & c'')

$$\underbrace{y_j = y_j' + (\omega_n)^j y_j'' \ ; \ y_{j+m} = y_j' - (\omega_n)^j y_j''}_{\text{fix-up cost} \Rightarrow 2m} \quad (j = 0, \cdots\cdots m-1)$$

(iv) Full FFT by Recursion

eg. $$F_{16} = \left[\ \cdots \ \right] \begin{bmatrix} \begin{smallmatrix} I & D \\ I & -D \end{smallmatrix} & 0 \\ 0 & \begin{smallmatrix} I & D \\ I & -D \end{smallmatrix} \end{bmatrix} \begin{bmatrix} \begin{smallmatrix} F_4 & 0 \\ 0 & F_4 \end{smallmatrix} & 0 \\ 0 & \begin{smallmatrix} F_4 & 0 \\ 0 & F_4 \end{smallmatrix} \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \left[P \right]$$

$(4m)^2 \to 2 \times (2m)^2 + \underset{2m}{\text{fix-up cost}} = 2 \times (2m^2 + m)$

Change the original $O(n^2)$ to FFT $O(n \log n)$.

• Interlacing Eigenvalue Theorem

(Thm) Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ & let $\vec{z}$ be a vector with $\|\vec{z}\| = 1$, $\alpha \in \mathbb{R}$. Then if we denote the eigenvalues of $A + \alpha \vec{z}\vec{z}^T$ (rank 1 update) by $\xi_1 \geq \xi_2 \geq \xi_3 \geq \cdots \geq \xi_n$, then $\boxed{\text{for } \alpha > 0}$

$$\xi_1 \geq \lambda_1 \geq \xi_2 \geq \lambda_2 \geq \cdots \geq \xi_n \geq \lambda_n \quad \text{with } \sum_{i=1}^{n} (\xi_i - \lambda_i) = \alpha$$

& If $\alpha < 0$, then $\lambda_1 \geq \xi_1 \geq \lambda_2 \geq \xi_2 \geq \cdots \geq \lambda_n \geq \xi_n$

• Positive Definite Matrix

1) Definition & three tests

$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

① $\lambda_1>0$, $\lambda_2>0$ eigenvalue test (all eigenvalue's >0)

② $a>0$, $ac-b^2>0$ determinant test (all subdeterminant >0) leading principal submat

③ $a>0$, $\frac{ac-b^2}{a}>0$ pivot test (all pivots >0)

④ * $x^TAx>0$, except of $x=0$ energy test / definition

Quandratic form

2) Graph of x^TAx – ellipsoid

$f(x,y)=\vec{x}^TA\vec{x}=ax^2+2bxy+cy^2$

For high-dimension x (eg. 3) Three axis – direction of eigenveltors

Length: values of eigenvalues

3) Test for minimum: In Linear Algebra: MIN ~ MATRIX OF 2nd DERIVS IS POS DEF

Additions: Pos def S, T gives a pos def (S+T)

Full rank A gives a pos def A^TA

Pos def A gives a pos def A^{-1} (because of eigenvalues pos)

Quadratic Form: x^TAx ($= x^T(\frac{A+A^T}{2})x$) just consider symmetry case

• Conditioning & Stability

1. [Defn] Condition number: for a * non-singular matrix (A), condition number $K(A) := \|A\|\cdot\|A^{-1}\|$, with diff. norms, $K_1(\cdot), \cdots K_\infty(\cdot) \cdots$

2. [Defn] A problem is well-conditioned if its soln will NOT be affected by small pertubations. (otherwise, ill-conditioned) (esp. for $Ax=b$,

$$\frac{\|x-\tilde{x}\|}{\|x\|} \doteq K(A)\left(\frac{\|A-\tilde{A}\|}{\|A\|}+\frac{\|b-\tilde{b}\|}{\|b\|}\right))$$

3. [Defn] An algorithm is said to be stable if it's guaranteed to produce accurate answers to all well-conditioned problems in its class.

• Similar matrices

1) A, B are called similar matrices, which means, for some M, $B=M^{-1}AM$
(eg. $S^{-1}AS=\Lambda$, means A is similar to Λ → Λ is the best matrix similar to A)

2) Similar matrices have the same eigenvalues & the same # of eigenvectors.
(proof. Suppose $Ax=\lambda x$, $\Rightarrow M^{-1}A\underbrace{MM^{-1}}_{I}x = M^{-1}\lambda x = \lambda M^{-1}x$
$\Rightarrow \underbrace{(M^{-1}AM)}_{B}(M^{-1}x)=\lambda(M^{-1}x)$, $\Rightarrow \lambda$ is an eigenvalue of B.)
A fixed matrix C produce a family of similar matrices BCB^{-1}, allowing all invertible B

3) BAD CASE: suppose a 2x2 square matrix with $\lambda_1=\lambda_2=4$
one small family has $\begin{bmatrix}4&0\\0&4\end{bmatrix}$, because $M^{-1}\begin{bmatrix}4&0\\0&4\end{bmatrix}M = 4M^{-1}IM=4I=\begin{bmatrix}4&0\\0&4\end{bmatrix}$
one big family includes $\begin{bmatrix}4&1\\0&4\end{bmatrix}$ ← Jordan Form (the most diagonal one)

⇒ ☆ Jordan Form: for every A, want B s.t. $B^{-1}AB$ is as nearly diagonal as possible.
① A has n independent eigenvectors, $\Lambda=S^{-1}AS$ is the Jordan Form of A
② A has s ($s<n$) independent eigenvectors, then
the Jordan Form $J=B^{-1}AB=\begin{bmatrix}\boxed{J_1}&&&0\\&\boxed{J_2}&&\\&&\ddots&\\0&&&\boxed{J_s}\end{bmatrix}$ # indep eigenvectors = # blocks
with every Jordan Block $J_i=\begin{bmatrix}\lambda_i&1&&0\\&\ddots&\ddots&\\0&&&\lambda_i\end{bmatrix}$ (one eigenvalue λ_i, 1's above)

Matrices are similar if they share the same Jordan Form — not otherwise
(NOT vice versa)

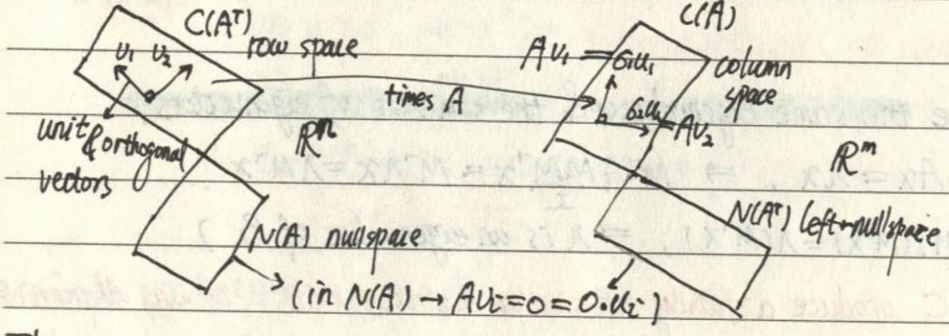
Addition: a special kind of similar matrices.
(L: a linear transformation $V\to V$, E, F are 2 bases of V
S – transition matrix correspondences to $[x]_E$, $[x]_F$ ($[x]_E=S[x]_F$)
(Page 8)
A – matrix representation of L w.r.t E. B, …… F.
Then $B=S^{-1}AS$. (B is similar to A.) → E as the basis for both domain & co-domain

In reality, by results of changing basis (Page 32) $\begin{cases}A=E^{-1}PE\\B=F^{-1}PF\end{cases}$ ($B_{out}^{-1}\cdots B_{in}$) with P – understand basis
$\Rightarrow B=\underbrace{(F^{-1}E)}_{S^{-1}}\underbrace{(E^{-1}PE)}_{A}\underbrace{(E^{-1}F)}_{S}$ (matrix case)

- The singular Value Decomposition (SVD)

1) Aim: decomposite a matrix $A \longrightarrow U\Sigma V^T$ $\begin{cases} A: m\times n \\ U: m\times m;\ V: n\times n \longrightarrow \text{(orthonomal basis with)} \\ \Sigma: m\times n \text{ squared. diagonal} \end{cases}$

$\Rightarrow$ means $AV = U\Sigma$



Then, suppose $v_1, v_2, \cdots v_r$ is an orthonormal basis in $C(A^T)$

$v_{r+1}, v_{r+2}, \cdots v_n$ is an orthonormal basis in $N(A)$

$u_1, u_2 \cdots u_r$ is an orthonormal basis in $C(A)$

$u_{r+1}, u_{r+2} \cdots u_m$ is an orthonormal basis in $N(A^T)$

$$AV = A\begin{bmatrix} v_1 & v_2 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_r & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} = U\Sigma$$

$\Rightarrow$ SVD: $A = U\Sigma V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T$ $(r = \text{rank}(A))$

(Special case – full rank – $A = Q\Sigma Q^T$ (symm pos def))

2) Proof & Construct of the SVD

For v's: $A^TA = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T\Sigma V^T$, means that v's are eigenvectors of A^TA (orthonormal)
>singular vector

For u's: $AA^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma\Sigma^T U^T$, means that u's are eigenvectors of AA^T

For σ's: $\sigma^2 = \lambda(A^TA)$ (totally m) (square root of eigenvalues of A^TA)
└ singular value

3) Rayleigh Quotient & norms

(i) suppose S is symmetric then $\max_x \frac{x^TSx}{x^Tx} = \lambda_{max}(S)$ (the largest eigenvalue of S)

proof: S has eigenvalues. suppose $\lambda_1, \lambda_2, \cdots \lambda_n$ with $\lambda_1 = \max_{1\le i\le n}\{\lambda_i\}$, & orthonomal eigenvectors $q_1, q_2 \cdots q_n$. Then every x can be written as:

$$\frac{x^TSx}{x^Tx} = \frac{(c_1q_1+c_2q_2+\cdots+c_nq_n)^T(\lambda_1c_1q_1+\cdots+\lambda_nc_nq_n)}{(c_1q_1+\cdots+c_nq_n)^T(c_1q_1+\cdots+c_nq_n)} = \frac{\sum_{i=1}^n \lambda_i c_i^2}{\sum_{i=1}^n c_i^2} \le \frac{\lambda_1\sum_{i=1}^n c_i^2}{\sum_{i=1}^n c_i^2} = \lambda_1$$

($x = c_1q_1+c_2q_2+\cdots+c_nq_n$, with $c_i \in \mathbb{R}$, then)

Let $x_1 = mq_1$, $m\in\mathbb{R}$ we can get $\frac{x^TSx}{x^Tx} = \lambda_1$, therefore $\max_x \frac{x^TSx}{x^Tx} = \lambda_{max}(S)$

(ii) norm of a matrix (Induced 2-norm / L-2 norm) $\|A\| := \max_x \frac{\|Ax\|}{\|x\|} = \delta_{max}(A) = \sqrt{\lambda_{max}(A^TA)}$ (the largest singular value of A)

proof: $\max_x \frac{\|Ax\|}{\|x\|} = \max_x \sqrt{\frac{\|Ax\|^2}{\|x\|^2}} = \max_x \sqrt{\frac{x^T(A^TA)x}{x^Tx}} \overset{(i)}{=} \sqrt{\lambda_{max}(A^TA)} \overset{SVD}{=} \delta_{max}(A)$

4) Principal Component Analysis (PCA by SVD)

Consider a data matrix $A_0 \in \mathbb{R}^{m\times n}$, with n samples & m measuring variables

$\Rightarrow$ Center the data – substract each row by the mean μ_i from row i (with sum=0) ← centered
↳ get centered A

☆ The "Sample covariance matrix" (2D) is $S = \frac{AA^T}{n-1}$ (with $m=2$)

$S_{1,1}$ & $S_{2,2}$ – variance, $S_{1,2} = S_{2,1}$ covariance

Essentials of PCA:

☆ Total variance $T = \sum_{i=1}^m \delta_i^2 = \sum_{i=1}^m S_i^2 = \text{trace}\, \frac{AA^T}{n-1}$ (δ_i: singular value)

☆ SVD of $A = \sum_{i=1}^r \delta_i u_i v_i^T$, i^{th} eigenvector u_i of S (i^{th} singular vector of A) accounts for a fraction $\frac{\delta_i^2}{T}$ of the total variance.

☆ Stop when those fractions are small (suppose $\delta_1^2 \ge \delta_2^2 \ge \cdots$), totally R directions that explain most of the data. (u_1 – the most significant direction of data) with R – "effective rank" of A, & u's are principal components of A in $\mathbb{R}^m$.

☆ The sample correlation matrix $C = \frac{DAA^TD}{n-1}$, D – a diagonalized matrix, with DA's each row having length $\sqrt{n-1}$.

Applications: Genetic Variation, Face recognition (Eigenfaces), Google page-ranking, & Finance

(Prove of "equivalence" of tests for pos. def. for real matrix)

① ⇒ ④ Since real symmetric matrix s.t. $A = \sum_i \lambda_i q_i q_i^T$

④ ⇒ ① Use corresponding orthogonal ei-vectors

① + ④ ⇒ ② (④ ⇒ ① ⇒ ②) Use $[x_k; \vec{0}]$ ⇒ all ei-values for leading sub-matrix > 0 ⇒ $\det = \Pi_i \lambda_i > 0$

② ⇒ ③ trivial. ③ ⇒ ④ Gauss-elimination $A = LDL^T \Rightarrow (L^T x)^T D (L^T x) > 0$.

Thm: $L: V \to W$ is injective

iff $\ker(L) = \{0_v\}$

• Linear Transformations

Def: Kernal ($L: V \to W$) → L is a linear tranformation

$\ker(L) = \{v \in V \mid L(v) = 0_W\}$ → subspace of V; $\vec{0}$ in W

Def: range/image "$L(V)$" → subspace of W

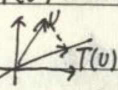
$\text{Range}(L(V)) = \text{Range}(L) = \{w \in W \mid \exists v \in V \text{ s.t. } L(v) = w\}$

1) Idea of linear transformation

Rules (the transformation is linear if it meets)

$T(v+w) = T(v) + T(w)$; $T(c \cdot v) = cT(v)$ for all c, v, w

(linear combination $T(cv + dw) = cT(v) + dT(w)$)

eg.1 projection $T: \mathbb{R}^2 \to \mathbb{R}$ (mapping) 

(non-eg.1, Shift whole plane by v_0 — non-linear because $T(0) = v_0 \neq 0$)

(non-eg.2, $T(v) = \|v\|$; $T: \mathbb{R}^3 \to \mathbb{R}$, non-linear because $T(-v) \neq -T(v)$)

eg.2 Rotation by θ, $T: \mathbb{R}^2 \to \mathbb{R}$

eg.3 Given matrix A, $T(v) = Av$

$L: V \to W$ — L: linear operator; V: domain; W: co-domain

☆ Derivative $T(f) = \frac{df}{dx}$ is linear

2) The Matrix of a Linear Transformation

☆ Construct "matrix of T" (A, represents lin. tr. $T: \mathbb{R}^n \to \mathbb{R}^m$)

(every linear transformation T can be assigned by a matrix A.)

① choose basis $v_1, v_2, \ldots v_n$ in the input space $\mathbb{R}^n$ (get inputs)

Columns of 1 to n will contains those out put $T(v_1)$ to $T(v_n)$

⇒ For every $v = \underbrace{c_1 v_1 + c_2 v_2 + \cdots + c_n v_n}_{\text{coordinates}}$, $T(v) = \sum_{i=1}^n c_i T(v_i) =$

$\begin{bmatrix} T(v_1) & T(v_2) & \cdots & T(v_n)\end{bmatrix}\begin{bmatrix}c_1\\c_2\\ \vdots \\ c_n\end{bmatrix}$ ($T(v_i)$ — $m\times 1$ columns)

② Further step: choose $w_1, w_2, \cdots w_m$ as a basis in $\mathbb{R}^m$ (output space) (get outputs)

Rules to get A: $T(v_i)$ is the combination $\underline{a_{1i}}w_1 + \underline{a_{2i}}w_2 + \underline{a_{3i}}w_3 + \cdots + \underline{a_{mi}}w_m$ (items of A)

Then we get $A\begin{pmatrix}\text{input}\\ \text{coords}\end{pmatrix} = \begin{pmatrix}\text{output}\\ \text{coords}\end{pmatrix}$

(some tips: Let $\begin{bmatrix} w_1 & w_2 & \cdots & w_m\end{bmatrix} A = \begin{bmatrix} T(v_1) & \cdots & T(v_n)\end{bmatrix}$ → get $Ac = [\ ]$ → output coords of w)

☆ change basis: $WB = V \Rightarrow B = W^{-1}V$; but $Vc = Wd$, know c, $d = W^{-1}Vc = Bc$

($V=\begin{bmatrix} T(v_1) & \cdots & T(v_n)\end{bmatrix}$) when $V=I$, $B=W^{-1}$, meaning that Large basis vectors have smaller coefficients.

(standard basis is changed to be another one)

☆ If A is a matrix for transformation T (in the standard basis), then $\boxed{B_{out}^{-1}AB_{in}}c = d$

$B_{out}^{-1}AB_{in}$ is the new matrix in new basis. $\Rightarrow T(B_{in}c) = B_{out}d$, $T(B_{in}c) = AB_{in}c$

We can choose B_{in} & B_{out} to make the matrix simple. (eg. Λ, Σ, Jordan form)

[take coordinates] ☆ $T(v) = \overset{I_m}{W}Ac = I_m A \boxed{I_n c} = Av$

e.g. (From PAGE 8) change basis of "$[\ ]_x$" (linear transformation operator)

Let U, V be matrices with i^{th} row being u_i / $[u_i]^T_{standard}$, v_i / $[v_i]^T_{standard}$, respectively

For general vector space (such as functions, matrices)

$[\ ]_v$ is a linear transformation with standard in/out bases (I), with the matrix V^{-1} (represent)

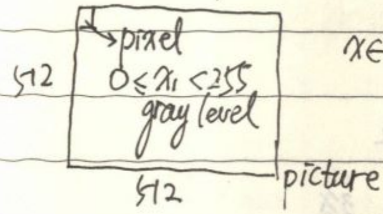
Now change $B_{in} = U$, $B_{out} = I$, we get new matrix $= V^{-1}U$.

($\Rightarrow [\ ]_u = V^{-1}U[\ ]_v$ / or Let A be $V^{-1}U$, i^{th} col of A — $A_i = [u_i]_v$)

the same expression

• Application: Image compression (intro)

IDEA in compression of image: throw away some small #s does NOT change heavily.



$x \in \mathbb{R}^{512^2}$

JPEG is the standard compression.

→ break the big matrix into 8x8 blocks

Video — sequence of images which are correlated.

Compression & Save by changing of basis.

original – standard basis $\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}\cdots\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}$

→ better basis $\begin{bmatrix}0\\\vdots\\0\end{bmatrix}\begin{bmatrix}1\\\vdots\\-1\\\vdots\\-1\end{bmatrix}\cdots\begin{bmatrix}1\\-1\\\vdots\\1\\-1\end{bmatrix}$ → high frequency

zero frequency

half one / half -one

good basis $\xrightarrow[\text{FWT}]{\text{FFT}}$ good compression (few basis is enough)

wavelet

☆ Wavelet basis

$$\begin{bmatrix}1\\\vdots\\\vdots\\1\end{bmatrix}\begin{bmatrix}1\\\vdots\\1\\-1\\\vdots\\-1\end{bmatrix}\begin{bmatrix}1\\1\\-1\\-1\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1\\1\\-1\\-1\end{bmatrix}\underbrace{\begin{bmatrix}1\\-1\\0\\0\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\1\\-1\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1\\-1\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\-1\end{bmatrix}}_{\text{less important}}$$

(or use Fourier basis)

$p = Wc$, $c = W^{-1}p$

process

signal p
↓ change basis ← $c = W^{-1}p$
coeffs c
↓
thresholding (throw away small #s)
↓
coeffs $\hat{c}$ (many zeros)
↓
reconstruct $\hat{x} = \sum \hat{c}_i v_i$

• Correlation = Function Space.

Because for basis (common) $1, x, x^2, \dots$ it is not orthogonal.

just satisfies $\int_{-1}^{1} x^{odd} x^{even} dx = \langle x^{odd}, x^{even} \rangle = 0$

☆ Three leading even-odd bases

① Fourier basis $1, \sin x, \cos x, \sin 2x, \cos 2x, \dots$

② Legendre basis $1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x, \dots$

③ Chebyshev basis $1, x, 2x^2 - 1, 4x^3 - 3x, \dots$

Comments

② — gotten by $\boxed{\text{Gram-schmidt}}$ $\begin{cases} \dfrac{\langle x^2, 1\rangle}{\langle 1, 1\rangle} = \dfrac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} dx} = \dfrac{1}{3}, \Rightarrow x^2 - \dfrac{1}{3} \\ \dfrac{\langle x^3, x\rangle}{\langle x, x\rangle} = \dfrac{\int_{-1}^{1} x^4 dx}{\int_{-1}^{1} x^2 dx} = \dfrac{3}{5}, \Rightarrow x^3 - \dfrac{3x}{5} \end{cases}$

③ — comes from $1, \cos\theta, \cos 2\theta, \dots$

• The Pseudo-inverse A^+ [Moore-Penrose Inverse]

1) 2-sided inverse. $AA^{-1}=I=A^{-1}A$ (r=m=n; squared, full rank matrix)

2) One-sided inverse.

(i) full column rank. $r=n<m \Rightarrow$ nullspace $=\{0\}$; indep. cols; 0 or 1 soln to $Ax=b$.

Has one-side inverse $A^{-1}_{left}=(A^TA)^{-1}A^T$ (because A^TA invertible)

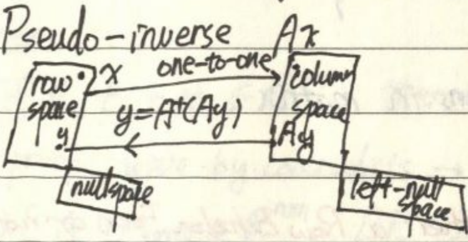
$A^{-1}_{left}A=I$; $AA^{-1}_{left}=A(A^TA)^{-1}A^T \rightarrow$ projection matrix on the col space.

(ii) full row rank: $r=m<n \Rightarrow N(A^T)=0$; indep. rows; infinit solns to $Ax=b$ (n-m free varibles)

Has one-side inverse $A^{-1}_{right}=A^T(AA^T)^{-1}$ (because AA^T invertible)

$AA^{-1}_{right}=I$; $A^{-1}_{right}A=A^T(AA^T)^{-1}A \rightarrow$ projection matrix on the row space.

3) Pseudo-inverse



☆ If $x\neq y$ both in row space, the $Ax\neq Ay$ in column space.

proof. suppose $Ax=Ay$. $\Rightarrow A(x-y)=0$

$\Rightarrow x-y$ in the nullspace, contradict!

$\rightarrow$ in row space $\Rightarrow x=y$

Pseudo-inverse used in linear regression.

How to find A^+?

From the SVD: $A=U\Sigma V^T$, $\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots \ \sigma_r & \\ 0 & & 0 \end{bmatrix}_{m\times n,\ \text{rank } r} \xrightarrow{\Sigma^+} \begin{bmatrix} \sigma_1^{-1} & & 0 \\ & \ddots \ \sigma_r^{-1} & \\ 0 & & 0 \end{bmatrix}_{n\times m,\ \text{rank } r}$ (transpose)

$A^+=V\Sigma^+U^T$

$A^+A=\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \rightarrow$ projection matrix onto $C(A^T)$

$$\begin{cases} \Sigma^+\Sigma_{n\times n}=\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ \Sigma\Sigma^+_{m\times m}=\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \end{cases}$$

MAT 2040 Addition (mainly some definitions)

• Linear System

(Equivalent systems) Systems which have the same solution set.

(Equation Operations) $\begin{cases} \text{swap two equations } R_i \leftrightarrow R_j \\ \text{multiple non-zero constant } R_i \leftrightarrow \frac{1}{\alpha} R_i \\ \text{"linear transformation" } R_j \leftrightarrow -\beta R_i + R_j \end{cases}$

↓ equivalent under these operations

• Matrix (def.)

$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & & a_{mn} \end{bmatrix} \triangleq \boxed{(a_{ij})_{m\times n}}$ $\begin{cases} \forall a_{ij} \in \mathbb{R} \text{ (all)} \text{ — real matrix} \\ \exists a_{ij} \in \mathbb{C} \setminus \mathbb{R} \text{ — complex matrix} \end{cases}$

(Row operations) similar to equation operations in matrix.

☆(Def) Row echelon form (how to know whether a Row Echelon Form or not)

1) If row k has non-zero entries (at least 1), # leading 0s in row k+1 is greater than # leading 0s in row k

2) Rows with all entries = 0 are below rows with some non-zero entries.

☆(Def.) Reduced row echelon form (rref)

1) First be row echelon form

2) The first non-zero entry in each non-zero row is 1 (leading 1)

3) Each leading 1 is the only non-zero entry in its column

☆(Def) consistence

$Ax=b$ (a linear system) is consistent if it has at least one solution.

↳ otherwise inconsistent

Homogeneous system: $Ax=0$; ∞ # solutions when A is underdetermined.

(left of this page, p35)

★ Elementary matrix (one row operation from I)

3 types of row operations $\to$ 3 types of elementary matrices

Remarks. (i) Doing elementary row operations $\Longleftrightarrow$ Multiply corresponding elementary matrices

(Foundamentals of LU decomposition)

(ii) Elementary matrizes are invertible.

$$\begin{cases} E_{R_iR_j}^{-1} = E_{R_iR_j}^{T} & \text{(permutation matrices)} \\ E_{\alpha R_i}^{-1} = E_{\frac{1}{\alpha}R_i}, & \text{with } \alpha \in \mathbb{R}\setminus\{0\} \\ E_{\beta R_i + R_j}^{-1} = E_{-\beta R_i + R_j} \end{cases}$$

Theorems:

1) If A, B both $\in \mathbb{R}^{n\times n}$, then AB is invertible $\boxed{\text{exactly when}}$ A, B both invertible.

(proof: argue by contradicts $\to$ prove B non-singular, then A non-singular)

2) If A, B both $\in \mathbb{R}^{n\times n}$, $AB = I_n$, then $BA = I_n$ (just use one side - enough)

- Vector spaces: (with defns of "addition" & "multiplication")

Eight axioms satisfied — V is a vector space. $(\forall u, v \in V)$

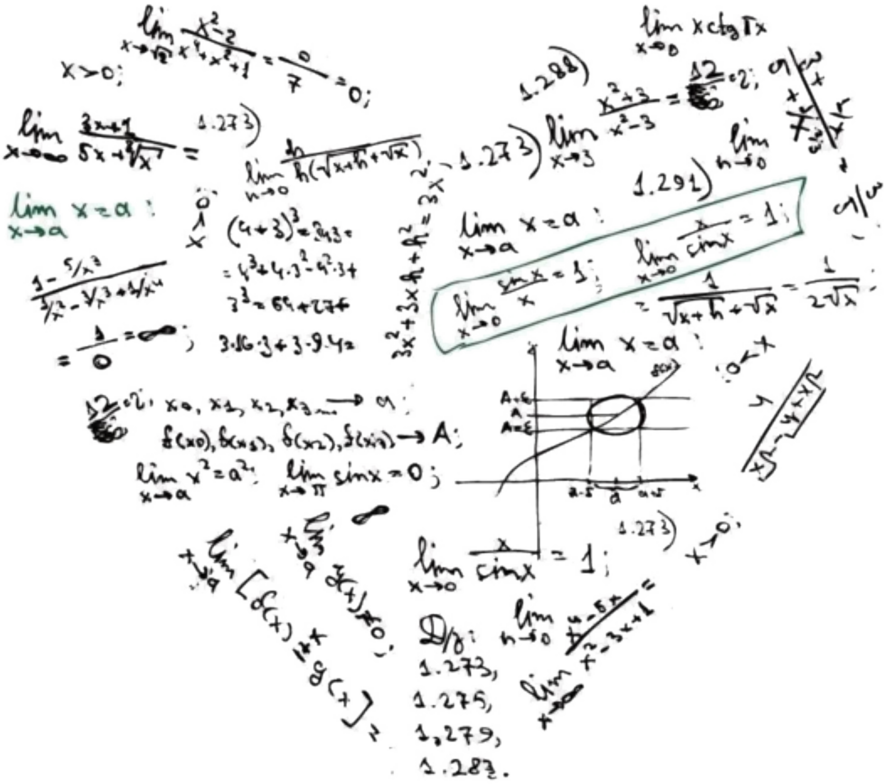
$$\begin{cases} ① u+v=v+u;\ ② u+(v+w)=(u+v)+w=u+v+w; \\ ③ \exists \text{ element } 0, \text{ s.t. } u+0=0+u=u;\ ④ \exists -u\in V, \text{ s.t. } -u+u=0 \\ ⑤ \alpha(u+v)=\alpha u+\alpha v;\ ⑥ (\alpha+\beta)u=\alpha u+\beta u \\ ⑦ \alpha(\beta u)=(\alpha\beta)u;\ ⑧ 1\cdot u=u. \end{cases}$$

not empty

eg. $C[a,b]$ (collection of all continuous funcs on $[a,b]$) is a space.

- Linear Transformation ($L: V \to W$)

3 properties
$$\begin{cases} L(0_V) = L(0(\text{in } v)) = 0_W \\ \text{Under Linear combinations } L(\alpha_1 x_1 + \dots + \alpha_n x_n) = \sum_{i=1}^{n} \alpha_i L(x_i) \\ L(-u) = -L(u) \end{cases}$$



Advanced Algebra

MAT 3040 Notebook

Youthy WANG

Advanced Linear Algebra (MAT3040, CUHKSZ)

- Vector spaces

1. (Defn) Let F be a field (e.g. $\mathbb{R}$ or $\mathbb{C}$, etc). A vector space V over F is a set V with 2 well-defined operators: $\begin{cases} \text{"}+\text{"}: V+V \to V \text{ (addition)} \\ \text{"}\bullet\text{"}: F\cdot V \to V \text{ (scalar multiplication)} \end{cases}$

satisfying
$$\begin{cases} \text{(i) } \underline{v}+\underline{w} = \underline{w}+\underline{v};\ \underline{v},\underline{w}\in V \\ \text{(ii) } (\underline{u}+\underline{v})+\underline{w} = \underline{u}+(\underline{v}+\underline{w}),\ \underline{u},\underline{v},\underline{w}\in V \\ \text{(iii) There is a } \underline{0}\in V \text{ s.t. } \underline{0}+\underline{v}=\underline{v},\ \forall \underline{v}\in V \\ \text{(iv) There is always } -\underline{u}\in V \text{ s.t. } -\underline{u}+\underline{u}=\underline{0},\ \forall \underline{u}\in V \\ \text{(v) } \alpha\cdot(\underline{u}+\underline{v}) = \alpha\cdot\underline{u}+\alpha\cdot\underline{v};\ \forall\alpha\in F,\ \underline{u},\underline{v}\in V \\ \text{(vi) } (\alpha+\beta)\cdot\underline{v} = \alpha\cdot\underline{v}+\beta\cdot\underline{v},\ \forall\alpha,\beta\in F \\ \text{(vii) } (\alpha\beta)\cdot\underline{v} = \alpha\cdot(\beta\cdot\underline{v}) \\ \text{(viii) } 1\cdot\underline{v}=\underline{v} \end{cases}$$

(see similar axioms in MAT2040)

operator (diff from previous +)

(1) examples of vector space. (eg.1) $V = F^n$ (with $F=\mathbb{R}$ or $\mathbb{C}$, etc) & "+": $\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}+\begin{bmatrix}\beta_1\\ \vdots\\ \beta_n\end{bmatrix} = \begin{bmatrix}\alpha_1+\beta_1\\ \vdots\\ \alpha_n+\beta_n\end{bmatrix}$; "•": $\alpha\cdot\begin{bmatrix}\beta_1\\ \vdots\\ \beta_n\end{bmatrix} = \begin{bmatrix}\alpha\beta_1\\ \vdots\\ \alpha\beta_n\end{bmatrix}$

with "$\underline{0}$" $:= \begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}$

(eg.2) $V = M_{m\times n}(F) = \{$all $m\times n$ matrices with entries in $F\}$

"+": $\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mn}\end{bmatrix} + \begin{bmatrix}b_{11} & \cdots & b_{1n}\\ \vdots & \ddots & \vdots\\ b_{m1} & \cdots & b_{mn}\end{bmatrix} = \begin{bmatrix}a_{11}+b_{11} & \cdots & a_{1n}+b_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn}\end{bmatrix}$; "•": $\alpha\cdot\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mn}\end{bmatrix} = \begin{bmatrix}\alpha a_{11} & \cdots & \alpha a_{1n}\\ \vdots & \ddots & \vdots\\ \alpha a_{m1} & \cdots & \alpha a_{mn}\end{bmatrix}$

with $\underline{0} = \begin{bmatrix}0 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix}_{m\times n}$ $(O_{m\times n})$

(eg.3) $V = C^\infty(\mathbb{R}^n) = \{$all smooth functions $f: \mathbb{R}^n \to \mathbb{R}\}$. "+": $\underbrace{f+g}_{\triangleq f(x)+g(x)} \in V$

& "•" $\underbrace{\alpha\cdot f}_{\triangleq \alpha f(x)} \in V$, with "$\underline{0}$": $\underline{0}(x) \equiv 0,\ \forall x\in V$ (constant-zero function)

$\alpha\in\mathbb{R}$ but NOT $\in\mathbb{C}$

(eg.4) Polynomials: $V = P_n(\mathbb{R}) = \{a_0 + \cdots + a_n x^n \mid a_i \in F\}$

(2) Vector Subspace (Defn)

Let V be a vector space on F. A subset $W \subseteq V$ is a vector subspace of V if $(W, +, \cdot)$ is itself a vector space.

(Thm) $W \subseteq F^n$ is a vector subspace iff $W = \mathrm{span}_F\{v_1, \dots v_k\}$. (Convention: If $S = \emptyset$, $W = \mathrm{span}_F S = \{\underline{0}\}$).

(Thm') Let V be a vector space over F. Then $W \subseteq V$ (subset of, generally) is a subspace of V iff $\forall w_1, w_2 \in W$, $\alpha, \beta \in F$, (we have) $\alpha w_1 + \beta w_2 \in W$.

(e.g.5) $V = M_{n\times n}(F)$, $\begin{cases} W_1 = \{\text{all invertible matrices}\}\ (GF(n, F))\ (\not\leq V) & \underline{0} \notin W \\ W_2 = \{\text{all singular matrices}\}\ (\not\leq V) & \begin{bmatrix}1&0\\0&0\end{bmatrix} + \begin{bmatrix}0&0\\0&1\end{bmatrix} \notin W \end{cases}$ (counter-e.g.)

Sets that can $\leq V$: {all symmetric matrices}, {all upper triangular matrices}, ...

(e.g.6) $V = C^\infty(\mathbb{R})$, $W = \{f \in V \mid f' = 2f\} \leq V$ (subspace of) → proof by defn

(Thm'')
- Suppose $W_1, W_2 \leq V$. Then, $W_1 \cap W_2 \leq V$ (easy proof)
  (More generally, with $W_i \leq V$, $i \in I$, we get $\bigcap_{i\in I} W_i \leq V$)

Note that $W_1 + W_2$ $(:= \{\underline{w}_1 + \underline{w}_2 \mid \underline{w}_1 \in W_1, \underline{w}_2 \in W_2\})$ is also a subspace.
(More generally, $\sum_{i\in I} W_i \leq V$)

(3)(Defn) Let $W_1, \dots W_k \leq V$, suppose the following holds:
$\exists\, \underline{w}_1 \in W_1, \dots, \underline{w}_k \in W_k$ s.t. $\underline{w}_1 + \dots + \underline{w}_k = \underline{0}$, then it must be $\underline{0} = \underline{w}_1 = \dots = \underline{w}_k$.
We then say $W_1 \oplus \dots \oplus W_k := W_1 + \dots + W_k$ is a direct sum.
(Intuition: NO redundancies in the direct sum of vector subspaces.)

2. Spanning Set

(1)(Defn) Let V be a vector space over F, $S \leq V$ is a (NOT necessarily finite) subset of V. Then, a linear combination of elements in S is of the form $\alpha_1 \underline{s}_1 + \alpha_2 \underline{s}_2 + \dots + \alpha_k \underline{s}_k$, with $\alpha_i \in F$, $\underline{s}_i \in S$.
must be a finite sum

The linear span (of S) then, is defined as the collection of all linear combinations of S.

i.e. $\text{span}_F(S) = \{ \sum_{i=1}^{k} \alpha_i \underline{s_i} \mid \alpha_i \in F,\ \underline{s_i} \in S,\ k \in \mathbb{N} \}$. We say S spans V, or S is the spanning set of V, if $\text{span}_F(S) = V$.

(proposition) $S \subseteq V$ $\boxed{\text{exactly when}}$ $\text{span}_F(S) \leq V$. (easy proof)

(proposition')

Suppose $S \subseteq V$, then $\begin{cases} (a)\ S \leq \text{span}(S) \\ (b)\ \text{span}(S) = \text{span}(\text{span}(S)) \end{cases}$ (skipped proof)

3. Linear Independence

(1) (Defn) Let V be a vector space, we say a subset $S \subseteq V$ is linearly independent (l.i.) in V if $\forall$ choice of $\boxed{\text{finite}}$ subset $\{\underline{s_1}, \cdots, \underline{s_n}\} \subseteq S$, only soln for $\alpha_1 \underline{s_1} + \alpha_2 \underline{s_2} + \cdots + \alpha_n \underline{s_n} = \underline{0}$ is $\alpha_1 = \cdots = \alpha_n = 0$, with $\alpha_i \in F$. !!

☆ (Note: when $S := \{\underline{s_1}, \cdots, \underline{s_k}\}$ is itself finite, same defn as 2040.)

(e.g.1) $V = \mathbb{F}[x]$, $S = \{1, x, x^2, x^3, \cdots\}$ ← infinite set.

(proof by comparing the coefficients)

(e.g.2) $V = C^\infty(\mathbb{R})$, $S = \{\cos x, \sin x\}$ (functional result)

(2) (Defn) Let V be a vector space, a subset $S \subseteq V$ is $\boxed{\text{a basis}}$ of V if (i) S spans V ; (ii) S is linearly independent.

(e.g.3) $V = \mathbb{F}[x]$, $S = \{1, x, x^2, \cdots\}$ is a basis of V.

or $V = C^\infty(\mathbb{R})$, $S = \{1, x, x^2, \cdots\}$ is NOT a basis of V.

$\sin(x) = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}(-1)^k \in V$ ← NOT a finite sum

→ but NOT $\in \text{Span}(S)$

(e.g.4) Suppose $\{\underline{v_1}, \underline{v_2}, \underline{v_3}\}$ is a basis of V, then $\{\underline{v_1} - \underline{v_2}, \underline{v_2} - \underline{v_3}, \underline{v_3}\}$ is also a basis of V. [construction of a new basis]

(3) (proposition) $S \subseteq V$ is a basis if and only if all non-zero elements $\underline{v} \in V$ can be expressed as $\underline{v} = \alpha_1 \underline{S_1} + \cdots + \alpha_k \underline{S_k}$, for some $\underline{s_i} \in S$, $\alpha_i \in F$ $\boxed{\text{UNIQUELY}}$ $(\alpha_i \neq 0)$

☆ Remark: uniqueness means if $\underline{v} = \alpha_1 \underline{S_1} + \cdots + \alpha_k \underline{S_k}$, with $\underline{S_i} \in S$
$= \beta_1 \underline{S_1'} + \cdots + \beta_\ell \underline{S_\ell'}$, with $\underline{S_j'} \in S$

then $k = \ell$, $\alpha_i = \beta_i$, $\forall 1 \le i \le k$ & $\underline{S_i'} = \underline{S_i}$, $\forall 1 \le i \le k$. (w.l.o.g. suppose after right arrangement)

(proof.) ("$\Rightarrow$") Since S spans V, all $\underline{v} \in V$ can be expressed as $\underline{v} = \alpha_1 \underline{S_1} + \cdots + \alpha_k \underline{S_k}$, $\underline{S_i} \in S$

For uniqueness, suppose $\alpha_1 \underline{S_1} + \cdots + \alpha_k \underline{S_k} = \beta_1 \underline{S_1'} + \cdots + \beta_\ell \underline{S_\ell'}$ $(= \underline{v})$, then uniqueness holds, for otherwise $\alpha_1 \underline{S_1} + \cdots + \alpha_k \underline{S_k} + (-\beta_1) \underline{S_1'} + \cdots + (-\beta_\ell) \underline{S_\ell'} = 0$, by getting together similar items $\exists$ some $\underline{S_i}$ (or S_i'), whose coefficient is non-zero, which means S is NOT l.i., contradictory!

("$\Leftarrow$") Suppose $\underline{v}$ can be expressed as $\alpha_1 \underline{S_1} + \cdots + \alpha_k \underline{S_k}$, then $\text{span}_F(S) = V$, & suppose $\exists\, \alpha_1 \underline{S_1} + \cdots + \alpha_m \underline{S_m} = \underline{0}$, with $\alpha_i \neq 0$ ($\exists$ such finite subset of S), consider vector $\underline{v}' = \beta_1 \underline{S_1} + \cdots + \beta_m \underline{S_m}$, then $\underline{v} = \underline{v} + \underline{0} = \sum_{i=1}^{m} (\alpha_i + \beta_i) \underline{S_i}$

"S is l.i." $\Longleftarrow$ contradictory with uniqueness

Remark: If $|S| < \infty$, the argument will be more simple (only use l.i. once).

(4) (Defn) Let V be a vector space, we say V is finitely generated, if $V = \text{span}_F(S)$, for some $\boxed{\text{finite set } S}$

(e.g.5) F^n, $M_{m \times n}(F)$, etc. are f.g. but $F[x]$ is NOT f.g.

(proposition) All finitely generated subspaces have a basis.

(proof.) By defn, $V = \text{span}(\{\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}\})$ for some $n \in \mathbb{N}$.

If $\{v_1, \ldots, v_n\}$ is l.i., it satisfies to be a basis. Otherwise, $\exists$ non-trivial $\alpha_1 v_1 + \alpha_2 \underline{v_2} + \cdots + \alpha_n \underline{v_n} = \underline{0}$ (with NOT all zeros $\alpha_1, \ldots, \alpha_n$). W.l.o.g., Let $\alpha_1 \neq 0$, then $\underline{v_1} = \left(-\frac{\alpha_2}{\alpha_1}\right) v_2 + \cdots + \left(-\frac{\alpha_n}{\alpha_1}\right) v_n$. then $v_1 \in \text{span}(\{\underline{v_2}, \ldots, \underline{v_n}\})$ precisely, $V = \text{span}(\{\underline{v_2}, \ldots, \underline{v_n}\})$

Continously doing it (stop if the "redundant" set is l.i., can stop because $\{\underline{v_n}\}$ is l.i.). ☒

☆ Remark: All vector spaces (including not f.g. cases) have bases. (axiom of choice $\Rightarrow$). Nonetheless, it's generally not simple (or impossible) to write down a basis for any arbitrary non-f.g. space explicitly.

Qn: Any basis with the same size for f.g. vector space? → dimension! (same finite number)

(5) (Thm) Any basis for f.g. vector space V have the same finite number of elements. (We call this number dimension of V.)

(proof.) Since V is f.g., V has bases. Suppose we have 2 bases $S_1=\{\underline{v_1},\dots,\underline{v_m}\}$ & $S_2=\{\underline{u_1},\dots,\underline{u_n}\}$ of V. W.l.o.g, let $m\le n$. By defn of basis, $\exists\, a_{ij}\in\mathbb{F}$ $1\le i\le n$, $1\le j\le m$ s.t. $\sum_{k=1}^{m} a_{ik}\underline{v_k}=\underline{u_i}$, $\forall\, 1\le i\le n$. Consider matrix $A\triangleq A_{ij}=a_{ij}$. then $\operatorname{rank}(A)\le m$, when $n>m$, rows are linearly dependent. Thus, $\exists\,\beta_1,\dots\beta_n$ s.t. Not all zeros s.t. $\sum_{i=1}^{n}\beta_i A_i^T=0^T_{n\times 1}$, which means $\sum_{i=1}^{n}\beta_i\underline{u_i}=\sum_{i=1}^{n}\beta_i A_i^T\underline{V}=\underline{0}$.
$\therefore$ Only $n=m$ is a possible choice.

(defn) Dimension: Let V be a vector space over $\mathbb{F}$. The dimension of V is defined by $\dim_{\mathbb{F}}(V)=\begin{cases} m, & \text{if } V \text{ is f.g. \& } m \text{ be the size of basis}\\ \infty, & \text{if } V \text{ is NOT f.g.}\end{cases}$

(e.g.) $\dim(\mathbb{F}^n)=n$; $\dim(M_{m\times n}(\mathbb{F}))=mn$; $\dim(\mathbb{F}[x])=\infty$ & $\dim(\{A\in M_{3\times 3}(\mathbb{F})\mid A^T=-A\}$ (skew-symmetric matrices) $)=3$.
→ One basis $\left\{\begin{bmatrix}0&1&0\\-1&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\-1&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&1\\0&-1&0\end{bmatrix}\right\}$

Qn: $\{\underline{w_1},\dots \underline{w_k}\}$ l.i. in V, can add some $\{\underline{v_1},\dots \underline{v_r}\}$ s.t. it's a basis? → generally true

(Thm) Basis Extension:
Let V be finite dimensional n. $\forall\,\{\underline{w_1},\dots,\underline{w_k}\}$ l.i. $\exists$ vectors $\{\underline{v_{k+1}},\dots,\underline{v_n}\}$ s.t. $\{\underline{w_1},\dots,\underline{w_k},\underline{v_{k+1}},\dots,\underline{v_n}\}$ is a basis.

(proof.) Take a basis $\{\underline{u_1},\dots \underline{u_n}\}$ of V, consider $\{\underline{w_1},\dots \underline{w_k},\underline{u_1},\dots,\underline{u_n}\}\triangleq S_1$
Then S_1 spans V.

If S_1 is l.i., done, otherwise $\alpha_1\underline{w}_1+\cdots+\alpha_k\underline{w}_k+\beta_1\underline{u}_1+\cdots+\beta_n\underline{u}_n=\underline{0}$, Not all α & βs are 0. Thus, $\exists\,\beta_j\neq 0$, assume $\beta_1 \Rightarrow \underline{u}_1\in\text{span}\{\underline{w}_1,\cdots,\underline{w}_k,\underline{u}_2,\cdots,\underline{u}_n\}$. (similar to previous case)
By "reduncance", we can get some $\underline{u}_i$'s, denoted as $\underline{v}_{k+1},\cdots \underline{v}_n$ (due to 'dimension') ...... ∎

(Thm) Complementation. Let $\dim(V)=n<\infty$ & $W\leq V$. Then $\exists\, W'$ s.t. $W'\leq V$ & $W\oplus W'=V$.

(Proof.) Let $\{\underline{w}_1,\cdots,\underline{w}_k\}$ be basis of W, $(k\leq n)$ (i.e. $\dim(W)=k\leq n$)

By basis extension, $\exists\{\underline{w}_1,\cdots,\underline{w}_k\}\cup\{\underline{w}_{k+1},\cdots,\underline{w}_n\}$ s.t. it forms a basis of V

Let $W'=\text{span}\{\underline{w}_{k+1},\cdots,\underline{w}_n\}\leq V$. Then, $W+W'=V$ (obviously by defn of basis) & $W\cap W'=\{\underline{0}\}$ (for otherwise $\alpha_1\underline{w}_1+\cdots+\alpha_k\underline{w}_k+(-\beta_1)\underline{w}_{k+1}+\cdots+(-\beta_{n-k})\underline{w}_n=\underline{0}$, contradictory with l.i. of basis). Then $W\oplus W'=V$.

• Linear Transformations & Vector Expressions

1. (1) Definitions of linear transformation:

(i) (Defn) Let V, W be 2 vector spaces of $\mathbb{F}$, a linear transformation $T: V\to W$ is a [mapping] from V to W satisfying ($T(a\underline{v}_1+b\underline{v}_2)=aT(\underline{v}_1)+bT(\underline{v}_2)$), $\forall a,b\in\mathbb{F}$, $\underline{v}_1,\underline{v}_2\in \underline{V}$.

(ii) E.g. 1 Indeed, All linear transformations $\mathbb{F}^n\to\mathbb{F}^m$, always can be represented by $T(\underline{x})=A\underline{x}$, with $A\in M_{m\times n}(\mathbb{F})$.

E.g. 2 Let $T_{1,2}: C^\infty(\mathbb{R})\to C^\infty(\mathbb{R})$, $T_1(f)=f^{(n)}$ n-th order derivatives; $T_2(f)=\int_a^x f(t)dt$ are both linear transformations.

Remarks: ① If target space $W=V$, linear transformation $T: V\to V$ is also called a [linear operator.]

② $T(\underline{0}_V)=\underline{0}_W$ for all linear transformations.

③ If $\left\{\begin{array}{l} T: V\to W \\ S: W\to U\end{array}\right.$ then $S\circ T: V\to U$ is also linear.
linear transformations

E.g. 3 $T: M_{n\times n}(\mathbb{R}) \to \mathbb{R}$, $T(A) = tr(A)$ (function)

(Note that $\det(A)$ is NOT a linear transformation $\forall n > 1$)

E.g. 4 $T: M_{m\times n}(\mathbb{R}) \to M_{k\times n}(\mathbb{R})$, $T(A) = CA$, $C \in M_{k\times m}(\mathbb{R})$

(2) Kernel & Image (Range):

(i) (Defn) Let $T: V \to W$ be linear transformation

The kernel of T is $\ker(T) = \{\underline{v} \in V \mid T(\underline{v}) = \underline{0}_W\} = T^{-1}(\{\underline{0}_W\})$

The Image of T is $Im(T) = \{\underline{w} \in W \mid \exists \underline{v} \in V \text{ s.t. } T(\underline{v}) = \underline{w}\} = T(V)$

(ii) E.g.s ① $T: \mathbb{R}^n \to \mathbb{R}^m$, $T(\underline{v}) = A\underline{v}$; $\ker(T) = Null(A)$ & $Im(T) = Col(A)$

② $T: C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, $T(f) = f''$; $\ker(T) = \{ax+b;\ a,b \in \mathbb{R}\} = span\{x, 1\}$

$Im(T) = C^\infty(\mathbb{R})$

(iii) Propositions: Let $T: V \to W$ is a linear transformation

Then $\ker(T) \leq V$ & $Im(T) \leq W$

Also, T is injective $\iff \ker(T) = \{\underline{0}_V\}$

T is surjective $\iff Im(T) = W$

(proof is too obvious to be omitted.)

If $T: V \mapsto W$ is bijective, we call it an isomorphism

☆ Remarks: for linear transformation, knowing $\{T(\underline{v}_1), \cdots, T(\underline{v}_n)\}$ with a basis $\{\underline{v}_1, \cdots, \underline{v}_n\}$, we can determine the function T.

[proposition'] If $T: V \mapsto W$ is an isomorphism, [$T^{-1}: W \mapsto V$ is also a linear transformation].

(proof.) $aT^{-1}(\underline{w}_1) + bT^{-1}(\underline{w}_2) \overset{\exists \underline{v}_1, \underline{v}_2}{=\!=} a\underline{v}_1 + b\underline{v}_2 = T^{-1}(\underbrace{T(a\underline{v}_1 + b\underline{v}_2)}) \overset{(T^{-1}\circ T)[a\underline{v}_1+b\underline{v}_2]}{}$ = (linear combinations)

$T^{-1}(aT(\underline{v}_1) + bT(\underline{v}_2)) = T^{-1}(a\underline{w}_1 + b\underline{w}_2)$. □ $(\forall a, b \in \mathbb{F},\ \underline{w}_1, \underline{w}_2 \in W)$

(iv) (Rank-Nullity Thm): Let $T: V \mapsto W$ be a linear transformation, $\dim(V) < \infty$. Then, $\underbrace{\dim(\ker(T))}_{\text{nullity of } T,\ \leq V} + \underbrace{\dim(Im(T))}_{\text{rank of } T,\ \leq W} = \dim(V)$

(consistence with matrix cases: $\dim(N(A)) = \dim(\ker(T))$)

(proof.) Since $\ker(T) \leq V$, by complementation, $\exists V_1 \leq V$ s.t. $\dim(\ker(T)) \oplus V_1 = V$. Now it becomes $\dim(V_1) = \dim(V) - \text{rank}(T)$.

Consider $T|_{V_1} : V_1 \mapsto W$ (only vectors in V_1), $T|_{V_1}(\underline{v_1}) := T(\underline{v_1})$, $\forall \underline{v_1} \in V_1$

restricted map. Note that $T(V_1) = \text{Im}(V_1) \leq W$. ($T|_{V_1} : V_1 \to T(V_1)$ true)

* Claim 1: $T|_{V_1}$ is an isomorphism. (Take $\underline{v_1}, \underline{v_2} \in V_1$, $T(\underline{v_1} - \underline{v_2}) = \underline{0}$, $\Rightarrow \underline{v_1} - \underline{v_2} \in \ker(T)$, with $\ker(T) \cap V_1 = \{\underline{0}_V\} \Rightarrow \underline{v_1} - \underline{v_2} = \underline{0}_V$, meaning that NO 2 diff vectors give the same value.)

Thus, $\dim(V_1) = \dim(T(V_1))$.

* Claim 2: $T(V_1) = \text{Im}(T)$. (Take $\underline{w} \in \text{Im}(T)$, $\exists \underline{v} \in V$ s.t. $T(\underline{v}) = \underline{w}$, if $\underline{w} \neq \underline{0}_W$, $\underline{v} \notin \ker(T)$, since $\underline{v} = \underline{v_1} + \underline{v_2}$, for $\underline{v_1} \neq \underline{0} \in V_1$ ($\underline{v} \notin \ker(T)$) & $\underline{v_2} \in \ker(T)$. $T(\underline{v_1}) = T(\underline{v}) - T(\underline{v_2}) = T(\underline{v})$ which means $\underline{w} \in T(V_1)$. If $\underline{w} = \underline{0}_W$, $\underline{0}_V \in V_1$ s.t. $T(\underline{0}_V) = \underline{0}_W$, thus, $\underline{0}_W \in T(V_1)$. Then, $\text{Im}(T) \subseteq T(V_1)$. Obviously, $T(V_1) = \text{Im}(V_1) \subseteq T(V) = \text{Im}(T)$, thus $T(V_1) = \text{Im}(T)$.) $\square$

2. Coordinate Vectors

(1) [Defn] Let V be a vector space over $\mathbb{F}$ with $\dim(V) = n < \infty$. Consider $\mathcal{B} = \{\underline{v_1}, \dots, \underline{v_n}\}$ be an ordered basis of V. Then, the coordinate vector of $\underline{v}$ is given by $\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ with $\underline{v} = \alpha_1 \underline{v_1} + \dots + \alpha_n \underline{v_n}$ $\xrightarrow{\text{denoted as}} [\underline{v}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{F}^n$.

Remark: Here ordered $\mathcal{B}$ means if $\mathcal{B}'$ has the same vectors but NOT the same order as $\mathcal{B}$, $\mathcal{B}' \neq \mathcal{B}$.

☆[proposition] The function $\phi : V \mapsto \mathbb{F}^n$ given by $\underline{v} \mapsto \phi(\underline{v}) = [\underline{v}]_{\mathcal{B}}$ is an isomorphism of vector spaces.

(e.g. checking l.i.: $\left\{ \begin{matrix} x + 3x^2 + 4x^3 \\ 1 + 5x^2 \\ 20 + x + 10x^3 \end{matrix} \right\} \to \left\{ \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 20 \\ 1 \\ 0 \\ 10 \end{bmatrix} \right\}$ is l.i. in $\mathbb{R}$)

general $\Rightarrow \exists$ tools for checking l.i. by isomorphism

(proof.) ϕ is a function, for basis has unique expression. This uniqueness definitely gives injectiveness ~~(with ordered basis)~~. Surjectiveness is too simple to be proved here.

(2) Changing Basis

(e.g. $V = P_2(\mathbb{R})$, $\underline{v} = 3+2x+4x^2$, with $B = \{1, x, x^2\}$ & $B' = \{x+x^2, 1+x, 1\}$

$[\underline{v}]_B = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$; $[\underline{v}]_{B'} = 3[1]_B + 2[x]_B + 4[x^2]_B = \begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix}$.)

(i) [Defn] Let B & B' be ordered basis of V. (Denoted as $\{\underline{v}_1, \dots, \underline{v}_n\}$ & $\{\underline{u}_1, \dots, \underline{u}_n\}$). Suppose $\underline{u}_j = \sum_{i=1}^{n} \alpha_{ij} \underline{v}_i$, $\forall j = 1, \dots, n$. The change of basis matrix is $C_{BB'} = (\alpha_{ij})_{n\times n}$ (meaning that $[\cdot]_B = C_{BB'}[\cdot]_{B'}$) ☆ [from $B' \to B$]

(ii) [Proposition] $C_{BB'}$ is invertible & $C_{BB'}^{-1} = C_{B'B}$ ↖ to prove this — (i) for $\underline{u}_j$ (basis) — (ii) for general $\underline{v}$

(proof.) $C_{BB'} C_{B'B} [\underline{v}]_B = C_{BB'} [\underline{v}]_{B'} = [\underline{v}]_B$. Thus, $C_{BB'} C_{B'B} = I_n$ (similarly, $C_{B'B} C_{BB'} = I_n$.) By defn, we know $C_{BB'}^{-1} = C_{B'B}$. □

(3) Matrix Rrepresentation of a Linear Transformation.

(i) [Defn] Let $T: V \mapsto W$ be a linear transformation, $A = \{\underline{v}_1, \dots, \underline{v}_n\}$ is an ordered basis of V & $B = \{\underline{w}_1, \dots, \underline{w}_m\}$ is an ordered basis of W.

The matrix representation of T w.r.t. A & B is given by $(T_{BA}) = (\beta_{ij})_{m\times n}$

with entries from $\begin{cases} T(\underline{v}_1) = \beta_{11}\underline{w}_1 + \dots + \beta_{m1}\underline{w}_m \\ \vdots \\ T(\underline{v}_n) = \beta_{1n}\underline{w}_1 + \dots + \beta_{mn}\underline{w}_m \end{cases}$ $\left(\begin{bmatrix} [T(\underline{v}_1)]_B & \cdots & [T(\underline{v}_n)]_B \end{bmatrix}_{m\times n} \right)$

(ii) Compositions of 2 Linear Trans.

v.sp. $V^n \xrightarrow{T} W^m \xrightarrow{S} U^p$

bases A B C

☆ $T_{BA}[\underline{v}]_A = [T(\underline{v})]_B$ ↳ to prove this — (i) for $\underline{v}_i$ — (ii) for general $\underline{v}$

(Thm) Functoriality. $\boxed{(S\circ T)_{CA} = S_{CB} \cdot T_{BA}}$ ← matrix multiplications

(proof.) Since $T_{BA}[\underline{v}]_A = [T(\underline{v})]_B$ & $S_{CB}[\underline{w}]_B = [S(\underline{w})]_C$ (from defn)

we get $[\underbrace{S\circ T(\underline{v})}_{w}]_C = [S(T(\underline{v}))]_C = S_{CB}[T(\underline{v})]_B = S_{CB}\cdot T_{BA}[\underline{v}]_A$

By defn, $(S\circ T)_{CA} = S_{CB}\cdot T_{BA}$. □

(iii) Change Basis in Linear Transformations

Given L.T. $T: V \mapsto V$ & Two bases of V: $\mathcal{A}$, $\mathcal{B}$.

Then, $(T)_{\mathcal{BB}} \cdot \mathcal{C}_{\mathcal{BA}} = (T)_{\mathcal{BA}}$ (more generally, $\mathcal{C}_{\mathcal{AB}} (T)_{\mathcal{BB}} \mathcal{C}_{\mathcal{BA}} = (T)_{\mathcal{AA}}$) recall in PAGE 33.

☆ $(T)_{\mathcal{AA}}$ is [similar] to $(T)_{\mathcal{BB}}$.

properties → { same characteristic polynomials, same eigen-values, same deteminant, diagnolizability (same) }

- Remark: To define the determinant of a linear operator $T: V \mapsto V$ by picking one basis $\mathcal{B}$ of V (randomly) & $\det(T) := \det(T_{\mathcal{BB}})$ ← NOT change with $\mathcal{B}$

• Quotient Spaces

1. Defn & Motivations

(1) IDEA: "patition" a big vector space V into union of smaller spaces.

(2) [Defn] Let V be a vector space, $W \leq V$. For any $\underline{v} \in V$ define a coset (with representative $\underline{v}$) by $\boxed{\underline{v} + W} \triangleq \{ \underline{v} + \underline{w} \mid \underline{w} \in W \} \subseteq V$

(Note, even if $\underline{v} \neq \underline{v}' \in V$, it's possible to have $\underline{v} + W = \underline{v}' + W$, when $\boxed{\underline{v} - \underline{v}' \in W}$) iff; 2 equal coset

(easily proved.)

One application of cosets: solve $A\underline{x} = b$ ($\underline{x}_{\text{particular}} + \text{Null}(A)$)

(3) [Defn] Quotient Space. Let $W \leq V$, the quotient space V/W is defined as the collection of all W-cosets $V/W := \{ \underline{v} + W \mid \underline{v} \in V \}$

Define a vec-space structure on V/W: ($\mathbb{F}$, "+", "·")

"+": $(\underline{v} + W) + (\underline{u} + W) := \boxed{(\underline{u} + \underline{v}) + W}$. & "·": $\alpha \cdot (\underline{u} + W) := \boxed{\alpha \underline{u} + W}$.

However, Note that $\exists\, \underline{u} \neq \underline{u}' \in V$ s.t. $\underline{u} + W = \underline{u}' + W$ (still well-defined $\mathbb{F}$, "+", "·" easily verified)

2. Theorems about Quotient Spaces

(1) (i) (Thm 1) Let V be a vec-space, $W \leq V$. Define a map $\pi: V \to V/W$ by $\pi(\underline{v}) := \underline{v} + W$. Then, π is a surjective linear transformation. (called canonical projection form)

(ii) $\ker(\pi) = W \leq V$ (Note that $0_{V/W} = \underline{0}_V + W$)

[Corollarly] If $\dim(V) < \infty$, $W \leq V$, then $\dim(V/W) + \dim(W) = \dim(V)$ "Quotient"

(proof.) With rank-nullity thm:
$$\underbrace{\dim(\ker(\pi))}_{\dim(W)} + \underbrace{\dim(\operatorname{Im}(\pi))}_{\dim(V/W)} = \dim(\underset{\text{Domain}}{V})$$

(e.g. $\underbrace{V = \mathbb{R}^3}_{\dim = 3}$, $W = \operatorname{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$ $\dim = 2$. $V/W = \operatorname{span}\left\{\underbrace{\begin{bmatrix}0\\0\\1\end{bmatrix} + W}_{\dim = 1}\right\}$

(2) (Thm 2) Let $T: V \mapsto W$ be a linear trans. & $V' \leq \ker(T)$ is a subspace of V. Then, a linear transformation can be defined as $\bar{T}: V/V' \mapsto W$ by $\bar{T}(\underline{v} + V') = T(\underline{v})$.

(proof.) $\bar{T}$ is well-defined. Suppose $\underline{v} \neq \underline{u}$ s.t. $\underline{v} + V' = \underline{u} + V'$. Then, $T(\underline{v}) - T(\underline{u}) = T(\underline{v} - \underline{u}) = \underline{0}_W$ since $\underline{v} - \underline{u} \in \ker(T)$. ☒

(check: $\bar{T}$ is a linear transformation)

3. Dual Vector Space

(1) [Defn] Let V be a vec. space over $\mathbb{F}$. The $\boxed{\text{dual of } V}$ is $V^* := \operatorname{Hom}_{\mathbb{F}}(V, \mathbb{F})$ where the latter is a collection of all linear transformations $\phi: V \mapsto \mathbb{F}$

(e.g. i) $V = M_{n\times n}(\mathbb{R})$, ϕ_{ij} = taking the i,j th entry $\in V^*$. Then, $\{\phi_{ij}\}_{i,j=1}^{n}$ forms a basis of V^* (recall $\boxed{\operatorname{Hom}_{\mathbb{F}}(V, W) \text{ with } \dim = \dim(V) \cdot \dim(W)}$)
proof: isomorphism between Hom & $M_{m\times n}(\mathbb{R})$

(e.g. ii) $V = \mathbb{R}[x]$ (space of polynomials)
$\phi_1(p(x)) := p(1)$, $\phi_1 \in V^*$.

(Remarks) ① Elements $\phi: V \mapsto \mathbb{F}$ in V^* are sometimes called linear functionals.

② vec. space structure for: $Hom_{\mathbb{F}}(V, \mathbb{F}) = V^*$ (linear combinations for ϕ & ψ)

③ To describe a function $f: A \mapsto B$, one needs only to specify $f(a)$, $\forall a \in A$

④ If we further know that any $f: V \mapsto W$ is a linear transformation, to describe, only need to specify $f(\underline{v_i})$, $\forall \underline{v_i}$ in basis $\mathcal{B}$.

(2) [Defn] Let V be a vec. space with basis $B = \{\underline{v_i} \mid i \in I\}$, $\forall i \in I$ define $\phi_i: V \mapsto F$ by

$$\begin{cases} \phi_i(\underline{v_j}) := \mathbb{1}_{\{i=j\}}, \ \forall j \in I \ \text{(values of basis)} \\ \forall \underline{v} \in V, \ \phi_i(\underline{v}) = \tilde{\alpha}_i, \text{ with } \underline{v} = \sum_{k \in I'} \alpha_k \underline{v}_k \quad (I' \subseteq I \text{ \& } I' \text{ is finite}, \ \tilde{\alpha}_i = \begin{cases} \alpha_i, & i \in I' \\ 0, & i \notin I' \end{cases}) \end{cases}$$

Then, by construction, $\phi_i \in V^*$, define $B^* = \{\phi_i \mid i \in I\}$

☆ (Thm) B^* as defined above, is a basis of V^*, if $\dim(V) < \infty$

(proof.) $\forall \phi \in V^*$, $\exists \beta_j \in \mathbb{F}$ s.t. $\phi(\underline{v_j}) = \beta_j \in \mathbb{F}$, $\forall j \in I$.

Then, $\phi = \sum_{j \in I} \beta_j \phi_j$, since I is finite ($\dim(V) < \infty$) (span)

$\{\phi_i\}_{i \in I}$ spans V^* (or we can show $\dim(V^*) = \dim(V) < \infty$)

Moreover, B^* is l.i. in V^*, since $\alpha_{i_1}\phi_{i_1} + \cdots + \alpha_{i_n}\phi_{i_n}$, $n \in \mathbb{N}$, $i_j \in I$, $j \in \{1, \cdots, n\}$.
true even for $\dim(V) = \infty$

if it is zero vec., take $\underline{v}_{i_1}, \cdots, \underline{v}_{i_n}$, respectively, $\Rightarrow \alpha_{i_1} = \cdots = \alpha_{i_n} = 0$. (l.i.)
const-zero functn.

(Remark.) If $\dim(V) = \infty$, B^* $\boxed{\text{MAY NOT BE}}$ a basis of V!

☆ [counter-e.g.] $\phi_i(x^j) = \mathbb{1}_{\{i=j\}}$ for polynomials in $\mathbb{R}[x]$, but $\phi(p) = p(1)$ cannot be written as $\sum_{j=1}^{k} \alpha_j \phi_{i_j}$, for otherwise, then $\phi(x^{i_k+1}) = 1 \neq 0 = \sum_{j=1}^{k} \alpha_j \phi_{i_j}(x^{i_k+1})$ with $i_1 < \cdots < i_k$. (In fact, $\phi = \sum_{i=1}^{\infty} \phi_i$ which is not allowed)

Conclusion: when $\dim(V) = \infty$, then V^* is somewhat "larger" than V

(3) Relationships between V, V^* & $(V^*)^*$ for finite dim case:

We can define an isomorphism $f: V \mapsto V^*$, but with a chosen basis $\mathcal{B} = \{\underline{v}_1, \cdots, \underline{v}_n\}$

However, if $F: V \mapsto (V^*)^*$, we have $F(\underline{v})[\phi] = \phi(\underline{v})$, $\forall \phi \in V^*$, independent of basis choice
functional

$\therefore$ $(V^*)^*$ is $\boxed{\text{closer to } V}$, compared with V^*, to some degree. $\boxed{\text{natural isomorphism}}$

4. Annihilators:

(1) [Defn] Let V be a vec. space, $S \subseteq V$ (subset). The annihilator of S is defined as $Ann(S) := \{\phi \in V^* \mid \phi(\underline{s}) = 0 \in F,\ \forall \underline{s} \in S\}$

(e.g. $\underline{0}_{V^*} \in Ann(S)$)

(2) [Proposition] (i) $Ann(S) \leq V^*$ (vec-subspace)

(ii) If $S \subseteq S' \subseteq V$, $Ann(S') \subseteq Ann(S)$

(3) [Thm] (i) $Ann(S) = Ann(span(S))$

(ii) If V s.t. $\dim(V) < \infty$ & $W \leq V$, then $[\dim(Ann(W)) + \dim(W) = \dim(V)]$

(the larger W, the smaller $Ann(W)$)

(proof.) $S \subseteq span(S) \Rightarrow Ann(S) \supseteq Ann(span(S))$, & if $\phi \in Ann(S)$, of course $\phi \in Ann(span(S))$ since ϕ is a linear trans. (Thus, $Ann(S) = Ann(span(S))$.)

Consider $\dim(V) = n < \infty$, $\dim(W) = k$ & $B^* = \{\phi_1, \cdots, \phi_n\}$ is a dual basis of V^*

Then, $Ann(W)$ has a basis $\{\phi_{k+1}, \cdots, \phi_n\}$ (without doubt that $span\{\phi_{k+1}, \cdots, \phi_n\} = Ann(W)$).

(Note that $f \in Ann(W)$, meaning that $r_1 = \cdots = r_k = 0$, with $\underline{v}_1, \cdots, \underline{v}_k$ taken.) ∎

(e.g. $V = M_{2\times2}(\mathbb{R})$, $S = \left\{\begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix}\right\}$, $Ann(S) = Ann(span(S))$,

then $\dim(Ann(S)) = \dim(V) - \dim(span(S)) = 3$.

Conclusion: $Ann(S) = span\{2\phi_{11} + \phi_{12},\ 3\phi_{11} - \phi_{21},\ 4\phi_{11} - \phi_{22}\}$, with ϕ_{ij}: take (i,j) entry.)

☆ Remark: If $\dim(V) < \infty$, $W \leq V$. we have $\dim(V/W) = \dim(V) - \dim(W) = \dim(Ann(W))$.

∴ It's possible to construct an isomorphism:

$\Phi : Ann(W) \mapsto V/W$

Then, let $B = \{\underbrace{\underline{v}_1, \cdots, \underline{v}_k}_{\text{basis of } W}, \underline{v}_{k+1}, \cdots, \underline{v}_n\}$ & then $\{\phi_{k+1}, \cdots, \phi_n\}$ is a basis of $Ann(W)$ & $\{\underline{v}_{k+1} + W, \cdots, \underline{v}_n + W\}$ is a basis of V/W. (We can define $\Phi(\phi_\ell) := \underline{v}_\ell + W$, $\ell = k+1, \cdots, n$) ← depend on bases!

Need a natural isomorphism. (indep. of basis)

$\boxed{\Psi : Ann(W) \mapsto (V/W)^*}$

(4) Let $f \in Ann(W)$, then $(f: V \mapsto \mathbb{F}) \in V^*$ is a linear trans. with $W \leq \ker(f)$
with this f, can define $\tilde{f}: V/_W \mapsto \mathbb{F}$ (proved) s.t. $\tilde{f}(\underline{v}+W) = f(\underline{v})$
[Defn] Define $\Phi: Ann(W) \mapsto (V/_W)^*$, with $\boxed{\Phi(f) = \tilde{f}}$, $f, \tilde{f}$ defined as above.
☆ propositions. (i) Φ is a linear transformation. (easily check)
(ii) Φ is an isomorphism, ("injective" + "surjective") if $\dim(V) < \infty$
(proof.) Suppose $f \in \ker(\Phi)$, then $\tilde{f}(\underline{v}+W) = f(\underline{v}) = 0$, $\forall \underline{v} \in V$. (since $\Phi(f) = \tilde{f}$ $= \underline{0}_{(V/_W)^*}$). $\therefore f = \underline{0}_{V^*}$ (or $\underline{0}_{Ann(W)}$) $\Rightarrow f$ is injective. (TRUE for $\dim(V) = \infty$ case!)
Since by rank-nullity thm. $\dim(Im(\Phi)) = \dim(Ann(W)) \overset{\text{proved}}{=} \dim(V/_W) = \dim(V/_W)^* \Rightarrow$
$Im(\Phi) = (V/_W)^*$ since $Im(\Phi) \leq (V/_W)^*$. Thus, f is surjective. ▣
(Remark.) If $\dim(V) < \infty$, it (Φ) is a natural isomorphism.
& $\dim(V) = \infty$, Φ is $\boxed{\text{still}}$ an isomorphism (need new proofs)
(proof.) Construct $\Psi: (V/_W)^* \mapsto Ann(W)$ as an inverse, namely, $\Phi \circ \Psi: (V/_W)^* \mapsto (V/_W)^*$
& $\Psi \circ \Phi: Ann(W) \mapsto Ann(W)$ are both identity maps. (payoff: need transpose of $T: V \mapsto W$)
(constructions below)

(5) Transpose of Linear Transformations:
[Defn] Let $T: V \mapsto W$ be a linear trans., define $T^t: W^* \mapsto V^*$ by:
let $f \in W^*$, $T^t(f)(\underline{v}) := f(T(\underline{v}))$, $\forall \underline{v} \in V$. ($T^t(f): V \mapsto \mathbb{F}$)
linear functional
(verifications: $T^t(f) \in V^*$ (linear trans.) & $T^t(\cdot)$ is a linear trans., etc.)
☆ [Proposition] Let V, W be finite dimensional, $T: V \mapsto W$, is a linear trans.
with transpose $T^t: W^* \mapsto V^*$. Consider: $\mathcal{A} = \{\underline{v}_1, \dots \underline{v}_n\}$, $\mathcal{B} = \{\underline{w}_1, \dots \underline{w}_m\}$,
$\mathcal{A}^* = \{\phi_1^*, \dots \phi_n^*\}$ & $\mathcal{B}^* = \{\psi_1^*, \dots \psi_m^*\}$ are bases of V, W, V^*, W^*, respectively.
Then, matrix representations: $(T: V \mapsto W) \rightsquigarrow (T)_{\mathcal{B}\mathcal{A}} = (\alpha_{ij})_{m \times n}$
$(T^t: W^* \mapsto V^*) \rightsquigarrow (T^t)_{\mathcal{A}^*\mathcal{B}^*} = (\beta_{ij})_{n \times m}$
are with $\alpha_{ij} = \beta_{ji}$, $\forall i, j$ ← matrix transpose!
(proof.) $T(\underline{v}_j) = \sum_{l=1}^{m} \alpha_{lj} \underline{w}_l$; $T^t(\psi_i^*) = \sum_{k=1}^{n} \beta_{ki} \phi_k^*$. On one hand, $T^t(\psi_i^*)(\underline{v}_j) = \psi_i^*(T(\underline{v}_j))$
$= \sum_{l=1}^{m} \alpha_{lj} \psi_i^*(\underline{w}_l) = \alpha_{ij}$, on the other $T^t(\psi_i^*)(\underline{v}_j) = \sum_{k=1}^{n} \beta_{ki} \phi_k^*(\underline{v}_j) = \beta_{ji}$. ▣

(proof cont'd) linear trans. $\pi: V \mapsto V/W$, then its transpose: $\pi^t: (V/W)^* \mapsto V^*$
$(\pi(\underline{v}) := \underline{v}+W)$

Claim: $\forall h \in (V/W)^*$, $\pi^t(h) \in Ann(W)$ (then $\pi^t: (V/W)^* \mapsto Ann(W)$)

$\pi^t(h)\cdot \underline{w} = h[\pi(\underline{w})] = h(\underline{w}+W) = h(\underline{0}_{V/W}) = 0$, $\forall w \in W$.
↳ $\in Ann(W)$ by definition!

Let $\Psi = \pi^t$, claim': $\left\{ \begin{array}{l} \Psi \circ \Phi \\ \Phi \circ \Psi \end{array} \right.$ are both identity maps. (need checking!)

verify: take any $f \in Ann(W)$, $\Phi(f) = \tilde{f} \in (V/W)^*$. Then $\Psi\circ\Phi(f) = \Psi(\tilde{f}) \in Ann(W)$

$\Psi\circ\Phi(f)(\underline{v}) = \Psi(\tilde{f})(\underline{v}) = \tilde{f}(\pi(\underline{v})) = \tilde{f}(\underline{v}+W) = f(\underline{v})$, $\forall \underline{v} \in V$.

Similary $\Phi\circ\Psi(\tilde{f}) = \tilde{f}$, $\forall \tilde{f} \in (V/W)^*$. They are identity maps!

- Eigenvalues & Eigenvectors

1. Characteristic Polynomials & Minimal Polynomials

(1) Recap: ($\boxed{\text{monic}}$: if a_n (leading coefficient) $= 1$)

(i) g.c.d. (Greastest common divisor) of $p_1, \dots p_r \in \mathbb{F}[x]$, is defined as a monic polynomial so that $g(x) \mid p_i(x)$, $1 \le i \le r$ with $\deg(g)$ is the highest possible degree.

(ii) [Unique Factorization]
↳ $f \in \mathbb{F}[x]$ can be factorized uniquely as products (with degrees) of monic, irreducible, distinct polynomials

(iii) To find g.c.d. ← Euclidean algorithm (tossing & turning divisions)

Note Bezout theorem is true for polynomials ($\exists\, a(x), b(x) \in \mathbb{F}[x]$ s.t $ag_1 + bg_2 = gcd(g_1, g_2)$.)

If $g_1, g_2, \dots g_n$ s.t. $gcd(g_1, \dots, g_n) = 1$
they are called relative prime.

(2) [Defn] Let $T: V \mapsto V$ be a linear operator, a non-zero vector $\underline{v} \in V$ is an eigenvector of T with eigenvalue λ if $T(\underline{v}) = \lambda \underline{v}$

(rmk.) By matrix representation, $T(\underline{v}) = \lambda \underline{v} \iff \boxed{T_{\mathcal{BB}}[\underline{v}]_{\mathcal{B}} = \lambda[\underline{v}]_{\mathcal{B}}}$ where $\mathcal{B}$ is an ordered basis of V with $\dim(V)=n<\infty$. & To find ei-values of $T: V\mapsto V$ $\iff$ find ei-values of matrix $T_{\mathcal{BB}}$.

★ [Defn] Let $\dim(V)=n<\infty$ & $T: V\mapsto V$ be a linear operator. Then the characteristic polynomial of T is defined as $\chi_T(x) := \chi_{T_{\mathcal{BB}}}(x) = \det(xI_n - T_{\mathcal{BB}})$ (with some chosen $\mathcal{B}$)

(rmk.) By similarity, $\chi_T(x)$ is independent of $\mathcal{B}$.

(3) Polynomials of Operators & Minimal Polynomials

(i) [Defn] Let $p(x)=\sum_{i=0}^{n} a_i x^i \in \mathbb{F}[x]$ be a polynomial and $T: V\mapsto V$ is a linear operator. Define $P(T): V\mapsto V$ by $p(T) := \sum_{i=1}^{n} a_i(\underbrace{T\circ\cdots\circ T}_{i\,\#}) + a_0 I$

(Rmk.) for detailed calculations, $T_{\mathcal{BB}}$ can be used & $T\circ\cdots\circ T = (T_{\mathcal{BB}})^i$.)
if $T, S: V\mapsto V$, generally $T\circ S \neq S\circ T$, nonetheless, $p(T)q(T) = q(T)p(T)$, $\forall p, q \in \mathbb{F}[x]$.

(ii) [Defn] Let $T: V\mapsto V$ be a linear operator, then the minimal polynomial $m_T(x)$ of T is a monic polynomial with $m_T(T)$ being [zero operator] & having smallest degree!

(e.g.) $m_T = x^2-5x-2$ for $T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x+2y\\3x+4y\end{bmatrix}$ (check 2 is the smallest degree! ← degree 1 fails)

(iii) [Thm] Cayley-Hamilton: $\chi_T(T) = 0$ (operator) with $T: V\mapsto V$ linear operator & V is finitely generated
(coro.) $\deg(m_T(x)) \leq \deg(\chi_T(x))$.)

(iv) Properties of minimal polynomials

[I.] If $\dim(V)<\infty$, then $m_T(x)$ always exists.

(proof.) Since $\mathrm{Hom}_{\mathbb{F}}(V,V) = n^2$ (with $\dim(V)=n<\infty$), now consider $\{\underbrace{I, T, \dots, T^{n^2}}_{n^2+1 \text{ elements}}\} \subseteq \mathrm{Hom}_{\mathbb{F}}(V,V)$, $\forall T\in \mathrm{Hom}_{\mathbb{F}}(V,V)$, the set must be l.d., that is, $\exists$ non-zero (all), wlog. $b_{n^2}\neq 0$, $b_0, \dots, b_{n^2}$ s.t. $\sum_{i=0}^{n^2} b_i T^i = 0$ (operator) $\Rightarrow \sum_{i=0}^{n^2}\frac{b_i}{b_{n^2}}x^i$ satisfies monic & "zero-operator" conditions. ◻

[II.] $m_T(x)$ is unique.

(proof.) Suppose $\exists\ p\neq q$ & p, q are both minimal polynomials. Back to defn. $\begin{cases} p(T), q(T) = 0_{V\mapsto V} \\ \deg(p)=\deg(q)\end{cases}$
let $h(x) = p(x)-q(x)$. Then $h(x)$ with $h(T) = 0_{V\mapsto V}$ & $\deg(h) \leq \deg(p)-1$ (or $\deg(q)-1$)
$\therefore$ It's possible to find a monic polynomial $(\hat{h})$ s.t. $(\hat{h}(T) = 0_{V\mapsto V}$ & $0<\deg(\hat{h}) < \deg(p))$ contra! ◻

[III.] $\forall$ polynomial $f(x) \in \mathbb{F}[x]$ s.t. $f(T) = \underline{0}_{V\mapsto V}$, then $m_T(x) \mid f(x)$.

(proof.) Since $\deg(f(x)) \geq \deg(m_T)$, then $\exists\, q, r \in \mathbb{F}[x]$, $\deg(r) < \deg(m_T)$ s.t. $f(x) = q(x) m_T(x) + r(x)$. Plug into T, $\underline{0}_{V\mapsto V} = f(T) = q(T)\underbrace{m_T(T)}_{\underline{0}_{V\mapsto V}} + r(T) \Rightarrow \underline{0}_{V\mapsto V}$

$\therefore$ r satisfies $r(T) = \underline{0}_{V\mapsto V}$ & then since $\deg(r) < \deg(m_T)$, the only possibility is: $r \equiv 0$. $\boxed{\text{III}}$

(r.m.k. Euclidean algorithm $\xRightarrow[\text{to}]{\text{close}}$ "minimal polynomial")

↳ to find the mininal polynomial: run through all factors of $\chi_T(\cdot)$. ← brutal search

2. Triangularizable Operators

(1) Definitions & Basic properties

(i) [Defn] Let $T: V \mapsto V$, we say T is diagonalizable if $\exists$ a basis $\mathcal{B}$ s.t. $T_{\mathcal{B}\mathcal{B}}$ is a diagonalizable matrix. Similarly, T is triangularizable if $\exists$ a basis $\mathcal{B}$ s.t. $T_{\mathcal{B}\mathcal{B}}$ is an upper triangular matrix.

(r.m.k.) ① this defn $\Rightarrow$ $\mathcal{B}$ is a basis of ei-vectors of T ⇐ relationship (kinship) with MAT2040 matrix case

② Obviously, $\chi_T(x)$ is NOT enough to determine diagonalizablity. but enough to determine triangularizablity

(ii) [Thm] $T: V \mapsto V$ is triangularizable exactly when $\chi_T(x)$ can be factorized into $\boxed{\text{linear factors}}$ (in real numbers)

↖ i.e., NO quadratic factors in its factorization form! ↪ (in $M_{n\times n}(\mathbb{R})$.)

(r.m.k.) ① The thm $\boxed{\text{still works}}$ in $M_{n\times n}(\mathbb{C})$, e.g. $T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Rightarrow (x+i)(x-i) = 0$

$\therefore \exists\, B \in M_{2\times 2}(\mathbb{C})$ s.t. $B^{-1} T B$ is an upper triangular matrix.

② If V is a vector space over $\mathbb{C}$ with $\dim(V) < \infty$, by fundamental thm of algebra, every $T: V \mapsto V$ is triangularizable.

(proof.) ($\Rightarrow$) Done by considering $\det(xI - T_{\mathcal{B}\mathcal{B}})$, which is $\prod_{i=1}^{n}(x - a_{ii})$.

($\Leftarrow$) recall T-invariant subspaces $W \leq V$. i.e. $T(W) \leq W$.

examples of T-invariant subspaces. If $\underline{v} \in V$ is an ei-vector of T, then $span\{\underline{v}\}$ is T-invariant. The λ-eigenspace $E_\lambda := \{\underline{v} \in V \mid T(\underline{v}) = \lambda \underline{v}\}$ ← space of λ-eivectors, is also T-invariant subspace of V. ★More generally, $\forall g \in \mathbb{F}[x]$, $g(T): V \mapsto V$, then $\boxed{ker(g(T)) \leq V \text{ is } T\text{-invariant}}$ ($g = \lambda x \Rightarrow E_\lambda = ker(g(T))$)

More precisely, let $\underline{v} \in ker(g(T))$, $g(T)(T(\underline{v})) = T \circ g(T)(\underline{v}) = T(\underline{0}_V) = \underline{0}_V$, then $T(\underline{v}) \in ker(g(T))$. Thus, $T(ker(g(T))) \leq ker(g(T))$.

Since when $T: V \mapsto V$ & $W \leq V$ is T-invariant, $T|_W : W \mapsto W$; $\tilde{T}: V/W \mapsto V/W$ $\tilde{T}(\underline{v}+W) = T(\underline{v}) + W$. Then, let $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ be basis of V & $\mathcal{C} = \{v_1, \dots, v_k\}$ be a basis of W, then $\underbrace{\{v_{k+1}+W, \dots, v_n+W\}}_{\triangleq \mathcal{A}}$ is a basis of V/W.

Then, $T_{\mathcal{B}\mathcal{B}} = \left[\begin{array}{c|c} (T|_W)_{\mathcal{C}\mathcal{C}} & * \\ \hline O & \tilde{T}_{\mathcal{A}\mathcal{A}} \end{array}\right]$ (easily checked)

[lemma] Under this setting, $\chi_T(x) = (\chi_{T|_W}(x)) \cdot (\chi_{\tilde{T}}(x))$

(proof. $\chi_T(x) = det(xI - (T|_W)_{\mathcal{C}\mathcal{C}}) \cdot det(xI - \tilde{T}_{\mathcal{A}\mathcal{A}})$)

By induction, $dim(V) = 1$ ✓, assume $dim(V) = k (\geq 1)$ case holds. Suppose now $T: V \mapsto V$ with $dim(V) = k+1$, we can write $\chi_T(x)$ as $(\chi_{T|_W}(x))(\chi_{\tilde{T}}(x))$ ← dim 1 ← dim k (hypothesis of induction) since $W = span\{v_1\}$ for some root λ_1 with $T(v_1) = \lambda_1 v_1$. By lemma & representation of $T_{\mathcal{B}\mathcal{B}}$, $T_{\mathcal{B}\mathcal{B}}$ with $dim(V) = k+1$ is triangularizable, induction $\Rightarrow \exists \mathcal{A} = \{v_2 + W, \dots, v_{k+1} + W\}$ making $\tilde{T}_{\mathcal{A}\mathcal{A}}$ triangularizable.

$\therefore \{v_1, \dots, v_n\} \triangleq \mathcal{B}$ makes $T_{\mathcal{B}\mathcal{B}}$ triangularizable. ▥

3. Proof of Cayley-Hamilton Theorem.

(1) [Proposition 1] If $T: V \mapsto V$ is triangularizable, then $\chi_T(T) = O_{V \mapsto V}$ [True for all $T: V \mapsto V$, over $\mathbb{C}$!

(proof.) Suppose $\mathcal{B}$ makes $T_{\mathcal{B}\mathcal{B}}$ triangularizable, with the form $\begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ O & & \lambda_n \end{bmatrix}$

Then $\chi_T(T) = \prod_{i=1}^{n} (T - \lambda_i I)$, with matrix representations, it is zero matrix.

Reason, let $A_i \in M_{n \times n}(\mathbb{R})$, $1 \leq i \leq n$ be upper-triangular matrices with $(A_i)_{i,i} = 0$.

Then $\prod_{i=1}^{n} A_i = \underline{O}_{M_{n\times n}(\mathbb{R})}$, since $\prod_{i=1}^{n} A_i \cdot \underline{x} = \underline{0}_{\mathbb{R}^n}$, $\forall \underline{x} \in \mathbb{R}^n$ (every time multiplying from backward, we get one more zero entry, that is, $\prod_{i=m}^{n} A_i \underline{x}$ has zero entries in $m, m+1, \dots, n$ positions)

(2) [Proposition 2] $T: V \mapsto V$ f.g. over $\mathbb{R}$. Then $\chi_T(T) = \underline{O}_{V\mapsto V}$.

(proof.) Let $M = T_{\mathcal{B}\mathcal{B}} \in M_{n\times n}(\mathbb{R})$, define $S: \mathbb{C}^n \mapsto \mathbb{C}^n$, with $S(\underline{x}) := M\underline{x}$, an linear operator with $\mathbb{C}$ vector space. By prop. 1 above, $\chi_S(x) = \chi_M(x) \in \mathbb{C}[x]$.

$\therefore \chi_S(x) = \chi_T(x)$, which gives $\chi_T(S) = \underline{O}_{\mathbb{C}^n\mapsto\mathbb{C}^n} \Rightarrow \chi_T(S_{\xi\xi}) = O_{n\times n}\ (\in M_{n\times n}(\mathbb{C}))$

$\xRightarrow[\xi]{\text{some basis}} \chi_T(M) = O_{n\times n}$ (thus, $\chi_T(T_{\mathcal{B}\mathcal{B}}) = O_{n\times n}\ (\in M_{n\times n}(\mathbb{R}))$). $\therefore \chi_T(T) = \underline{O}_{V\mapsto V}$.

(r.m.k. By Cayley-Hamilton thm, $m_T(x) \mid \chi_T(x)$. we can find the $m_T(x)$ from factorization $\chi_T(x) = \prod_{i=1}^{N} p_i^{e_i}(x)$ with distinct irreducible polynomials.)

$\therefore m_T(x) = \prod_{i=1}^{N} p_i^{f_i}(x)$, with $0 \le f_i \le e_i$

(3) [Proposition 3] Let $T: V \mapsto V$ be s.t. $\chi_T(x) = \prod_{i=1}^{N} p_i^{e_i}(x)$, with irreducible, distinct polynomials p_i, & $e_i > 0$. Then, $m_T(x) = \prod_{i=1}^{N} p_i^{f_i}(x)$, with $0 < f_i \le e_i$. all basic factors show!

(proof.) Only consider $n=2$ case, $\chi_T(x) = p_1^{e_1}(x)\, p_2^{e_2}(x)$, let $p_T(x) = p_2^{f_2}(x)$, (with $f_2 \le e_2$)

Only consider complex cases (real cases - same trick above).

Then, $\chi_T(x) = [\prod_{i=1}^{m}(x-\beta_i)]^{e_1} \cdot [\prod_{i=1}^{k}(x-\theta_i)]^{e_2}$, $\beta_i \neq \theta_j \in \mathbb{C}$, $T_{\mathcal{B}\mathcal{B}} \triangleq M$ has an ei-vector in $\mathbb{C}^n$ with an ei-value μ_1. Compute $p_T(M)\underline{v} = p_2^{f_2}(M)\underline{v} = [\prod_{i=1}^{k}(M-\theta_i I)]^{f_2}\underline{v}$ (for ei-vec $\underline{v} \in \mathbb{R}^n$)

$= \prod_{i=1}^{k}[(\mu_1-\theta_i)]^{f_i}\underline{v}$. Since $\mu_1 \in \{\beta_1 \dots \beta_m\}$, $p_T(M)\underline{v} \neq 0$. which means $p_T(\cdot) \neq m_T(\cdot)$

nonzero!!

4. Primary Decomposition

(1) [Thm] Let $T: V \mapsto V$ be a linear operator ($\dim(V) < \infty$)

Suppose $m_T(x) = p_1^{e_1}(x) \cdots p_k^{e_k}(x)$ with distinct, irreducible & monic p_i. $1 \le i \le k$

Let $V_i := \ker(p_i^{e_i}(T)) \le V$.

① $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$

② V_i is a [T-variant subspace] of V, $\forall i$

③ The minimum polynomial of $[T|_{V_i}: V_i \to V_i$ is $m_{T|_{V_i}}(x) = p_i^{e_i}(x)]$

(proof.) Let $q_i(x) = \frac{m_T(x)}{p_i^{e_i}(x)}$, then g.c.d $(q_1, q_2, \dots, q_k) = 1$

g.c.d. $(q_i, p_i^{e_i}) = 1$. & $m_T(x) \mid q_i q_j(x)$, $\forall i \neq j$.

Bezout thm, $\sum_{i=1}^{k} a_i(x)q_i(x) = 1$, with some $a_1, a_2, \dots, a_k \in \mathbb{F}[x]$

$\therefore \sum_{i=1}^{k} a_i(T)q_i(T) = I$. Thus, $\forall \underline{v} \in V$, $\underline{v} = \sum_{i=1}^{k} \underbrace{a_i(T)q_i(T)(\underline{v})}$ denote as $\underline{v_i}$.

Claim 1: $\underline{v_i} \in \ker(P_i^{e_i}(T)) = V_i$, since $P_i^{e_i}(T)a_i(T)q_i(T)(\underline{v}) = a_i(T)\underbrace{P_i^{e_i}(T)q_i(T)}_{m_T(T)}(\underline{v}) = \underline{0}$

$\therefore V = V_1 + V_2 + \dots + V_k$ $\quad \boxed{f(T)g(T) = g(T)f(T)}$ (but generally $AB \neq BA$)

It's a direct sum, since if $\underline{v_1'} + \underline{v_2'} + \dots + \underline{v_k'} = \underline{0}$, then $\underline{v_i'} = \underline{0}$ by multiplying $q_i(T)$. $\forall 1 \le i \le k$

(since $q_i(T)\underline{v_i'} \neq 0$ by Bezout's thm, $b(x)q_i(x) + c(x)p_i^{e_i}(x) = 1$.) multiple $b(T)q_i(T)$ - better can change the order!

For ③, firstly, $m_{T|_{V_i}}(x) \mid p_i^{e_i}(x)$ since $p_i^{e_i}(T|_{V_i})(\underline{v}) = 0$, $\forall \underline{v} \in V_i$

let $m_{T|_{V_i}}(x) = p_i^{f_i}(x)$, $0 < f_i \le e_i$. Consider ① again, $\forall \underline{v} \in V$, $\underline{v} = \underline{v_1} + \dots + \underline{v_k} \Rightarrow$

$p_1^{e_1}(T) \cdots p_i^{f_i}(T) \cdots p_k^{e_k}(T)(\underline{v}) = \underline{0}$ thus $m_T(x) \mid p_1^{e_1}(x) \cdots p_i^{f_i}(x) \cdots p_k^{e_k}(x) \Rightarrow f_i = e_i$

[Coro.] $T: V \mapsto V$ linear, $\dim(V) < \infty$. Then $\left(T \text{ is diagonalizable}\right) \Leftrightarrow \left(m_T(x) = (x-\mu_1) \cdots (x-\mu_k) \text{ with distinct } \mu_1, \dots, \mu_k\right)$

(Proof.) "$\Leftarrow$" PDT ($V = V_1 \oplus \dots \oplus V_k$) where $V_i = \ker(T - \mu_i I)$ $\to$ μ_i eigen-space!

Thus, take $\mathcal{B}_i :=$ basis of V_i, $\forall 1 \le i \le k$.

$\mathcal{B} := \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_k$ is a basis of V. (easily checked)

Meanwhile, every vector in $\mathcal{B}_i$ is an ei-vector. then $T_{\mathcal{B}\mathcal{B}} = \mathrm{diag}(\mu_\cdot, \mu_\cdot, \dots)$ (diagonalizable)

"$\Rightarrow$" easily done. (Note "m_T" is this corollary but NOT "χ_T")

(2) [Thm] (Jordan Normal/Canonical Form)

Suppose $T: V \mapsto V$ be s.t. $m_T(x) = (x-\mu_1)^{e_1} \cdots (x-\mu_k)^{e_k}$. $\exists$ a basis $\mathcal{A}$ of V

s.t. $T_{\mathcal{A}\mathcal{A}}$ has a [blocked Jordan normal form.] (i.e., [Jordan blocks diagonal])

$J_i = \begin{bmatrix} \mu_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \mu_j & 1 \\ 0 & & & 0 \end{bmatrix}$ $\leftarrow$ the most diagonal ones

$1 \le m \le k$ (see previous 2040 Notes for more details)

(r.m.k. the sizes of every J_r need NOT be the same!

May $\exists$ 2 J_i, J_j, $i \neq j$ with the same diagonal entries!

Every $\mu_1, \dots, \mu_k$ must be in $\boxed{\text{at least 1}}$ J_r)

Above corollary (diagonalizable) is a special case of this JNF

(proof.) Case 1: let $m_T(x) = x^m \Rightarrow \underset{\leftarrow K_0}{0} \le \underset{\leftarrow K_1}{\ker(T)} \le \dots \le \underset{\leftarrow K_m}{\ker(T^m)} = V$ (all NOT equality.)

Let $W_k = K_k / K_{k-1}$, $k = 1, \dots, m$. $\mathcal{B}_m' = \{\underline{u}_1^m + K_{m-1}, \dots, \underline{u}_i^m + K_{m-1}\}$ is a basis of W_m

consider $\mathcal{C}_m' = \{T(\underline{u}_1^m) + K_{m-2}, \dots, T(\underline{u}_i^m) + K_{m-2}\}$, it's in W_{m-1}, since $T(\underline{u}_j^m) \in K_{m-1}$ $\forall j$ (easy!)

Moreover, $\mathcal{C}_m'$ is l.i. in W_{m-1} since only $\alpha_1 = \cdots = \alpha_i = 0$ gives $\alpha_1 \underline{u}_1^m + \cdots + \alpha_i \underline{u}_i^m \in \ker(T^{m-1})$

(because of the defn of basis of $\mathcal{B}_m'$).

By construction method above, $\{\underline{u}_1^m + K_{m-1}, \dots, \underline{u}_i^m + K_{m-1}\} \xrightarrow{\text{extension}} \{T(\underline{u}_1^m) + K_{m-2}, \cdots, T(\underline{u}_i^m) + K_{m-2}, \underline{v}_1 + K_{m-2}$

$\dots, \underline{v}_j + K_{m-2}\} \Rightarrow \cdots \Rightarrow \{T^{m-1}(\underline{u}_1^m) + K_0, \dots, T^{m-1}(\underline{u}_i^m) + K_0, \dots, T^{m-2}(\underline{v}_j) + K_0, \dots \underline{y}_1 + K_0, \dots \underline{y}_r + K_0\}$

$\Rightarrow \{T^{m-1}(\underline{u}_1^m), \dots, T(\underline{u}_1^m), \underline{u}_1^m, \dots, T^{m-1}(\underline{u}_i^m), \dots, T(\underline{u}_i^m), \underline{u}_i^m, T^{m-2}(\underline{v}_1), \dots, \underline{v}_1, \dots T^{m-2}(\underline{v}_j), \dots, \underline{v}_j,$

$\mathcal{A} \hat{=}$ $\underline{y}_1, \dots \underline{y}_r\}$ is a basis of V (since from $V/W_{m-1}, \dots, W_2/W_1, W_1/W_0\}$ $\hookrightarrow \{\underline{0}\}$

Construct $T_{\mathcal{A}\mathcal{A}}$, that is $T^m(\underline{u}_1^m) = \underline{0}$, $T^{m-1}(\underline{u}_1^m) = 1 \cdot T^{m-1}(\underline{u}_1) + 0 \cdots$, $\cdots$ $\underline{u}_1^m \to 1$ at m^{th} entry $\dots$

$$\begin{bmatrix} \boxed{\begin{smallmatrix} 0 & 1 & \\ & \ddots & 1 \\ & & 0 \end{smallmatrix}} \leftarrow \text{w.r.t. } \underline{u}_1^m \, (m \times m) & & 0 \\ & \ddots \quad \boxed{0} \leftarrow \text{w.r.t. } \underline{y}_1 \, (1\times1) & \\ 0 & & \boxed{0} \leftarrow \text{w.r.t. } \underline{y}_r \, (1\times1) \end{bmatrix}$$

Jordan blocks

Totally, we don't know # of blocks.
(many cases)
however, $\exists$ $m \times m$ block ("must") → largest

Case II : $m_T(x) = (x-\mu)^m$, consider $S := (T - \mu I d)$ $\quad (T: V \mapsto V)$

$\Rightarrow m_S(x) = x^m \Rightarrow \exists \mathcal{B}$ s.t. $S_{\mathcal{B}\mathcal{B}} = \begin{bmatrix} \boxed{\begin{smallmatrix} 0 & 1 \\ & 0 \end{smallmatrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{0} \end{bmatrix} = T_{\mathcal{B}\mathcal{B}} - \mu I$

Case III : $m_T(x) = (x-\mu_1)^{m_1} \cdots (x-\mu_k)^{m_k}$

By Primary Decomposition thm, $V = V_1 \oplus \cdots \oplus V_k$. $V_i = \ker((T - \mu_i I)^{m_i})$

& $T|_{V_j} : V_j \mapsto V_j$, with $m_{T|_{V_j}}(x) = (x - \mu_j)^{m_j}$. By case II, $\exists$ basis $\mathcal{B}_i$ of V_i s.t

$(T|_{V_i})_{\mathcal{B}_i\mathcal{B}_i} = [\text{ blocked form }]$. Let $\mathcal{B} := \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ be a basis of V (why?)

$(T)_{\mathcal{B}\mathcal{B}} = \left[\begin{array}{c|c|c} (T|_{V_1})_{\mathcal{B}_1\mathcal{B}_1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & (T|_{V_k})_{\mathcal{B}_k\mathcal{B}_k} \end{array}\right]$ which is JNF.

(r.m.k Not all matrices over $\mathbb{R}$ have JNF!)

(ii) Find / Calculate JNF : $\chi_T(x) = (x-\mu_1)^{e_1} \cdots (x-\mu_n)^{e_n}$

work with $(T - \mu_1 I)^{m_1}$, $1 \le m_1 \le e_1$, stop when $\dim(\ker(T - \mu_1 I)^{m_1}) = \dim(V_1) \hat{=} d_1$

(i.e. $\dim(\ker(T-\mu_1 I)^{m_1}) = \dim(\ker(T - \mu_1 I)^{m_1+1})$), $\{\underline{v}_1, \dots \underline{v}_{d_1}\} \to$ basis of $(T - \mu_1 I)^{m_1}$ $(m_{T|_{V_1}} = (\cdot)^{m_1})$
$\hookrightarrow$ generalized ei-vectors!

Similarly for $(x-\mu_2)^{e_2}, \dots, (x-\mu_n)^{e_n}$.

• Inner Product Space

IDEA : generalize "•" (dot product) in $\mathbb{R}^n$

(1) [Defn] Let V be a vector space over $\mathbb{R}$, a $\boxed{\text{bilinear form}}$ on V is a function

$$B: V\times V \mapsto \mathbb{R} \text{ s.t. } \begin{cases} B(\underline{u}+\underline{v}, \underline{w}) = B(\underline{u},\underline{w}) + B(\underline{v},\underline{w}) \\ B(\underline{u}, \underline{v}+\underline{w}) = B(\underline{u},\underline{v}) + B(\underline{u},\underline{w}) \end{cases} \text{"sum"}$$

$$B(\lambda \underline{u}, \underline{v}) = \lambda B(\underline{u},\underline{v}) = B(\underline{u}, \lambda\underline{v}), \text{ for all } \lambda\in\mathbb{R} \quad \text{"prod"}$$

we say B is $\begin{cases} \text{symmetric} \\ \text{positive definite} \\ \text{non-degenerate} \end{cases}$ if $\begin{cases} B(\underline{v},\underline{u}) = B(\underline{u},\underline{v}), \ \forall \underline{u},\underline{v}\in V \\ B(\underline{v},\underline{v}) \geq 0 \ \forall \underline{v}\in V \text{ equality holds iff } \underline{v}=\underline{0}_V \\ B(\underline{u},\underline{v})=0, \ \forall \underline{v}\in V \Rightarrow \underline{u}=\underline{0}_V. \end{cases}$

(e.g.s $\begin{cases} \text{Dot product on } \mathbb{R}^n, \text{ i.e. } B(\underline{u},\underline{v}) = \langle \underline{u},\underline{v}\rangle = \underline{u}^t\underline{v} \quad (\text{s.p.d \& non-degenerate}) \\ \text{Lebesque integral on } V=C^\infty([0,1]), \text{ i.e. } \int_{[0,1]} f\cdot g\, d\mu = B(f,g) \quad (\text{s.p.d \& non-degenerate}) \end{cases}$

(generally, on $L^2(\mu)$) "a.e. on [0,1]" but in C^∞, fine.

(r.m.k There are $\boxed{\text{NO positive definite bilinear form}}$ for vector space V over complex field $\mathbb{C}$.)

explanation: $(B(\underline{v},\underline{v})>0, \ \forall \underline{v}\neq 0$ but $B(i\underline{v}, i\underline{v}) = -B(\underline{v},\underline{v})<0$, contradictory!$)$)

↳ reason why $\|\underline{v}\| := B(\underline{v}, \overline{\underline{v}})$ ← other forms for complex numbers! sesqui-linear!

[Defn'] Let V be a vec. space over $\mathbb{C}$, a $\boxed{\text{sesqui}}$linear form on V, is a function ($\text{sesqui} = 1.5$)

$$B: V\times V \mapsto \mathbb{C} \text{ s.t. } \begin{cases} B(\underline{x}, \lambda\underline{y}) = \lambda B(\underline{x},\underline{y}) \ \& \ B(\lambda\underline{x},\underline{y}) = \bar{\lambda} B(\underline{x},\underline{y}) \\ B(\underline{x}+\underline{y}, \underline{z}) = B(\underline{x},\underline{z}) + B(\underline{y},\underline{z}) \\ B(\underline{x}, \underline{y}+\underline{z}) = B(\underline{x},\underline{y}) + B(\underline{x},\underline{z}) \end{cases}$$

($\bar{\lambda}$: conjugate in $\mathbb{C}$)

☆ Note: a sesqui-linear form is $\begin{cases} \text{conjugate symmetric if } B(\underline{x},\underline{y}) = \overline{B(\underline{x},\underline{y})} \\ \text{positive definite if } B(\underline{x},\underline{x}) \geq 0 \text{ only "=" when } \underline{x}=\underline{0}_V. \end{cases}$

(e.g. $V=\mathbb{C}^n$, $B(\underline{v},\underline{w}) = \langle \overline{\underline{v}}, \underline{w}\rangle = \underline{v}^H\underline{w}$ (pos. def., conjugat symmetric)

(2) [Defn] (Inner Product Space) Let V be a vec. space over $\mathbb{R}$ (or $\mathbb{C}$). An inner product of V is a bilinear (or sesqui-linear) form. $B: V\times V\mapsto \mathbb{R}$ (or $\mathbb{C}$) on V s.t.

B is (conjugate) symmetric & positive definite. (Denoted as $(V, \langle , \rangle)$).

(r.m.k $\begin{cases} \text{orthogonality } \langle \underline{u}, \underline{v} \rangle = 0 \\ \text{norm of } \underline{v}: \|\underline{v}\| = \sqrt{\langle \underline{v}, \underline{v} \rangle}. \\ \text{Unit vector}: \|\underline{v}\| = 1 \end{cases}$ Cauchy-Schwarz Inequality (special case of Hölder) $|\langle \underline{u}, \underline{v} \rangle| \leq \|\underline{u}\| \|\underline{v}\|$

Triangle Inequality $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\|$

☆ Gram-Schmidt Process:

$\{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n\}$: $\underline{w}_1 = \frac{\underline{v}_1}{\|\underline{v}_1\|}$, $\underline{u}_k = \underline{v}_k - \underbrace{[\langle \underline{v}_k, \underline{w}_1 \rangle \underline{w}_1 + \cdots + \langle \underline{v}_k, \underline{w}_{k-1} \rangle \underline{w}_{k-1}]}_{\text{orthogonal complement}}$,

$\underline{w}_k = \frac{\underline{u}_k}{\|\underline{u}_k\|}$, sequentially, $\forall k = 2, 3, \cdots, n$

(3) [Theorem] Riesz-Representation: Let $(V, \langle , \rangle)$ be a inner product space

Define a function $\phi: V \mapsto V^*$ $(\underline{v} \mapsto \phi_{\underline{v}}: V \mapsto \mathbb{F})$ as $\phi_{\underline{v}}(\underline{w}) = \langle \underline{v}, \underline{w} \rangle$

Then, ϕ is an $\boxed{\text{injective}}$ $\boxed{\mathbb{R}\text{-linear transformation}}$

$$\phi_{r\underline{v}+\alpha\underline{w}} = r\phi_{\underline{v}} + \alpha\phi_{\underline{w}}, \ \forall r, \alpha \in \mathbb{R}$$

(r.m.k., If V is a $\mathbb{C}$-vec. space, ϕ is NOT $\mathbb{C}$-linear, e.g. $\phi(i\underline{v}) \neq i\phi(\underline{v})$)

(proof.) well-definedness: $\phi_{\underline{v}}(a\underline{w}_1 + b\underline{w}_2) = \langle \underline{v}, a\underline{w}_1 + b\underline{w}_2 \rangle = a\phi_{\underline{v}}(\underline{w}_1) + b\phi_{\underline{v}}(\underline{w}_2)$

(either bi-linear / sesqui-linear), thus $\phi_{\underline{v}} \in V^*$. Suppose $\phi_{\underline{u}} = \phi_{\underline{v}} \Rightarrow \forall \underline{w} \in V$,

$\langle \underline{u} - \underline{v}, \underline{w} \rangle = 0$. Take $\underline{w} = \underline{u} - \underline{v} \Rightarrow \underline{u} - \underline{v} = 0_V$. (Thus, $\underline{u} = \underline{v}$).

$\mathbb{R}$-linear transformation: $\phi_{r\underline{v}+\alpha\underline{w}}(\underline{u}) = \langle r\underline{v} + \alpha\underline{w}, \underline{u} \rangle = r\langle \underline{v}, \underline{u} \rangle + \alpha\langle \underline{w}, \underline{u} \rangle$ $\forall \underline{u} \in V$, $\alpha, r \in \mathbb{R}$

Thus $\phi_{r\underline{v}+\alpha\underline{w}} = r\phi_{\underline{v}} + \alpha\phi_{\underline{w}}$, $\forall r, \alpha \in \mathbb{R}$, $\underline{v}, \underline{w} \in V$.

(Coro. If V is a real inner product, finitely generated space

$\phi: V \mapsto V^*$ is an isomorphism.)

(proof. any injective linear transformation $T: V \mapsto W$ with $\dim(V) = \dim(W) < \infty$ is an isomorphism.

☆ For complex vector space (inner product, f.g.): treat V with $\dim(V) = n$ as a $\dim = 2n$ real vector space. (e.g. $\mathbb{C}^n = \text{span}_{\mathbb{C}}\{\underline{e}_1, \cdots, \underline{e}_n\} = \{z_1\underline{e}_1 + \cdots + z_n\underline{e}_n \mid z_1 = a_1 + b_1 i, \ldots, z_n = a_n + b_n i\}$

$= \cdots = \text{span}_{\mathbb{R}}\{\underline{e}_1, i\underline{e}_1, \cdots, \underline{e}_n, i\underline{e}_n\}$) More generally, if $\dim_{\mathbb{C}}(V) = n$ basis $\{\underline{e}_1, \cdots, \underline{e}_n\}$

$\dim_{\mathbb{R}}(V) = 2n$ simutaneously, basis $\{\underline{e}_1, i\underline{e}_1, \dots \underline{e}_n, i\underline{e}_n\}$.

(Coro') If V is a complex inner product, f.g. space. $\phi: V \mapsto V^*$ is an isomorphism if V, V^* are treated as real inner product spaces.

(4) Adjoint Operator / Map

(i) [Defn / Thm] Let V be a f.g. inner product space, & $T: V \mapsto V$ be a linear operator. $\forall \underline{v} \in V$, $\exists$ a unique $\underline{a}_{\underline{v}} \in V$ s.t. $\langle \underline{a}_{\underline{v}}, \underline{w} \rangle = \langle \underline{v}, T(\underline{w}) \rangle$, $\forall \underline{w} \in V$.

(proof.) Consider $\theta_{\underline{v}}: V \mapsto \mathbb{F}$ l.t. (i.e., $\theta_{\underline{v}} \in V^*$), define $\theta_{\underline{v}}(\underline{w}) = \langle \underline{v}, T(\underline{w}) \rangle$, $\forall \underline{w} \in V$.

(check: $\theta_{\underline{v}} \in V^*$). By Riesz-representation theorem: $\exists\, \underline{a}_{\underline{v}} \in V$, unique, s.t. $\phi(\underline{a}_{\underline{v}}) = \theta_{\underline{v}}$ (since f.g. $\Rightarrow$ surjective). $\therefore \phi(\underline{a}_{\underline{v}})(\underline{w}) = \langle \underline{a}_{\underline{v}}, \underline{w} \rangle \overset{\text{also}}{=} \theta_{\underline{v}}(\underline{w}) = \langle \underline{v}, T(\underline{w}) \rangle$. $\forall \underline{w}$

[Defn'] Let V be a f.g. inner product space, $T: V \mapsto V$ (linear operator)

from thm above

Then, the [adjoint map] $T^*: V \mapsto V$ of T is defined as $[T^*(\underline{v}) := \underline{a}_{\underline{v}}]$

(r.m.k Riesz representation thm guarantees the existence & uniqueness)

In short, $T^*: V \mapsto V$ s.t. $[\langle T^*(\underline{v}), \underline{w} \rangle = \langle \underline{v}, T(\underline{w}) \rangle]$ $\forall \underline{v}, \underline{w} \in V$.

(e.g. $\langle \underline{v}, \underline{w} \rangle \triangleq \underline{v}^t \underline{w}$, $T(\underline{v}) = A\underline{v}$, then $T^*(\underline{w}) = A^t \underline{w}$ in $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$)

(ii) Properties of A-M.

[Prop I] The map $T^*: V \mapsto V$ is linear.

(proof.) $\langle T^*(a\underline{v} + b\underline{w}), \underline{u} \rangle = \langle a\underline{v} + b\underline{w}, T(\underline{u}) \rangle = \bar{a}\langle \underline{v}, T(\underline{u}) \rangle + \bar{b}\langle \underline{w}, T(\underline{u}) \rangle$

$= \langle aT^*(\underline{v}) + bT^*(\underline{w}), \underline{u} \rangle$, thus, $T^*(a\underline{v} + b\underline{w}) = aT^*(\underline{v}) + bT^*(\underline{w})$. (since $\forall \underline{u}$)

[Prop II] Let V be f.g. inner product space, $\mathcal{B}$ is an orthonomal basis of V.

Then, matrix representation of $T: V \mapsto V$ & $T^*: V \mapsto V$ are related by $[(T^*)_{\mathcal{B}\mathcal{B}} = (T)^H_{\mathcal{B}\mathcal{B}}]$

(proof.) Suppose $(T^*)_{\mathcal{B}\mathcal{B}} = (a_{ij})$ & $(T)_{\mathcal{B}\mathcal{B}} = (b_{ij})$, $\mathcal{B} = \{\underline{e}_1, \dots, \underline{e}_n\}$

$\langle T^*(\underline{e}_k), \underline{e}_l \rangle = \langle \underline{e}_k, T(\underline{e}_l) \rangle \Leftrightarrow \langle \sum_{i=1}^{n} a_{ik}\underline{e}_i, \underline{e}_l \rangle = \langle \underline{e}_k, \sum_{j=1}^{n} b_{jl}\underline{e}_j \rangle \Rightarrow \overline{a_{lk}} = b_{kl}$. ▥

(iii) Self-Adjoint Map:

[Defn] Let V be inner product space. A linear operator $T: V \mapsto V$ is self-adjoint if $[T^* = T]$

(Note: in $(\mathbb{R}^n, \langle\cdot,\cdot\rangle)$, self-adjoint $T(\underline{x}) = A\underline{x} \Leftrightarrow A$ is symmetric)

All nice results can be generalized to self-adjoint operators.

Namely. T is diagonalizable, always has orthonormal bases, real ei-values.

For V over $\mathbb{C}$ (e.g. $\mathbb{C}^n$) $T(\underline{v}) = A\underline{v}$ is self-adjoint if $A^H = A$.

[Proposition I] All ei-values for self-adjoint map $T: V \mapsto V$ are real. (V – f.g.)

(proof.) Suppose $T(\underline{v}) = \lambda\underline{v}$, $(\underline{v} \neq \underline{0}_V)$ $\langle T(\underline{v}), \underline{v}\rangle = \langle \underline{v}, T(\underline{v})\rangle \Rightarrow \bar{\lambda} = \lambda$.

[Proposition II] If $\begin{cases} T(\underline{v}) = \lambda\underline{v} \\ T(\underline{w}) = \mu\underline{w} \end{cases}$ & $\underline{v}, \underline{w} \neq 0$, $\lambda \neq \mu$ (distinct), $\langle \underline{v}, \underline{w}\rangle = 0$

(proof.) $\langle T(\underline{v}), \underline{w}\rangle = \langle \underline{v}, T(\underline{w})\rangle \Rightarrow (\underbrace{\lambda - \mu}_{\text{both real}})\langle \underline{v}, \underline{w}\rangle = 0 \Rightarrow \langle \underline{v}, \underline{w}\rangle = 0$. ▥

[proposition III] Let V be a f.g. inner product space, & $T: V \mapsto V$ is self-adjoint Then. T always has an ei-vector.

(proof.) Case I. V over $\mathbb{C}$ space, then from fundamental thm of algebra ~.

Case II. V over $\mathbb{R}$ space, $\forall$ self-adjoint $T: V \mapsto V$, $V = \text{span}\{\underline{v}_1, \dots, \underline{v}_n\}$, (basis) $\dim(V) = n$. Consider. $V_{\mathbb{C}} := \text{span}_{\mathbb{C}}\{\underline{v}_1, \dots, \underline{v}_n\}$, $T_{\mathbb{C}}: V_{\mathbb{C}} \mapsto V_{\mathbb{C}}$ defined by $[T_{\mathbb{C}}(\underline{v}) = T_{\mathbb{C}}(\sum_{i=1}^n z_i \underline{v}_i) := \sum_{i=1}^n z_i T(\underline{v}_i)]$ since from defn, $T_{\mathbb{C}}(\underline{v}_i) \overset{\forall i}{=} T(\underline{v}_i) \Rightarrow T_{\mathcal{B}\mathcal{B}} = (T_{\mathbb{C}})_{\mathcal{B}\mathcal{B}}$ for $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$. Define $\langle\cdot,\cdot\rangle_{\mathbb{C}}: V_{\mathbb{C}}^2 \mapsto \mathbb{C}$ given by $\langle \sum_i z_i \underline{v}_i, \sum_j w_j \underline{v}_j\rangle_{\mathbb{C}} := \sum_i \sum_j \bar{z}_i w_j \langle \underline{v}_i, \underline{v}_j\rangle$. Then, $T_{\mathbb{C}}: V_{\mathbb{C}} \mapsto V_{\mathbb{C}}$ is a self-adjoint operator on the inner product space. Apply case I, $T_{\mathbb{C}}$ has an ei-vector with ei-value $\mu_i \in \mathbb{C}$, then $\mu_i \in \mathbb{R}$ since $\chi_{T_{\mathbb{C}}} = \chi_T \Rightarrow \mu_i \in \mathbb{R}$ is the ei-value of T. ▥

[proposition IV] Let $T: V \mapsto V$ be self-adjoint f.g. V. If $[U \leq V$ is T-invariant, $U^\perp \leq V$ must be T-invariant.$]$

(proof.) Take $\underline{w} \in U^{\perp}$, $\langle \underline{w}, \underline{v} \rangle = 0$, $\forall \underline{v} \in U$. $\langle T(\underline{w}), \underline{v} \rangle = \langle \underline{w}, T(\underline{v}) \rangle = 0$ since $\forall \underline{v} \in U$ $T(\underline{v}) \in U$ (T-invariant U). Thus, $T(\underline{w}) \in U^{\perp}$ by defn. ▥

[Thm] Let V be f.g. inner-product vec. space, $T: V \mapsto V$ is [self-adjoint.]
Then, V has an [orthonormal basis of ei-vectors] with [real ei-values.]

(proof.) Do induction on $\dim(V) = n$, when $n = 1$, obvious
Suppose true for all self-adjoint $T: V \mapsto V$ with $\dim(V) = n \leq k \in \mathbb{N}$
Let $S: W \mapsto W$, $\dim(W) = k+1$, self-adjoint. By prop III, S has ei-value μ with ei-vector $\underline{v} \in W$. Let $U = \operatorname{span}\{\underline{v}\}$, then U is S-invariant, thus (prop IV) $U^{\perp}$ is S-invariant
$S|_{U^{\perp}}: U^{\perp} \mapsto U^{\perp}$ (with $\dim(U^{\perp}) = k$) is self-adjoint. $\Rightarrow \exists$ an orthonormal basis $\{u_1, \dots, u_k\}$ of $U^{\perp}$. then $\{\underline{v}, \underline{u}_1, \dots, u_k\}$ is (let $\|\underline{v}\| = 1$) an orthonormal basis of U.

(5) Unitary Operators (only for vec. space over $\mathbb{C}$)
(i) [Defn] Let V be a $\mathbb{C}$-inner product space. An operator $T: V \mapsto V$ is called [unitary]
if [$\langle T(\underline{u}), T(\underline{v}) \rangle = \langle \underline{u}, \underline{v} \rangle$, $\forall \underline{u}, \underline{v} \in V$.]
(r.m.k If $T: V \mapsto V$ is unitary, $\|T(\underline{u})\| = \|\underline{u}\|$, T perserves the norm of $\underline{u}$)
(e.g. $V = \mathbb{C}^n$, $\langle \underline{u}, \underline{v} \rangle := \underline{u}^H \underline{v}$. $T(\underline{x}) = A(\underline{x})$ is $\boxed{\text{unitary}}$ iff $A^H A = \underline{I}$) → orthogonal/unitary matrix
(ii) [Thm] Let $T: V \mapsto V$ be unitary. V f.g. Then, V has an [orthonomal basis of ei-vectors of T.] Moreover, the ei-values [$\mu \in \mathbb{C}$ of T must have $\|\mu\| = 1$.]
(lemma 1) let V be $\mathbb{C}$ vec. space. ($T: V \mapsto V$ is unitary) $\Longleftrightarrow$ ($T^* \circ T: V \mapsto V$ is identity)
(proof.) $T^* \circ T$ is identity $\Leftrightarrow \langle T^* \circ T(\underline{u}), \underline{v} \rangle = \langle \underline{u}, \underline{v} \rangle$ $\forall \underline{u}, \underline{v}$ $\Leftrightarrow \langle T(\underline{u}), T(\underline{v}) \rangle = \langle \underline{u}, \underline{v} \rangle$ $\forall \underline{u}, \underline{v}$ $\Leftrightarrow T$ is unitary.
check "$\Leftarrow$"
(lemma 2) let $\underline{v}$ be an ei-vector of a unitary operator $T: V \mapsto V$. Then, its ei-value $\lambda \in \mathbb{C}$ s.t. $|\lambda| = 1$. Thus, $\langle T(\underline{v}), T(\underline{v}) \rangle = \langle \underline{v}, \underline{v} \rangle$ gives $\bar{\lambda}\lambda \langle \underline{v}, \underline{v} \rangle = \langle \underline{v}, \underline{v} \rangle$
$\Rightarrow |\lambda|^2 = |\lambda| = 1$.

(lemma 3) Let $T: V \mapsto V$ be unitary, $U \leq V$ is T-invariant. Then, $U \leq V$ is [T^*-invariant.] Moreover, with this lemma, $U^\perp$ is T-invariant.

(proof.) Assume lemma 3 is true, $\langle \underline{u}, T(\underline{w}) \rangle = \langle T^*(\underline{u}), \underline{w} \rangle = 0$, with any $\underline{w} \in U^\perp$, $\underline{u} \in U$ & by defn. $T(\underline{w}) \in U^\perp \Rightarrow U^\perp$ is T-invariant.

proving lemma 3: with lemma 1, $T^* \circ T = T \circ T^* = Id$, which means $T^{-1} = T^*$.

$\Rightarrow \chi_T(x) = 0$ does NOT have a root $\underset{\Downarrow}{0}$. By Cayley-Hamilton thm. $\chi_T(T) = O_{Hom(V,V)}$

$\Rightarrow T^n + a_{n-1}T^{n-1} + \cdots + a_1 T = -a_0 I \quad (a_0 \neq 0) \Rightarrow \frac{-1}{a_0}(T^n + a_{n-1}T^{n-1} + \cdots + a_1 T) = I$

$\therefore T^* = -\frac{1}{a_0}(T^{n-1} + a_{n-1}T^{n-2} + \cdots + a_1 I)$ since U is T-invariant, U is T^*-invariant explicitly, from what T^* is. $\boxed{\text{ } }$

Thus, from above lemmas, thm can be proved similarly to that w.r.t. self-adjoint map.

(6) Normal Operator

(i) [Defn] Let V be $\mathbb{C}$-vec. space. Say $T: V \mapsto V$ is normal if $T^* \circ T = T \circ T^*$

(r.m.k. Self-adjoint & unitary operators are normal ← "generalization")

(ii) (lemma 1) Let $T: V \mapsto V$ be normal, then

⟨i⟩ $\|T(\underline{v})\| = \|T^*(\underline{v})\|$, $\forall \underline{v} \in V$.

⟨ii⟩ $T - \lambda I$ is also normal, $\forall \lambda \in \mathbb{C}$

⟨iii⟩ $T(\underline{v}) = \lambda \underline{v} \Rightarrow T^*(\underline{v}) = \bar{\lambda} \underline{v}$

⟨iv⟩ If $\begin{cases} T(\underline{v}) = \lambda \underline{v} \\ T(\underline{w}) = \mu \underline{w} \end{cases}$ & $\lambda - \mu \neq 0 \Rightarrow \langle \underline{v}, \underline{w} \rangle = 0$

(proof.) $\|T(\underline{v})\|^2 = \langle T(\underline{v}), T(\underline{v}) \rangle = \langle T^* \circ T(\underline{v}), \underline{v} \rangle = \langle T \circ T^*(\underline{v}), \underline{v} \rangle = \langle T^*(\underline{v}), T^*(\underline{v}) \rangle = \|T^*(\underline{v})\|^2$. For the 2nd one, $(T - \lambda I)^* = T^* - \bar{\lambda} I$, directly take product by definition. $\underbrace{\|T(\underline{v}) - \lambda \underline{v}\|}_{0 =} = \|(T - \lambda I)(\underline{v})\| \overset{\langle i \rangle, \langle ii \rangle}{=\!=} \|(T^* - \bar{\lambda} I)(\underline{v})\| \Rightarrow T^*(\underline{v}) = \bar{\lambda}\underline{v}$.

For the last one, with ⟨iii⟩, $\langle T^*(\underline{v}), \underline{w} \rangle = \langle \underline{v}, T(\underline{w}) \rangle \Rightarrow (\lambda - \mu)\langle \underline{v}, \underline{w} \rangle = 0$

(iii) [Thm] Let V be a f.g. $\mathbb{C}$-inner product space & $T: V \mapsto V$ is normal. Then, $\exists$ a basis $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ (orthonormal) s.t. $\underline{v}_i$ is an ei-vector of T, $\forall i$

(proof.) Similar to those for self-adjoint & unitary operators. (need T-invariance $U^\perp$)

(lemma 2) Let $\underline{w}$ be ei-vec of T (normal), $U := \text{span}\{\underline{w}\}$, $U^\perp \leq W$ is T-invariant.

(proof.) $\forall\, \underline{u} \in U$, let $\underline{v} \in U^\perp$, $\langle \underline{u}, T(\underline{v})\rangle = \langle T^*(\underline{u}), \underline{v}\rangle = \bar{\lambda}\langle \underline{u}, \underline{v}\rangle = 0$

(lemma 3) let U be defined as in lemma 2, then $U^\perp \leq W$ is T^*-invariant.

With above 2 lemmas, do induction (since it's safe to write $T^*|_{U^\perp} \circ T|_{U^\perp} = T|_{U^\perp} \circ T^*|_{U^\perp} : U^\perp \mapsto U^\perp$ is normal for $T|_{U^\perp}$, then use induction)

(iv) Spectral Theorem: Let $T: V \mapsto V$ be normal on a f.g. $\mathbb{C}$-inner product space V.
Then, $\exists$ $\boxed{\text{self-adjoint}}$ operators $P_1, \cdots, P_k : V \mapsto V$, s.t. ← "decomposition" of T

① $P_i^2 = P_i$, $\forall\, i = 1, 2, \cdots, k$ (idempotent / projection matrices)

② $P_i \cdot P_j = 0$, if $i \neq j$

③ $\sum_{i=1}^{k} P_i = Id$, & $\left[\sum_{i=1}^{k} \lambda_i P_i = T, \text{ for some } \lambda_1, \cdots, \lambda_k \in \mathbb{C}\right]$ → decompose "T" into a linear combination of projection operators P_i

(Proof.) Since $T: V \mapsto V$ is normal, T is diagonalizable

$\therefore m_T(x) = \prod_{i=1}^{k}(x - \lambda_i)$ (distinct) $\therefore$ We have a primal decomposition

$V = V_1 \oplus \cdots \oplus V_k$, $V_i \triangleq \ker(T - \lambda_i I)$, $i = 1, \cdots, k$ (eigen-spaces)

Recall the proof of PDT: let $q_i(x) = \frac{m_T(x)}{(x - \lambda_i)}$, Bezout's theorem gives that $\exists\, a_i(x)$, $1 \leq i \leq k$ s.t. $\sum_{i=1}^{k} a_i(x) q_i(x) = 1$. Let $P_i \triangleq a_i(T) q_i(T) \Rightarrow \sum_{i=1}^{k} P_i = Id$.

If $i \neq j$, $m_T(x) \mid a_i(x) a_j(x) q_i(x) q_j(x) \Rightarrow P_i P_j = 0$, (also $P_j P_i = P_i P_j = 0$).

Now, $P_i = P_i \cdot Id = \sum_{j=1}^{k} P_i P_j = P_i^2$. Moreover, P_i is self-adjoint since it's normal & has ei-values 0, 1 (real), where V is f.g. Consider $T(\underline{v}) = T \circ Id(\underline{v}) = T\left(\sum_{i=1}^{k} P_i(\underline{v})\right) = \sum_{i=1}^{k} T \circ P_i(\underline{v}) = \sum_{i=1}^{k} \lambda_i P_i(\underline{v})$, i.e. $T(\underline{v}) = \left(\sum_{i=1}^{k} \lambda_i P_i\right)(\underline{v})$, since $V_i = \ker(T - \lambda_i I)$ & $P_i(\underline{v}) \in V_i$.

- Tensor Product

(1) Motivations & Introductions:

(i) Understand "k-linear" maps: $f: V_1 \times V_2 \times \cdots \times V_k \mapsto W$, i.e. $[f(\underline{v}_1, \cdots, a\underline{v}_i + b\underline{v}_i', \cdots, \underline{v}_k) = af(\underline{v}_1, \cdots, \underline{v}_i, \cdots, \underline{v}_k) + bf(\underline{v}_1, \cdots, \underline{v}_i', \cdots, \underline{v}_k)]$

(e.g.s Bi-linear form $V \times V \mapsto \mathbb{R}$; $\underset{\text{cross-prod}}{\underline{u} \times \underline{v}}: \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$ (2-linear)

$\det(\cdot): \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n\text{ terms}} \mapsto \mathbb{R}$ with $\det(\underline{v}_1, \cdots, \underline{v}_n) = \det[\underline{v}_1 | \cdots | \underline{v}_n]$ (n-linear map))

Qn: can a k-linear map $f: V_1 \times V_2 \times \cdots \times V_k \mapsto W$ be a linear transformation between vec. spaces ? → Mostly NO! ← for non-trivial cases

(e.g. Bi-linear form $B(a(\underline{v}_1, \underline{v}_2)) = a^2 B(\underline{v}_1, \underline{v}_2)$ ⇐ never true for non-trivial B)

Soln: Construct $\boxed{\text{a tensor product space}}$ $V_1 \otimes \cdots \otimes V_k$ s.t. $\exists$ an $\boxed{\text{injection map}}$ $i: V_1 \times \cdots \times V_k \mapsto (V_1 \otimes \cdots \otimes V_k)$ "bigger" & $\forall$ k-linear map $f: V_1 \times \cdots \times V_k \mapsto W$, $\exists$ a corresponding linear transformation $\Phi: V_1 \otimes V_2 \otimes \cdots \otimes V_k \mapsto W$ s.t. $\boxed{f = \Phi \circ i}$

(i.e., $f(\underline{v}_1, \cdots, \underline{v}_k)$ can be understood as $\boxed{\Phi(\underline{v}_1 \otimes \underline{v}_2 \otimes \cdots \otimes \underline{v}_k)}$, completely)

containing all information of "f"

(2) Construction of Tensor Product Space

(i) [Defn] Let V, W be 2 vec. spaces over $\mathbb{F}$. Define a set $[\mathcal{S} := \{(\underline{v}, \underline{w}) \mid \underline{v} \in V, \underline{w} \in W\}$ & a vec. space $\mathcal{X} := \text{span}(\mathcal{S})]$. (WARNING: elements in $\mathcal{X}$ have form $(a_1(\underline{v}_1, \underline{w}_1) + \cdots + a_k(\underline{v}_k, \underline{w}_k))$ but no-relationships $\boxed{\text{among diff. } (\underline{v}_i, \underline{w}_i)\text{s}}$. e.g. $(\underline{v}, \underline{0}_W)$, $(\underline{0}_V, \underline{w})$ & $(\underline{v}, \underline{w})$ are linearly independent!)

[Defn'] Let $\mathcal{Y} \leq \mathcal{X}$ be a vec. subspace spanned by the following vectors:

$(\underline{v}, \underline{w}_1 + \underline{w}_2) - (\underline{v}, \underline{w}_1) - (\underline{v}, \underline{w}_2)$; $(\underline{v}_1 + \underline{v}_2, \underline{w}) - (\underline{v}_1, \underline{w}) - (\underline{v}_2, \underline{w})$;

$(k\underline{v}, \underline{w}) - k(\underline{v}, \underline{w})$ & $(\underline{v}, l\underline{w}) - l(\underline{v}, \underline{w})$, for all possible $\underline{v}, \underline{v}_1, \underline{v}_2 \in V$, $\underline{w}, \underline{w}_1, \underline{w}_2 \in W$, $k, l \in \mathbb{F}$

($\mathcal{Y}$ – a big vec. subspace, with uncountable dimension)

Then, define $\boxed{\text{tensor product}}$ $[V \otimes W := \mathcal{X}/\mathcal{Y}]$

For $(\underline{v}, \underline{w}) \in V \times W$, corresponding $[\underline{v} \otimes \underline{w} := (\underline{v}, \underline{w}) + \mathcal{Y}] \in \mathcal{X}/\mathcal{Y}$

"quotient"

(e.g. $(2\underline{v}) \otimes \underline{w} = (2\underline{v}, \underline{w}) + \mathcal{Y} = 2(\underline{v}, \underline{w}) + \mathcal{Y} = 2(\underline{v} \otimes \underline{w})$

similarly, $(\underline{v}_1 + \underline{v}_2) \otimes \underline{w} = \underline{v}_1 \otimes \underline{w} + \underline{v}_2 \otimes \underline{w}$

> use $\mathcal{Y}$ to make "+" "&" "·" operations unique! for whatever in V, W or $V \otimes W$.

(WARNING: a general element in $V \otimes W$ is of the form

$\underline{v}_1 \otimes \underline{w}_1 + \cdots + \underline{v}_k \otimes \underline{w}_k$ (NOT $\underline{v} \otimes \underline{w}$)

operations diff. from those in ordinary $V \times W$ spaces)

← weak version

(ii) [Thm] (Universal property) Let $f: V\times W \mapsto \mathcal{U}$ be a bi-linear map. $\exists\ \Phi: V\otimes W \mapsto \mathcal{U}$ s.t. Φ is linear, $[\Phi(\underline{v}\otimes\underline{w}) = f(\underline{v},\underline{w})]$ $\forall(\underline{v},\underline{w})\in V\times W$.

(proof.) Consider a linear transformation $T: \mathcal{X}\mapsto\mathcal{U}$, given by $T(\sum_i a_i(\underline{v}_i,\underline{w}_i)) \triangleq \sum_i a_i f(\underline{v}_i,\underline{w}_i)$ for all finite sums in $\mathcal{X}$. $T((\underline{v}_1+\underline{v}_2,\underline{w})-(\underline{v}_1,\underline{w})-(\underline{v}_2,\underline{w}))=0$, $\forall \underline{v}_1,\underline{v}_2\in V$ $\underline{w}\in W$. Thus, similarly all elements in $\mathcal{Y}$ s.t. $T(\underline{y})=0$, if $\underline{y}\in\mathcal{Y}$. (check all elements spaming $\mathcal{Y}$). $\therefore \mathcal{Y}\leq \ker(T)$, it's reasonable to define $\tilde{T}(\underline{v}\otimes\underline{w}) = T((\underline{v},\underline{w}))$ (thus $f(\underline{v},\underline{w})$)

↖well-definedness of $\tilde{T}$

Only remaining is to check linearity of T (thus, for sure $\tilde{T}\triangleq\Phi$)

☆

(Applications) ① V has basis $\{\underline{v}_1,\dots,\underline{v}_n\}$, W has basis $\{\underline{w}_1,\dots,\underline{w}_m\}$ $\Rightarrow V\otimes W$ has basis $\{\underline{v}_i\otimes\underline{w}_j \mid {1\leq i\leq n \atop 1\leq j\leq m}\}$

$\dim(V\otimes W) = \dim(V)\dim(W)$

(If $\dim(V)<\infty$, $\dim(W)<\infty$)

(check: linear independence $\Leftarrow$ universal property)

ⓘⓘ If $T: V\mapsto V'$ & $S: W\mapsto W'$ are 2 linear transformations, we can construct a linear map $T\otimes S: V\otimes W\mapsto V'\otimes W'$, by $T\otimes S(\underline{v}\otimes\underline{w}) \triangleq f(\underline{v},\underline{w})$ where a bi-linear map $f: V\times W\mapsto V'\otimes W'$ is defined by $f(\underline{v},\underline{w}) = T(\underline{v})\otimes S(\underline{w})$ } thus $T\otimes S(\underline{v}\otimes\underline{w}) = T(\underline{v})\otimes S(\underline{w})$

ⓘⓘⓘ Let $T: \mathbb{R}^m\mapsto\mathbb{R}^m$ s.t. $T(\underline{v}) = A\underline{v}$ & $S: \mathbb{R}^n\mapsto\mathbb{R}^n$ s.t. $S(\underline{w}) = B\underline{w}$

Then, under $\mathcal{C} := \{\underline{e}_i\otimes\underline{f}_j \mid {1\leq i\leq m \atop 1\leq j\leq n}\}$, where $\mathcal{A} = \{\underline{e}_1,\dots,\underline{e}_m\}$, $\mathcal{B} = \{\underline{f}_1,\dots,\underline{f}_n\}$

$$(T\otimes S)_{\mathcal{C}\mathcal{C}} = (T)_{\mathcal{A}}\otimes(S)_{\mathcal{B}} \quad (= A\otimes B)$$

tensor product ↑ ; Kronecker product ↓

Kronecker product is a special case of tensor product.

(3) Determinants

(i) [Defn] Let $f: \underbrace{V\times\cdots\times V}_{\#:p} \mapsto W$ be a p-linear map. We say f is $\boxed{\text{alternating}}$ if $f(\dots,\underline{v},\dots,\underline{v},\dots) = \underline{0}_W$ (whenever) $\exists$ 2 repeated terms on the left.

(e.g. $f: \mathbb{R}^3\times\mathbb{R}^3\mapsto\mathbb{R}^3$ s.t. $f(\underline{u},\underline{v}) = \underline{u}\times\underline{v}$ "cross product")

[Defn'] The (p-linear tensor product space) of V is $\Lambda^p V := V^{\otimes p}/\mathcal{Z}$, where $V^{\otimes p} \triangleq \underbrace{V\otimes\cdots\otimes V}_{\#:p}$

← define by quotient

& $\mathcal{Z} = \mathrm{span}\{(\dots,\underline{v},\dots,\underline{v}\dots)\in V^{\otimes p}\}$. Also define $(\underline{v}_1\wedge\underline{v}_2\wedge\cdots\wedge\underline{v}_p := \underline{v}_1\otimes\cdots\otimes\underline{v}_p + \mathcal{Z})$

(egs / arithmetic) ⓘ $v_1 \wedge \cdots \wedge v \cdots \wedge v \cdots \wedge v_p = 0_{\wedge^p V}$; thus further we get "swap" for ⓘ

← switching order

as ⓘⓘ $v_1 \wedge \cdots \wedge w \wedge \cdots \wedge v \wedge \cdots \wedge v_p = - v_1 \wedge \cdots \wedge v \wedge \cdots \wedge w \wedge \cdots \wedge v_p$ (proved by using p-linearity & ⓘ) "2 Gold Rules": same inputs ⇒ 0, swapping inputs ⇒ "−" sign & p-linearity

(Hint: from these 3 properties, it's enough to define "determinant")

(ii) Properties of $\wedge^p V$:

[Prop. I] Let $B = \{v_1, \cdots, v_n\}$ be a basis of V. Then, a basis of $\wedge^p V$ is

$\left[\{ v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_p} \mid i_1 < i_2 < \cdots < i_p \}\right]$ (dimension: $\dim(\wedge^p V) = \binom{n}{p}$)

(proof.) Use "2 Gold Rules" directly. Note that $p \leq n$, is a must.

[Prop. II] Suppose $T_1, \cdots, T_p : V \mapsto W$ are p linear transformations, $\exists$ a linear transformation $T_1 \wedge \cdots \wedge T_p : \wedge^p V \mapsto \wedge^p W$

Given by $\left[(T_1 \wedge \cdots \wedge T_p)(v_1 \wedge \cdots \wedge v_p) = T_1(v_1) \wedge \cdots \wedge T_p(v_p)\right]$, $\forall\, v_1 \wedge \cdots \wedge v_p \in \wedge^p V$.

(iii) The Determinant of a Linear Transformation

[Defn] Let V be a vector space with $\dim(V) = n < \infty$, and $T: V \mapsto V$ be linear, thus $\underbrace{T \wedge \cdots \wedge T}_{\#: n \text{ terms}} : \wedge^n V \mapsto \wedge^n V$ is linear & $T^{\wedge n}(\alpha) = \Delta\, \alpha$, for some $\Delta \in F$, $\forall\, \alpha \in \wedge^n V$.

$T^{\wedge n}$ 1-dimensional

The determinant of T, $\det(T)$, is defined as $\left[\det(T) = \Delta\right]$, then.

(e.g. special case $T: \mathbb{R}^n \mapsto \mathbb{R}^n$, $T(x) = Ax$, $\det(T) = \det(A)$, with $e_1 \wedge \cdots \wedge e_n$ used

⇒ $T^{\wedge n}(e_1 \wedge \cdots \wedge e_n) = \det(A)\, e_1 \wedge \cdots \wedge e_n$.)

It's also possible to check all the properties of $\det(A)$ from $\det(T)$.

↳ e.g. $\det(T \circ S) = \det(T)\det(S)$, if $T(x) = Ax$ & $S(x) = Bx$.

$(T \circ S)^{\wedge n}(e_1 \wedge \cdots \wedge e_n) = \det(T)\, S(e_1) \wedge \cdots \wedge S(e_n) = \det(T)\det(S)\, e_1 \wedge \cdots \wedge e_n$.