Tutorial 0

TA information: Haiyu CHEN. hychen@math.cuhk.edu.hk.

The following answer is author's opinion on subjects and could be bias.

1. Organization of tutorial.

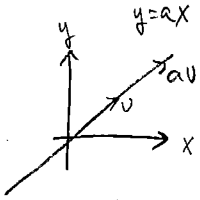
   - Hightlight important thingsin lecture
   - work out examples or sample problems
   - answer questions.
   - ANYTHING YOU PROPOSE!

   Do not hasitate to let me know your advice! I will change the style to suit your need!

2. What is linear algebra?

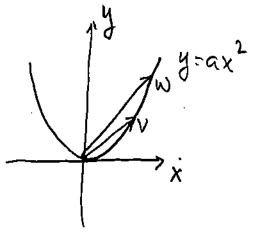
   Algebra : $+, -, \times, \div$, other operations

   linear : line



1' $v \in V$, $av \in V$. scaling.

2' $v, w \in V$, $v+w \in V$ addition



$v+w \notin \cup$

$av \notin \cup$

not linear.

This is studied in algebraic geometry.

3. MATH1030 V.S. MATH2040?

   1030 concrete vectors in $\mathbb{R}^n$, $\mathbb{C}^n$, matrices.

   2040
   - abstraction of 1030 so that it has wider applications!
   - A vector is <u>anything</u> that satisfy properties 1', 2' above.
     so we don't need to limit ourselves to $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{F})$.
   - We can do linear algebra as long as we have a field $\mathbb{F}$ and things that behaves like a vector.

4. What does the subject linear algebra consist of (for a first course)?

Linear algebra

① Vector spaces V.
- Classification of vector spaces? — dimension — basis
  - span
  - Linear independence

  What are the objects of a given type, up to some equivalence?

② Maps between them $V \to V$.
- Classification of linear transformations?
  - invertibility
  - Jordan normal form
    Jordan decomposition.
  - primary decomposition
    - eigenvalue
    - eigenspace, Cayley-Hamilton
    - diagonalizability.
  - matrix representation — 1130.
- kernel, range, rank-nullity thm.

③ inner product spaces
- notion of angle, length?
  - orthogonality
  - Gram-Schmidt process

④ interactions between ② and ③.
- adjoint operators — matrix transpose
- normal, self adjoint, unitrary operators — symmetry, unitary matrices.
- Spectrum theorem: diagonalization of above linear operators.

⑤* Bilinear forms, quadratic forms and multilinear algebra.

Applications to later courses: Topology of manifold
Functional analysis
Abstract algebra
Modules and representation theory.
applied math, physics, computer science etc.

The dictionary

These connections are rough connections, or generalizations.

2040	1030
vector space	vectors
linear combination	column space
dimension	rank
standard basis	reduced echelon forms
linear maps	matrices
kernel	null space
range	column space
composition	matrix multiplication.
invertible maps	invertible matrices
change of basis	similar matrices.
orthonormal basis	orthonormal matrices
eigenspace	eigen vector, eigenspace.
norm, inner product	length, angle.
invariant space	block matrices
adjoint operator	matrix transpose
self-adjoint operator	symmetric matrices.
	Hermitian matrices
spectral decomposition	congruent matrices.

Tutorial 1: Vector space, subspace, span and linear independence.

Def 1.1. (field).

A <u>field</u> is a set $\mathbb{F}$ with two binary operations $+, \cdot : \mathbb{F}\times\mathbb{F} \to \mathbb{F}$ satisfying the following axioms: ($a,b,c \in \mathbb{F}$ unless specified).

(F1). $(a+b)+c = a+(b+c)$. associativity for $+$
$(ab)c = a(bc)$. associativity for $\cdot$.

(F2). $a+b = b+a$ commutativity for $+$
$ab = ba$ commutativity for $\cdot$

(F3). $\exists\ x \in \mathbb{F}$ s.t. $x + a = a$ for all $a \in \mathbb{F}$. Denote x as "0". identity for $+$
$\exists\ y \in \mathbb{F}$ s.t. $y \cdot a = a$ for all $a \in \mathbb{F}$ Denote y as "1". identity for $\cdot$.

(F4). $\forall\, a \in \mathbb{F}$, $\exists\, b \in \mathbb{F}$ s.t. $a+b=0$. Denote b as "$-a$". inverse for $+$.
$\forall\, a \neq 0$ in $\mathbb{F}$, $\exists\, b \in \mathbb{F}$ s.t. $ab = 1$ Denote b as a^{-1}. inverse for $\cdot$

(F5). $a(b+c) = ab + ac$ distributivity of $\cdot$ over $+$.

Remark (i) Note that (F1)–(F3) is symmetric for $(\mathbb{F}, +)$ and $(\mathbb{F}, \cdot)$.
(F4) is not as we require $a \neq 0$.
(F5) is not either.

(ii). We usually say the triple $(\mathbb{F}, +, \cdot)$ is a field as there can be more than one definition of $+$ and $\cdot$ making the same set into a different field.
If there is no danger of ambiguity, we will only write $\mathbb{F}$.

(iii). If you take MATH2070, $(\mathbb{F},+)$ and $(\mathbb{F}\setminus\{0\}, \cdot)$ are abelian groups.

Def 1.2. (vector space).

Let $(\mathbb{F}, \oplus, \cdot)$ be an arbitrary field. A <u>vector space</u> is a set V with a binary operation $+ : \mathbb{F}\times\mathbb{F} \to \mathbb{F}$ and a <u>scalar multiplication</u> $* : \mathbb{F}\times V \to V$ satisfying the following axioms: ($u,v,w \in V$, $a,b \in \mathbb{F}$ unless specified).

(VS1). $u+(v+w) = (u+v)+w$. associativity for $+$.

(VS2). $u+v = v+u$ commutativity for $+$.

(VS3) $\exists\, x \in V$ s.t. $x+u = u$ for all $u \in V$ Denote x as "0" identity for $+$

(VS4) $\forall\, u \in V$, $\exists\, y \in V$ s.t. $u+y=0$ Denote y as "$-u$" inverse for $+$.

((VS1)–(VS4) make $(V,+)$ an abelian group. Note that $*$ is not relevant so far.)

(VS 5) $a*(b*v) = (a\underset{\uparrow \text{ in } \mathbb{F}}{\cdot} b)*v$ compatibility of $\cdot$ in $\mathbb{F}$ and $*$.

(VS 6) $1*v = v$ compatibility of $1=1_{\mathbb{F}}$ for $\cdot$ in $\mathbb{F}$ and $*$.

(VS 7) $a*(u+v) = a*u + a*v$ distributivity of $*$ over $+$ in V.

(VS 8) $(a\oplus b)*v = a*v + b*v$ distributivity of $*$ over $\oplus$ in $\mathbb{F}$.

Remark. (i) It is very important to distinguish $\oplus$ in $\mathbb{F}$ with $+$ in V and $\cdot$ in $\mathbb{F}$ with $*: \mathbb{F}\times V \to V$!!!

(ii) If you take MATH 2070, $(V,+)$ is an abelian group.

(iii) Once you are fluent, we write $\oplus$ as $+$, $*$ as $\cdot$, keeping in mind that they are different. We shall do it from next tutorial.

Q1. Show the following properties for a vector space V over $\mathbb{F}$.

(i) $0 \in V$ is unique, called the <u>zero vector</u>.

(ii) $\forall u \in V$, the inverse $-u$ is unique, called the <u>additive inverse</u> of u.

(iii) If $u+v = w+v$, then $u=w$. <u>cancellation law</u>.

(iv). $0*v = 0$ for all $v \in V$, where the first $0=0_{\mathbb{F}}$ in $\mathbb{F}$, and second $0=0_V$ in V.

(v). $a\cdot 0 = 0$ for all $a \in \mathbb{F}$. where both $0 = 0_V$.

(vi). $(-1)*v = -v$ for all $v \in V$, where -1 is the additive inverse to 1 with respect to $\oplus$ in $\mathbb{F}$ and $-v$ is the additive inverse to v with respect to $+$ in V.

Pf. Exercise. We will just show (iv).

$$0_V + 0_{\mathbb{F}}*v = 0_{\mathbb{F}}*v = (0_{\mathbb{F}}+0_{\mathbb{F}})*v = 0_{\mathbb{F}}*v + 0_{\mathbb{F}}*v$$

From (iii), we have $0_V = 0_{\mathbb{F}}*v$.

Q2. Further properties.

(vii) $-(-v) = v$.

(viii). $(-a)*v = -(a*v) = a*(-v)$. $a \in \mathbb{F}$, $v \in V$.

(ix). If $a*v = 0$, then either $a=0$ or $v=0$

(x). For $u,v \in V$, $\exists!\, w \in V$ s.t. $u+w = v$.

(xi). $(a\oplus b)*(u+v) = a*u + a*v + b*u + b*v$, $a,b \in \mathbb{F}$, $u,v \in V$

Q3. Let X be an arbitrary set, with power set $P(X)$. Then the boolean algebra $(P(X), \Delta)$ is an $\mathbb{F}_2$-vector space.

We will explain every terminology here.

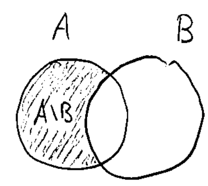
(i) $P(X)$ is a <u>boolean algebra</u> because it is closed under finite union $\cup$, finite intersection $\cap$ and complement $\setminus$, i.e.,
For $A, B \in P(X)$, we have
- $A \cup B \in P(X)$.
- $A \cap B \in P(X)$
- $X \setminus A \in P(X)$.

(ii). $\mathbb{F}_2$ is a <u>field</u>.
- As set, $\mathbb{F}_2 = \{0, 1\}$.
- addition table and multiplication table.

$\oplus$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

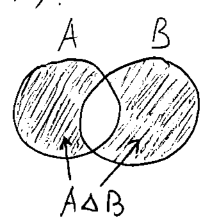
One may verify this is indeed a field.



(iii). Define the <u>symmetric difference</u> $\Delta : P(X) \times P(X) \longrightarrow P(X)$

$$(A, B) \longmapsto (A \setminus B) \cup (B \setminus A).$$

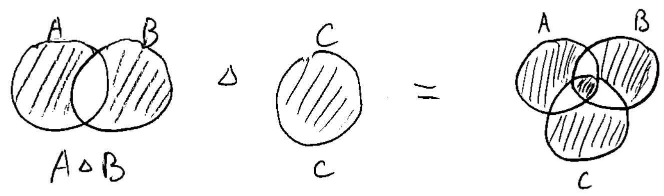
where $A \setminus B$ is defined to be $A \cap (X \setminus B)$ using (i).



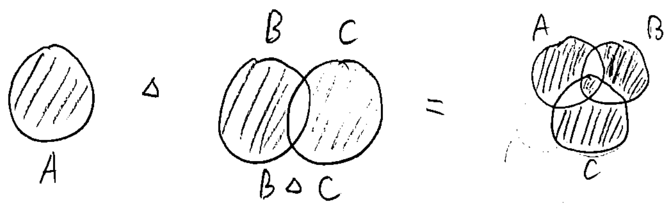
(iv). Check $(P(X), \Delta)$ is a vector space over $(\mathbb{F}_2, \oplus, \cdot)$:

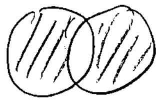
(VS1) $(A \Delta B) \Delta C = A \Delta (B \Delta C)$.

LHS:



RHS:



(VS2). $A \triangle B = B \triangle A$.  obvious

(VS3). $\exists \phi \in \mathcal{P}(X)$, $\phi \triangle A = A$. as $(\phi \setminus A) \cup (A \setminus \phi) = \phi \cup A = A$ identity is ϕ.

(VS4). $\forall A \in \mathcal{P}(X)$, $\exists A \in \mathcal{P}(X)$ s.t. $A \triangle A = \phi$. as $(A \setminus A) \cup (A \setminus A) = \phi \cup \phi = \phi$. inverse is itself.

Define $*: \mathbb{F}_2 \times \mathcal{P}(X) \to \mathcal{P}(X)$

$$(0, A) \mapsto \phi.$$
$$(1, A) \mapsto A.$$

(VS5). $a * (b * A) = (a \cdot b) * A$.

a	b	LHS	LHS.
0	0	$0 * (0 * A) = 0 * \phi = \phi$.	$(0 \cdot 0) * A = 0 * A = \phi$.
0	1	$0 * (1 * A) = 0 * A = \phi$	$(0 \cdot 1) * A = 0 * A = \phi$.
1	0	$1 * (0 * A) = 1 * \phi = \phi$	$(1 \cdot 0) * A = 0 * A = \phi$.
1	1	$1 * (1 * A) = 1 * A = A$.	$(1 \cdot 1) * A = 1 * A = A$.

(VS6). $1 * A = A$ directly from definition.

(VS7). $a * (A \triangle B) = (a * A) \triangle (a * B)$

$a = 0$. $LHS = \phi$. $RHS = \phi \triangle \phi = \phi$.

$a = 1$. $LHS = A \triangle B$. $RHS = A \triangle B$.

(VS8). $(a \oplus b) * A = (a * A) \triangle (b * A)$

a	b	LHS	RHS
0	0	$(0 \oplus 0) * A = \phi$	$(0 * A) \triangle (0 * A) = \phi \triangle \phi = \phi$
0	1	$(0 \oplus 1) * A = 1 * A = A$.	$(0 * A) \triangle (1 * A) = \phi \triangle A = A$.
1	0	$(1 \oplus 0) * A = (0 \oplus 1) * A = A$.	$(1 * A) \triangle (0 * A) = (0 * A) \triangle (1 * A) = A$. $\leftarrow$ alternative method using commutivity of $\oplus$ and $\triangle$.
1	1	$(1 \oplus 1) * A = 0 * A = \phi$.	$(1 * A) \triangle (1 * A) = A \triangle A = \phi$.

Q4. Recall that $(\mathbb{R}, +)$ is a vector space over itself $(\mathbb{R}, +, \cdot)$, where $* = \cdot$.
Show that it is not possible to extend it as a vector space over $(\mathbb{C}, +, \cdot)$.

pf. [(VS1)-(VS4) do not involve scalar multiplication, so it should be fine.
(VS6) is also fine as $1_{\mathbb{R}} = 1_{\mathbb{C}} = 1$.]

Suppose for $z \in \mathbb{C}$, $x \in \mathbb{R}$, we can define $*: \mathbb{C} \times \mathbb{R} \to \mathbb{R}$ s.t. $z * x \in \mathbb{R}$ and when $z \in \mathbb{R}$, $z * x = z \cdot x$. Then

$$f: \mathbb{C} \to \mathbb{R}$$
$$x \mapsto x * 1$$

is an injection, as $x * 1 = y * 1$ implies $(x - y) * 1 = 0$ so $x = y$. It is also surjective as $f|_{\mathbb{R}} = id_{\mathbb{R}}$.
But this is not possible because $f(i) = f(x)$ for some $x \in \mathbb{R}$ implies $i = x \in \mathbb{R}$ ($\Rightarrow\Leftarrow$). $\square$

Subspaces.

Def. 1.3. A vector <u>subspace</u> of a vector space $(V,+)$ over $\mathbb{F}$ is a subset $U \subseteq V$ that is also a vector space where $+$ and scalar multiple $*$ is restricted to U.

$$+|_U : U \times U \to U. \qquad *|_U : \mathbb{F} \times U \to U.$$

Remark. It will be clumsy to check (VS1)-(VS8) every time, as most of them are automatic.

First, we need $+|_U$ and $*|_U$ is well-defined, i.e.,

☆ (i). For $u, v \in U$, $u+v \in U$.

☆ (ii). For $a \in \mathbb{F}$, $u \in U$, $a*u \in U$.

Next, we need (VS1)-(VS8).

(VS1) automatic.

(VS2) automatic.

(VS3). Not automatic, so we need

☆ (iii). $\exists\, 0 \in U$ s.t. $0+u=u$ for all $u \in U$.

(VS4) follow from (ii) and $-u=(-1)*u$

(VS5)-(VS8) automatic.

Hence we only need to check (i)-(iii) in practice.

Span

(From now on we always assume V is over a field $\mathbb{F}$ without mentioning the field.)

Def. 1.4 (linear combination)

Let V be a vector space, and $v_i \in V$ for $i \in I$ ($|I|$ may be infinite). An <u>$\mathbb{F}$-linear combination</u> of $\{v_i\}_{i\in I}$ is

$$v := \sum_{i\in I} a_i * v_i \quad \text{for some } a_i \in \mathbb{F} \text{ and only } \underline{\underline{\text{finitely}}} \text{ many } a_i \neq 0$$

As the summation is finite, we have $v \in V$.

Def. 1.5. (span).

Let V be a vector space, and $S \subseteq V$ be a subset. We define the <u>span</u> of S, denoted $\langle S \rangle$ or span S to be all $\mathbb{F}$-linear combinations in S.

$$\operatorname{span} S := \left\{ \sum_{i\in I} a_i * s_i \;\middle|\; s_i \in S,\ \text{only finitely many } a_i \neq 0,\ a_i \in \mathbb{F} \right\}$$

Linear independence

Def 1.6. (linear independence).

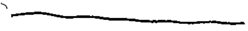
Let V be a vector space, and $S \subseteq V$. We say S is a <u>linearly independent</u> subset if for every $v_1, \dots, v_n \in V$

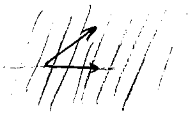
$$\sum_{i=1}^{n} a_i * v_i = 0 \text{ implies } a_1 = a_2 = \cdots = a_n = 0, \quad a_i \in \mathbb{F}.$$

i.e., every finite subset of S is linearly independent. We say S <u>linearly dependent</u> otherwise.

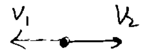
Remark: Geometric interpretation.

Let $V = \mathbb{R}^n$, $\mathbb{F} = \mathbb{R}$.

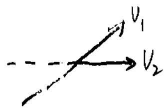
span $\{ \to \}$ =  $\mathbb{R}^1$.

span $\{$  $\}$ =  $\mathbb{R}^2$.

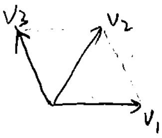
span $\{$  $\}$ = $\mathbb{R}^3$



are linearly dependent, as $v_1 + (-1)v_2 = 0$, but $(1,-1) \neq 0$.



are linearly independent, as $a_1 v_1 + a_2 v_2 \neq 0$ unless $a_1 = a_2 = 0$.



are linear dependent as $v_2 = v_1 + v_3 \Rightarrow v_1 + (-1)v_2 + v_3 = 0$ but $(1,-1,1) \neq 0$.

Q5. Recall $\mathbb{C}$ is a v.s. over $\mathbb{R}$ or $\mathbb{C}$.

(i) Is $\{1+i, 1-i\}$ linearly independent over $\mathbb{R}$? Yes

(ii) Is $\{1+i, 1-i\}$ linearly independent over $\mathbb{C}$? No.

Q6. Let V be a vector space. Show that a subset $W \subseteq V$ is a vector subspace of V iff span $W = W$.

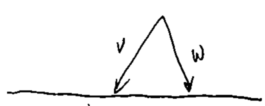
Pf. If W is a subspace of V, then every linear combination in W must stay in W. So span $W \subseteq W$. Obviously $W \subseteq$ span W.

Conversely, if $W =$ span W, then for $u, w \in W$, $u + w \in$ span $W = W$; for $a \in \mathbb{F}$, $w \in W$, $a * w \in$ span $W = W$; $0 = 0_{\mathbb{F}} * w \in$ span $W = W$. Hence W is a subspace. $\square$

Rmk. convex combination

$\{(1-t)v + tw \mid t \in [0,1]\}$.

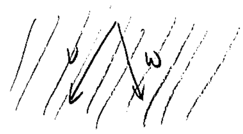
affine combination



$\{(1-t)v + tw \mid t \in \mathbb{R}\}$
$=$
$\{av + bw \mid a+b=1, a, b \in \mathbb{R}\}$

1 degree of freedom

linear combination



$\{av + bw \mid a, b \in \mathbb{R}\}$.

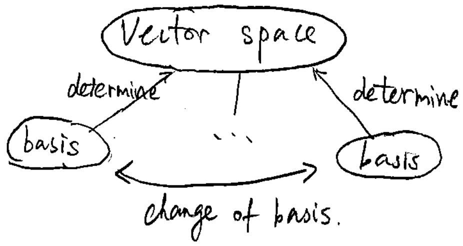
2-degree of freedom.

Tutorial 2 Basis and Dimension

(The "proofs" are sketchy. You are NOT supposed to write like this in exam!).

Linear independence, span, and basis

1. Vector spaces are very regularly behaved because of the notion of basis. We understand the whole vector space if we understand the basis. Everytime you have trouble proving something, use a basis!



Upshot 2.1. Vector space V.S. basis.

- Basis uniquely determines vector space.
- Vector space can have many basis, they are related by change of basis.

2. Basis are special to vector spaces! They have too many properties that is not at all common for other algebraic structures. Basis are well-behaved.

Prop. 2.2. Let V be a finite dimensional vector space.

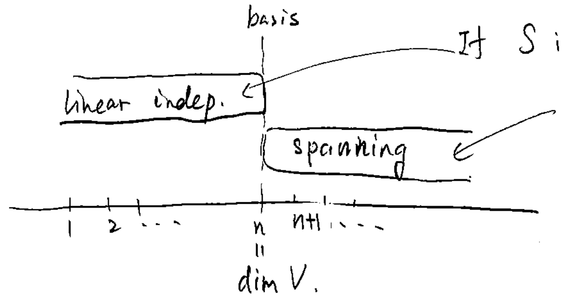
i). (existence). Every V has a basis.

ii). (extension). Every linearly indep. set can be extended to a basis.

iii). (reduction) Every spanning set can be reduced to a basis.

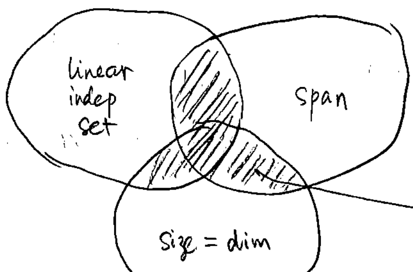
iv). The dimension, i.e., the size of basis, is well-defined for V. (All basis has same size).

Basis = linear independence + span = maximal linear independent set = minimal spanning set.



If S is lin. indep. set, and $T \subseteq S$, then T is lin. indep.

If S is spanning, and $S \subseteq T$, then T is spanning.



shaded area are basis

knowing 2 implies the remaining 1.

They are all consequences of the following central theorem.

Thm 2.3. (replacement theorem).

Let V be a finite dimensional vector space. If

$S_1 = \{v_1, \cdots, v_n\} \subseteq V$: linear independent set.

$S_2 = \{w_1, \cdots, w_m\} \subseteq V$: spanning set.

then $|S_1| \leq |S_2|$.

[The proof uses a replacement argument, explaining the name of thm.].

"pf." Initial step: As $v_1 \in V = \mathrm{span}\{w_1, \cdots, w_m\}$, we have $w_1 \in \mathrm{span}\{v_1, w_2, w_3, \cdots, w_m\} = V$.

(Replace w_1 by v_1). (Because $v_1 = \sum_{i=1}^{m} a_i w_i \Rightarrow w_1 = a_1^{-1}\left(\sum_{i=2}^{m} a_i w_i - v_1\right) \in \mathrm{span}\{v_1, w_2, w_3, \cdots, w_n\}$

$\Rightarrow \mathrm{span}\{v_1, w_2, w_3, \cdots, w_m\} = \mathrm{span}\{v_1, w_1, w_2, w_3, \cdots\} = V$).

Induction step. For $1 < i \leq \min\{n, m\}$, we are in situation

(Replace w_i by v_i)

$$\mathrm{span}\{\underbrace{v_1, \cdots, v_{i-1}}_{\text{linearly indep.}}, w_i, \cdots, w_n\} = V.$$

$v_i \in \mathrm{span}\{v_1, \cdots, v_{i-1}, w_i, \cdots, w_n\}$, and $v_i \notin \mathrm{span}\{v_1, \cdots, v_{i-1}\}$.

so $\exists\, w_j \in \{w_i, \cdots, w_n\}$, WLOG w_i, s.t. $w_i \in \mathrm{Span}\{v_1, \cdots, v_{i-1}, v_i, w_{i+1}, \cdots, w_m\}$.

If $m < n$, then we are in situation

$\mathrm{span}\{v_1, \cdots, v_m\} = V$, but $v_{m+1} \notin \mathrm{span}\{v_1, \cdots, v_m\} = V$ contradition. $\square$.

Q1. If $\{v_1, v_2, v_3\}$ is a basis, then $\{v_1+v_2, v_2+v_3, v_3\}$ is also a basis.

(The same proof work for any $M \in GL_3(\mathbb{F})$, $M\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ is also a basis).

pf. (lin. indep.) $0 = (a_1\ a_2\ a_3)\begin{pmatrix} v_1+v_2 \\ v_2+v_3 \\ v_3 \end{pmatrix}$

$$= (a_1\ a_2\ a_3)\left[\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}\right]$$

$$\overset{\text{why?}}{=} \left[(a_1\ a_2\ a_3)\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\right]\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$= (a_1\ \ a_1+a_2\ \ a_2+a_3)\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$\overset{\text{lin. indep}}{\underset{\text{of } v_1, v_2, v_3}{\Longrightarrow}} (0,0,0) = (a_1\ \ a_1+a_2\ \ a_2+a_3) = (a_1\ a_2\ a_3)\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

As $\det\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \neq 0$, $(a_1\ a_2\ a_3) = (0,0,0)$ by multiplying inverses on both sizes.

(span). One may check $v_1, v_2, v_3 \in \operatorname{Span}\{v_1+v_2, v_2+v_3, v_3\}$ respectively. But we will do it in a systemetic way.

$$
\begin{aligned}
&(a_1\ a_2\ a_3)\begin{pmatrix} v_1+v_2 \\ v_2+v_3 \\ v_3 \end{pmatrix} \\
&= (a_1\ a_2\ a_3)\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \Rightarrow \operatorname{span}\{v_1+v_2, v_2+v_3, v_3\} \subseteq \operatorname{span}\{v_1, v_2, v_3\}. \\
&= (a_1\ \ a_1+a_2\ \ a_2+a_3)\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}
\end{aligned}
$$

$$
\text{so}\quad (b_1\ b_2\ b_3)\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = (b_1\ b_2\ b_3)\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} v_1+v_2 \\ v_2+v_3 \\ v_3 \end{pmatrix} \Rightarrow \operatorname{span}\{v_1, v_2, v_3\} \subseteq \operatorname{span}\{v_1+v_2, v_2+v_3, v_3\}.
$$

$\square$.

Q2. Recall the inclusion-exclusion principle.

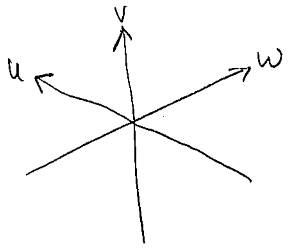
For finite dim. vector subspaces U, V of W.

$$\dim(U+V) = \dim U + \dim V - \dim(U \cap V).$$

This is always true.

Prove or give a counterexample: For f.d. vector subspaces U, V, W of W.

$$\dim(U+V+W) = \dim U + \dim V + \dim W - \dim(U\cap V) - \dim(U \cap W) - \dim(V \cap W) + \dim(U \cap V \cap W).$$



Where is the problem?

Reason: pairwise linear indep. $\not\Rightarrow$ joint linear indep.

You should also try to prove $n=2$ case using replacement thm and see how it fails for $n \geq 3$.

Tutorial 3: Linear maps, null space, range, rank-nullity theorem

(The "pf"'s are sketchy. You should NOT write like this in exam.)

<u>Linear maps</u> (Throughout, all v.s. are over a same field $\mathbb{F}$).

Def 3.1. (linear maps)

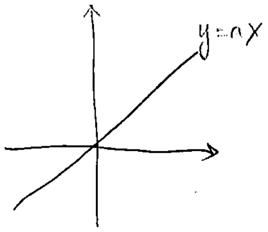
A (set theoretic) map $T: V \to W$ between two vector spaces V, W is <u>linear</u> if

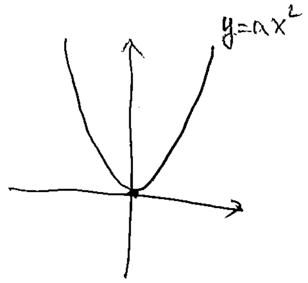
(i) $T(v_1+v_2) = Tv_1 + Tv_2$. the first $+_V$ is addition in V, the second is $+_W$ in W.

(ii). $T(av) = a(Tv)$ $\quad a \in \mathbb{F}$.

Remark. (i) In general, a map between algebraic (vector space) structures is a set map that prescrve the structure. (linear)

(ii). $x \longmapsto y = ax$ is linear $\qquad x \mapsto y = ax^2$ is not linear





Eg. Differentiation and integration are linear maps.

Take $\mathbb{F} = \mathbb{R}$. $V = C(\mathbb{R})$ = space of all continuous functions $[0,1] \to \mathbb{R}$.

"pf." (i) $(f_1+f_2)' = f_1' + f_2'$ $\quad$ so $'$ is linear from $C(\mathbb{R})$ to $C(\mathbb{R})$.

$(af_1)' = af_1'$

(ii). $\int_0^1 f_1+f_2 = \int_0^1 f_1 + \int_0^1 f_2$

$\int_0^1 af = a\int_0^1 f$ $\quad$ so $\int_0^1$ is linear from $C(\mathbb{R})$ to $C(\mathbb{R})$ $\square$.

Eg. Matrix transpose is linear map.

Over $\mathbb{F}$ arbitrary, take $V = M_n(\mathbb{F})$. $\quad {}^t : M_n(\mathbb{F}) \to M_n(\mathbb{F})$ is linear.

"pf". $(A+B)^t = A^t + B^t$.

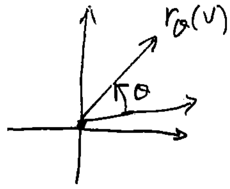
$(aA)^t = aA^t$. $\square$.

Eg. Linear map may not look linear: find a linear map $\mathbb{R} \to \mathbb{R}_{>0}$ over $\mathbb{R}$?

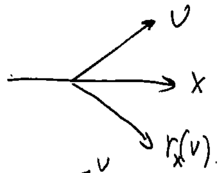
Cx: $\varphi_t: \mathbb{R} \to \mathbb{R}$, $a \mapsto a+t$ $\quad$ translation is NOT linear. Why?

Eg. Rotation, reflection, projection are all linear.

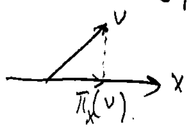
Take $\mathbb{F} = \mathbb{R}$, $V = \mathbb{R}^2$.



<u>Rotation by θ</u>: $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \xmapsto{r_\theta} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.



<u>Reflection about x-axis</u>: $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \xmapsto{r_x} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$



<u>Projection onto x-axis</u>: $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \xmapsto{\pi_x} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

"pf." This is a special case of the below example. $\square$

Eg. Left-multiplication by a matrix is linear.

Over any $\mathbb{F}$, $V = \mathbb{F}^n$. Let $A \in M_n(\mathbb{F})$. Define the left-multiplication map

$$\begin{aligned} \mathcal{L}_A : \mathbb{F}^n &\to \mathbb{F}^n \\ v &\mapsto Av \end{aligned}$$

This is a linear map.

"pf" $A(v_1 + v_2) = Av_1 + Av_2$.

$A(a v_1) = a(Av_1)$ $\square$

Remark. <u>Every</u> linear map between finite dimensional v.s. is actually of this form, called matrix representation. We will study that later.

<u>Linear maps and basis</u>.

$$\begin{tikzcd} S \arrow[r, "f"] \arrow[d, hook] & W \\ (\operatorname{Span} S =) V \arrow[ur, "\varphi(f)"'] & \end{tikzcd}$$

$\boxed{\text{Upshot}}$: Every linear map is determined by its value on a basis!

This makes our life easier. To be precise, we state this in the

Proposition 3.2. Let V be a vector space with basis $S = \{e_i\}_{i \in I} \subseteq V$. Let W be another v.s. Then

$$\underset{\substack{\uparrow \\ \text{linear maps}}}{\operatorname{Hom}_{\mathbb{F}}(V, W)} \underset{\substack{\uparrow \\ \text{set bijection}}}{\cong} \underset{\substack{\uparrow \\ \text{set maps/functions}}}{\operatorname{Map}_{\text{Set}}(S, W)}.$$

"pf". Define $\varphi : \operatorname{Map}_{\text{Set}}(S, W) \to \operatorname{Hom}_{\mathbb{F}}(V, W)$ called <u>extend by linearity</u>.

$$f \longmapsto \varphi(f)$$

where $\varphi(f)(v) = \varphi(f)\left(\sum_{\text{finite}} a_i e_i\right) = \sum_{\text{finite}} a_i f(e_i)$, $v = \sum_{\text{finite}} a_i e_i$

This is well-defined as $\{a_i, i\in I\}$ are unique by property of basis.

For injectivity, $\varphi(f_1)=\varphi(f_2) \Rightarrow \varphi(f_1)\big|_S = \varphi(f_2)\big|_S \Rightarrow f_1=f_2$.

For surjectivity, take $g\in \mathrm{Hom}_{\mathbb{F}}(V,W)$, define $f(e_i):=g(e_i)$, $e_i\in S$, so $f\in \mathrm{Map}_{Set}(S,W)$.

Then $g=\varphi(f)$. $\square$

Remark. Whenever you want to do something for linear maps, it suffices to do it on a basis, and extend by linearity.

Eg. $\mathrm{Hom}_{\mathbb{F}}(V,W)=\{$all linear maps between two v.s. $V, W\}$ is again a vector space.

"pf." $(f_1+f_2)(v):=f_1(v)+f_2(v)$.

$(f_1+f_2)(av):= f_1(av)+f_2(av) = a(f_1(v)+f_2(v))$.

$\Rightarrow f_1+f_2\in \mathrm{Hom}_{\mathbb{F}}(V,W)$.

$(af_1)(v) = a\cdot f(v)$.

$(af)(bv) = a\cdot f(bv) = ab f(v)$

$\Rightarrow af_1 \in \mathrm{Hom}_{\mathbb{F}}(V,W)$.

$\left(\begin{array}{l} 0(v):=0_W \text{ is also linear} \\ \Rightarrow 0\in \mathrm{Hom}_{\mathbb{F}}(V,W). \end{array}\right)$ Checking axioms are omitted. $\square$

Eg. Matrix multiplication: composition of linear maps are linear. $U\xrightarrow{T} V \xrightarrow{S} W$

For $T\in \mathrm{Hom}_{\mathbb{F}}(U,V)$, $S\in \mathrm{Hom}_{\mathbb{F}}(V,W)$ we have $S\circ T\in \mathrm{Hom}_{\mathbb{F}}(U,W)$.

As compositions are not commutative, even for $U=V=W$, $T\circ S\neq S\circ T$ in general.

$\boxed{\text{Upshot}}$: This is the reason why matrix multiplication is not commutative.

<u>Rank-nullity theorem</u> (/First isomorphism theorem if you take 2070)

basis for V:	$\underbrace{v_1\ v_2}_{\ker T}$	$v_3, \dots, v_n$.
basis for $V/_W$:		$v_3+\ker T, \dots, v_n+\ker T$.
basis for range$T$:		$T(v_3), \dots, T(v_n)$.

* Thm 3.3: For a linear map $T: V\to W$. We have

not examinable

(i) $V=\ker T \oplus V/_{\ker T}$ ← quotient space.

(ii) $V/_{\ker T} \cong \mathrm{range}(T)$. (isomorphism = linear bijection)

As a result, $\dim V = \dim\ker T + \dim \mathrm{range}\, T$ if $\dim V<\infty$.

For this to make sense, we need to define <u>quotient space</u>.

This page is Not examinable

Def. 3.4 (quotient vector space).

For a v.s. V and a subspace $W \subseteq V$. We define the <u>quotient vector space</u> by

$$V/W := \{ v+W \mid v \in V \}, \text{ where } v+W := \{ v+w \mid w \in W \} \text{ called } \underline{\text{cosets}}.$$

[Note: $v+W = w+W$ iff $v-w \in W$]

addition: $(v_1+W)+(v_2+W) := (v_1+v_2)+W$.

Scalar multiplication: $a(v+W) := av+W$

You may verify this is also a vector space.

Problem: $\oplus$ does not make sense unless we realize V/W as a subspace of V.

Proposition 3.5. Any basis $\underset{\subseteq V}{\mathcal{E}_1}$ of W can be extended to $\mathcal{E}_1 \cup \mathcal{E}_2$ of V. We have

$$\overline{\mathcal{E}_2} = \{ v+W \mid v \in \mathcal{E}_2 \}$$

is a basis of V/W.

pf. Exercise. $\square$

Proposition 3.6. If $\mathcal{E}_1$ is a basis of W, and $\overset{\subseteq V/W}{\overline{\mathcal{E}_2}} = \{ v+W \mid v \in \overset{\subseteq V}{\mathcal{E}_2} \}$ is a basis of V/W.

Then $\mathcal{E}_1 \cup \mathcal{E}_2$ is a basis of V.

pf. Exercise. $\square$

Cor 3.7. We can find a vector subspace $W' = \operatorname{span} \mathcal{E}_2$ of V with

(i) $W' \cong V/W \leftarrow \operatorname{span} \overline{\mathcal{E}_2}$.

(ii) $V = \underset{\operatorname{span}\mathcal{E}_1}{W} \oplus \underset{\operatorname{span}\mathcal{E}_2}{W'}$.

by the above extension of basis technique. (So V/W can be viewed as a subspace of V if we do not distinguish W' and V/W as above.)

This finishes the proof of (i) in Thm 3.3.

For (ii), define a map,

$$V/\ker T \longrightarrow \operatorname{range}(T)$$

$$v+\ker T \longmapsto T(v).$$

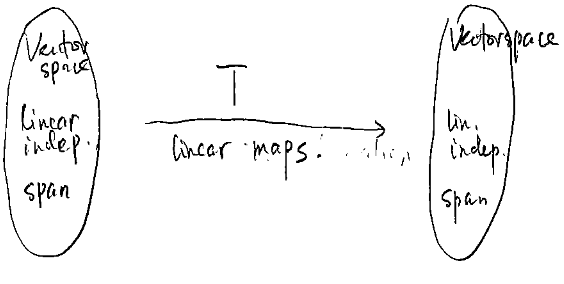
This is well-defined as if $v-w \in \ker T$, $T(v) = T(v-w+w) = \underset{0}{T(v-w)} + T(w) = T(w)$.

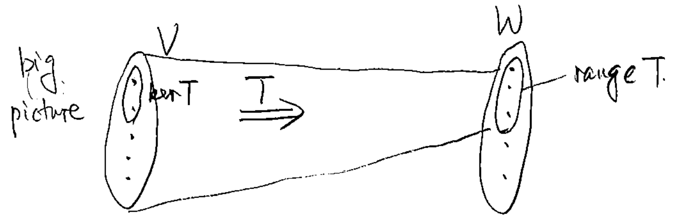
This is injective as $T(v) = T(w) \Rightarrow T(v-w) = 0 \Rightarrow v-w \in \ker T \Rightarrow v-w+\ker T = 0+\ker T$ in $V/\ker T$.

This is surjective by definition of range. $\square$ 3.3.

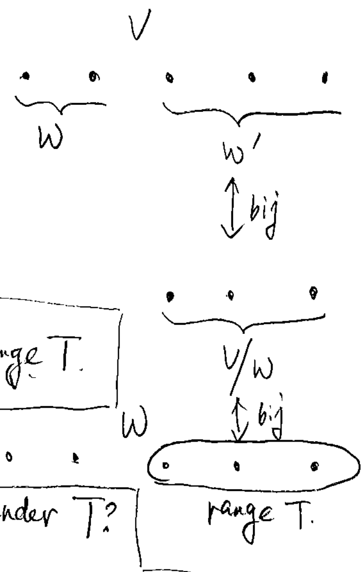
This page in inside syllabus.

Linear maps & lin. indep / span





dots means basis.



Note that we get no control on what is outside range T.
We may WLOG replace W by range T.

Q: How do linear indep. sets / spanning sets behave under T?

Ans: Summarized in following

Prop 3.8. Let $V \xrightarrow{T} W$ be the linear maps between two vector spaces V, W.
Let $S \subseteq V$ be a subset.

(i) If T injective, then $T(S)$ lin. indep $\Rightarrow$ S linearly independent in V. (actually $\overset{HW2}{\Leftrightarrow}$).

(ii) If T surjective, then S spanning set $\Rightarrow$ $T(S)$ is spanning set of W.

pf. (i). Let $S' \subseteq S$ by any finite subset. Then $T(S') \subseteq T(S)$ is a finite subset of a linearly independent set, hence linearly independent. Now for $s_i \in S'$, $a_i \in \mathbb{F}$,

$$\sum_{i=1}^n a_i s_i = 0 \Rightarrow T\left(\sum_{i=1}^n a_i s_i\right) = T(0) = 0. \Rightarrow \sum_{i=1}^n a_i T(s_i) = 0 \xRightarrow[\text{As } T \text{ is inj, } T(s_i) \text{ are distinct in } T(S')]{T(S') \text{ lin. indep.}} a_i = 0 \text{ for all } i = 1, \dots, n.$$

This is true for every finite subsets S'. By def, S is lin. indep.

(ii). For any $w \in W$, $\exists v \in V$ s.t. $T(v) = w$ as T surjective. As span $S = V$ we have $v = \sum_{i=1}^n a_i s_i$ for some $a_i \in \mathbb{F}$, $s_i \in S$. But then

$$w = T(v) = T\left(\sum_{i=1}^n a_i s_i\right) = \sum_{i=1}^n a_i T(s_i) \in \operatorname{span} T(S).$$

□

Remark. (i) This is actually "abstract non-sense". Trivially true but very useful for writting proofs.

(ii). Notice the "duality" between lin. indep. V.S. span again!

The first "duality" is $\boxed{\text{lin indep}}$ $\boxed{\text{span}}$ For $S' \subseteq S$, S lin. indep $\Rightarrow$ S' lin. indep. S' spanning $\Rightarrow$ S spanning.

(iii) If T is bijective, both are iff (HW2). (c.f. Q1 of Tutorial 2).

This page is inside syllabus

Pf of prop 3.5-3.6. Let $W \subseteq V$ with $\mathcal{E}_1$ basis of W, extend to a basis $\mathcal{E}_1 \cup \mathcal{E}_2$ of V. Let $W' = \operatorname{span} \mathcal{E}_2$. Obviously $V = W \oplus W'$. We wish to show $W' \cong V/W$ by basis maps $\mathcal{E}_2 \mapsto \bar{\mathcal{E}}_2$.

Consider the natural quotient map $V \xrightarrow{T} V/W$, $v \longmapsto \bar{v} = v + W$. T is obviously linear and surjective.

By prop 3.8, one direction for spanning set is automatic.

It remains to show

(i). $\mathcal{E}_2$ is lin. indep. $\Leftrightarrow T(\mathcal{E}_2) = \bar{\mathcal{E}}_2$ is linearly indep.

(ii). $T(\mathcal{E}_2) = \bar{\mathcal{E}}_2$ spans $V/W \Rightarrow \mathcal{E}_2$ spans W'.

Indeed, for (i), ($\Rightarrow$): $\sum a_i \bar{v}_i \overset{(v_i \in \mathcal{E}_2)}{=} \bar{0} \Rightarrow \sum a_i(v_i + W) = \bar{0} \Rightarrow (\sum a_i v_i) + W = 0 + W \Rightarrow \sum a_i v_i \in W \Rightarrow \sum a_i v_i = \sum b_j w_j$, $w_j \in \mathcal{E}_1$.
As $\mathcal{E}_1 \cup \mathcal{E}_2$ are lin. indep. $\Rightarrow a_i = b_j = 0 \ \forall i, j$.

($\Leftarrow$). $\sum a_i v_i \overset{(v_i \in \mathcal{E}_2)}{=} 0 \Rightarrow \sum a_i(v_i + W) = 0 + W \Rightarrow \sum a_i \bar{v}_i = 0$, and $\bar{v}_i \neq \bar{v}_j$ as $v_i - v_j \notin W$ ($v_i \notin \operatorname{span}\{\mathcal{E}_1 \cup \{v_j\}\}$). As $\bar{v}_i \in \bar{\mathcal{E}}_2$ distinct, lin. indep, $a_i = 0 \ \forall i$.

for (ii), for $w' \in W' = \operatorname{span} \mathcal{E}_2$. As $\bar{\mathcal{E}}_2$ spans V/W.
$\bar{w'} = \sum a_i \bar{v_i}$ for some $v_i \in \mathcal{E}_2$, $a_i \in \mathbb{F}$. $\Leftrightarrow w' - \sum a_i v_i \in W$. But $w' - \sum a_i v_i \in W'$ and $W \cap W' = \{0\}$.

$\Rightarrow w' - \sum a_i v_i = 0$. $\Rightarrow W' \subseteq \operatorname{span} \mathcal{E}_2$. $\square$.

Remark. (i) You may see prop 3.8 simplifies half of the proof!

(ii) The "abstract nonsense" like prop 3.8 and (ii) of 3.3 are studied systematically in a subject called category theory.

(iii) This concludes the proof of rank-nullity.

(iv). By prop 3.8 (i), range(T) measures "How many linearly independent variables T preserved."
ker(T) measures "How much T fails to measure linear independence."
For vector spaces, dim ker T + dim range T = total size of linearly independent set in V = dim V.

Warning: If V is not finite dimensional, we need to assume (AoC) (Axiom of choice.) in order to show every linearly independent set can be extended to a basis.

This page is non-examinable.

Def. Partially ordered set (poset) is a set P with a order $\leq$ s.t.

(i) $a \leq a$

(ii) $a \leq b$, $b \leq a \Rightarrow a = b$.

(iii). $a \leq b$, $b \leq c$, $c \leq a$

$\leq$ is called totally ordered if any two elements are comparible, i.e., either $a \leq b$ or $b \leq a$ for $\forall a, b$.

Axiom: (Zorn's lemma) Let P be a poset.

If Any chain/totally ordered set P has an upper bound in P, then the set P has a maximal element.

Remark. This is equivalent to Axiom of choice (AoC).

Prop. Every lin. indep. can be extended to a basis of V. (Every vector space has a basis.)

pf. let $S \subseteq V$ be lin. indep. Consider

$$\mathcal{A} = \{ T \subseteq V : S \subseteq T,\ T \text{ is lin. indep} \}.$$

Every chain in $\mathcal{A}$ has an upper bound. (take union).

Then by Zorn's lemma, $\mathcal{A}$ has a maximal element R.

We claim $\langle R \rangle = V$.

Suppose $\langle R \rangle \subsetneq V$, then $\exists v \in V$ $v \notin \langle R \rangle$.

Then $R \subsetneq R \cup \{v\} \in \mathcal{A}$. contradiction of maximality of R.

So we must have $\langle R \rangle = V$, and R is a basis.

This proves every vector space has a basis.

HW1 Feedback

$W \leq V$.

Q1.3.31 (a). If $v+W$ is a subspace of V, then $v \in W$.

Attempt: As $v = v+0 \in v+W$. and w is a subspace.

$-v \in v+W$.

so $-2v \in W$.

As W is a subspace.

$(-\frac{1}{2}) \cdot (-2v) = v \in W$.

This is false proof if $\mathbb{F} = \mathbb{F}_2 = \{0,1\}$. Here,

$2=0$. So no notion of $-\frac{1}{2}$ in $\mathbb{F}_2$!

(b). What would you do to prove $A=B$ two sets equal?

$A \subseteq B$: $\forall a \in A, \exists b \in B$ s.t. $a=b$.

$B \subseteq A$: $\forall b \in B, \exists a \in A$ s.t. $a=b$.

Q1.3.11. $W = \{ f \in P(\mathbb{F}) \mid f(x)=0$ or $f(x)$ has degree $n \} \leq P(\mathbb{F})$ for $n \geq 1$?

What is the degree of a polynomial?

Eg. $\underbrace{7x^2y^3}_{\text{degree } 5} + \underbrace{x^5}_{\text{deg } 5} + \underbrace{7}_{\text{deg } 0}$ is of degree 5

If $f = \sum_{i=1}^{n} f_i$ where f_i are monomials.

$\deg f = \max\{\deg f_i\}$ convention: degree of constant function is 0.

where degree of a monomial is the sum of powers of all indeterminates.

$$f = f_1 + f_2 + f_3$$

$$f_1 = 7x^2y^3, \quad f_2 = x^5, \quad f_3 = 7.$$

$$\deg f_1 = 2+3=5 \quad \deg f_2 = 5. \quad \deg f_3 = 0.$$

$\deg f = \max\{5,5,0\} = 5$.

Q1.3.26. What is the difference between $+$ and $\oplus$?

Def. Let $U_1, \cdots, U_m \subseteq V$ subspaces.

The <u>direct sum</u> $U_1 \oplus \cdots \oplus U_m$ is the <u>sum</u> $U_1 + \cdots + U_m$ AND

$u_1 + \cdots + u_m = 0 \quad u_i \in U_i \Rightarrow u_i = 0$ for all $i=1, \cdots, n$.

Prop. Let $U_1, U_2 \leq V$ subspaces. Then $U_1 + U_2$ is a direct sum iff $U_1 \cap U_2 = \{0\}$.

Warning: This is NOT true for $n > 2$, i.e., it will NOT be sufficient to check.

$U_i \cap U_j = \{0\}$ for all pair of (i, j) with $i \neq j$.

Ex: $\mathbb{F}^3$ as v.s over $\mathbb{F}$.

$U_1 := \{(x, y, 0) \in \mathbb{F}^3 : x, y \in \mathbb{F}\}$.

$U_2 := \{(0, 0, z) \in \mathbb{F}^3, z \in \mathbb{F}\}$.

$U_3 := \{(0, y, y) \in \mathbb{F}^3 : y \in \mathbb{F}\}$.

Note that $U_1 \cap U_2 = \{0\}$, $U_2 \cap U_3 = \{0\}$, and $U_1 \cap U_3 = \{0\}$.

But $(0, 0, 0) = \underset{\underset{U_1}{\cap}}{(0, 1, 0)} + \underset{\underset{U_2}{\cap}}{(0, 0, 1)} + \underset{\underset{U_3}{\cap}}{(0, -1, -1)}$.

Direct sum and direct product

① If $U_1, U_2 \leq V$ subspace, with $U_1 \cap U_2 = \{0\}$, then $U_1 \oplus U_2$ is the direct sum.

- $U_1 \oplus U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$ as sets.
- $+ : (U_1 \oplus U_2) \times (U_1 \oplus U_2) \to U_1 \oplus U_2$

  $(a_1 + b_1, a_2 + b_2) \mapsto \underbrace{(a_1 + a_2)}_{U_1} + \underbrace{(b_1 + b_2)}_{U_2}$.
- $* : \mathbb{F} \times (U_1 \oplus U_2) \longmapsto U_1 \oplus U_2$

  $a * (a_1 + b_1) \longmapsto \underbrace{(a a_1)}_{U_1} + \underbrace{(a b_1)}_{U_2}$.

② For any two vector spaces V_1, V_2, not necessarily subspace of another vector space, we can form a direct product / cartesian product $V_1 \times V_2$:

- $V_1 \times V_2 = \{(a, b) \mid a \in V_1, b \in V_2\}$ as sets.
- $+ : (V_1 \times V_2) \times (V_1 \times V_2) \mapsto V_1 \times V_2$

  $((a_1, b_1), (a_2, b_2)) \mapsto (\underbrace{(a_1 + b_1)}_{V_1}, \underbrace{(a_2 + b_2)}_{V_2})$.
- $* : \mathbb{F} \times (V_1 \times V_2) \mapsto V_1 \times V_2$

  $a * (a_1, b_1) \mapsto (\underbrace{a a_1}_{V_1}, \underbrace{a b_1}_{V_2})$

③. As $U_1, U_2 \leq V$ with $U_1 \cap U_2 = \{0\}$, then $U_1 \oplus U_2 \cong U_1 \times U_2$

Pf. We define a linear map $T: U_1 \oplus U_2 \longrightarrow U_1 \times U_2$ by

$$\sum_{i=1}^{n} (a_i + b_i) \longmapsto \left(\sum_{i=1}^{n} a_i , \sum_{i=1}^{n} b_i \right).$$
$$a_i \in U_1,\ b_i \in U_2$$

This is well-defined as $\sum_{i=1}^{n}(a_i+b_i) = \sum_{j=1}^{m}(a_j'+b_j') \Rightarrow \underbrace{\sum_{i=1}^{n} a_i - \sum_{j=1}^{m} a_j'}_{\in U_1} = \underbrace{\sum_{j=1}^{m} b_j' - \sum_{i=1}^{n} b_i}_{\in U_2} \subseteq U_1 \cap U_2 = \{0\}$.

Hence $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} a_j'$ and $\sum_{j=1}^{m} b_j' = \sum_{i=1}^{n} b_i$.

T is injective as $\left(\sum_{i=1}^{n} a_i, \sum_{i=1}^{n} b_i\right) = (0,0) \Rightarrow \sum_{i=1}^{n} a_i = 0$ and $\sum_{i=1}^{n} b_i = 0. \Rightarrow \sum_{i=1}^{n}(a_i+b_i) = 0$.

T is surjective by definition. $\square$

Remark. We also say $\oplus$ we defined in ① is internal direct sum as U_1 and U_2 are subspaces of another vector space.

$\times$ we defined in ② is external direct sum as there is no such restriction.

Because of ③ we see internal direct sum is just a special case of direct product. In some literature, $\times$ is also denoted as $\oplus$.

Fact of life: $\{0\}$ is not a linearly independent set as $a \cdot 0 = 0 \not\Rightarrow a = 0$.

Direct sum, sum and basis. Throughout, let V be a vector space and $S \subseteq V$ be a linearly independent set.

Fact: Suppose $S = S_1 \amalg S_2$. Then $\underset{\text{span}}{\langle S \rangle} = \langle S_1 \rangle \oplus \langle S_2 \rangle$, with S, S_1, S_2 being basis of respective vector space.

$$\underbrace{\underbrace{\cdot\ \cdot\ \cdot\ \cdot}_{S_1}\ \ \underbrace{\cdot\ \cdot}_{S_2}}_{S}$$

Pf. We use prop. at the top of P3-9.

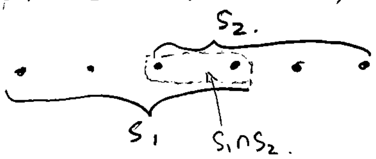
$\langle S_1 \rangle + \langle S_2 \rangle \subseteq \langle S \rangle$ is obvious, as $\langle S_1 \rangle + \langle S_2 \rangle$ is the smallest vector space containing $\langle S_1 \rangle$ and $\langle S_2 \rangle$.

$\langle S \rangle \subseteq \langle S_1 \rangle + \langle S_2 \rangle$ is also clear, as elements of $\langle S \rangle$ are of the form $\sum a_i s_i$ $s_i \in S$.

But $S = S_1 \amalg S_2$, $\sum a_i s_i = \sum a_i s_i' + \sum a_i s_i''$ where $s_i' \in S_1$, $s_i'' \in S_2$.

As the union is disjoint, $\langle S_1 \rangle \cap \langle S_2 \rangle = 0$, otherwise $\sum a_i s_i = \sum b_j s_j'$ $s_i \in S_1, s_j' \in S_2$ gives $\sum a_i s_i - \sum b_j s_j' = 0$. a non-trivial linear combination of elements in S. $\underset{S \text{ is lin. indep.}}{\times}$ $\square$.

Fact. $\langle S_1 \cup S_2 \rangle = \langle S_1 \rangle + \langle S_2 \rangle$; $\langle S_1 \cap S_2 \rangle = \langle S_1 \rangle \cap \langle S_2 \rangle$. (Rmb we assume S is lin. indep).



pf. Exercise.

Remark. (The set of all subspaces, $\cap$, $+$) form a bounded modular lattice (see page 13).

The "zero element" in a vector space may not be 0.

Ex: $(P(X), \Delta)$ is a v.s. over $\mathbb{F}_2$. Its zero element is ϕ.

Ex: $(\mathbb{R}_{>0}, \cdot)$ is a v.s. over $\mathbb{R}$.

$$+ : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0}$$
$$(x, y) \longmapsto xy$$

$$* : \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0}$$
$$(z, x) \longmapsto x^z = e^{z \ln x}.$$

Its zero element is 1.

If V is infinite dimensional, $\dim V$ is not defined. $\infty - \infty$ is not defined!

$$\underset{\uparrow \text{a set.}}{M_n(\mathbb{F})} \neq \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \nwarrow \text{an element.}$$

$$M_n(\mathbb{F}) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} : a_{ij} \in \mathbb{F} \right\}.$$

Q: What is a basis for $M_n(\mathbb{F})$?

A: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longleftrightarrow \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$.

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1+a_2 & b_1+b_2 \\ c_1+c_2 & d_1+d_2 \end{pmatrix}. \qquad \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix} = \begin{pmatrix} a_1+a_2 \\ b_1+b_2 \\ c_1+c_2 \\ d_1+d_2 \end{pmatrix}.$$

$$r\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ra & rb \\ rc & rd \end{pmatrix} \qquad r\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} ra \\ rb \\ rc \\ rd \end{pmatrix},$$

So we have $M_n(\mathbb{F}) \cong \mathbb{F}^{n^2}$ as vector spaces.

$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ is a basis for $\mathbb{F}^{n^2}$, so $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & 0 \\ 0 & \cdots & & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & & 0 \\ 0 & \cdots & & 0 \end{pmatrix} \cdots \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & & & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$

is a basis for $M_n(\mathbb{F})$.

Q. $M_n(\mathbb{F}) \overset{\varphi}{\cong} \mathbb{F}^{n^2}$, then why do we care about $M_n(\mathbb{F})$? $\mathbb{F}^{n^2}$ seems more convenient?

A: We see addition is not dependent on positions of entries, as they are defined entriwise. But multiplication does!

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2+b_1c_2 & a_1b_2+b_1d_2 \\ c_1a_2+d_1c_2 & c_1b_2+d_1d_2 \end{pmatrix} \xrightarrow{\varphi} \begin{pmatrix} a_1a_2+b_1c_2 \\ a_1b_2+b_1d_2 \\ c_1a_2+d_1c_2 \\ c_1b_2+d_1d_2 \end{pmatrix}$$

But this is not our usual definition for $\mathbb{F}^{n^2}$:

$$\begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix} \times \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 \\ b_1b_2 \\ c_1c_2 \\ d_1d_2 \end{pmatrix}$$

different!

Also, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix}$ can be defined, but $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix}$ is not usually defined!

So the correct way to view a matrix is not through vector spaces, but <u>Linear maps</u>! In particular, $M_n(\mathbb{F})$ represent $\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^n)$.

That is why we have <u>matrix multiplication</u> and <u>matrix-vector multiplication</u>.

$$\updownarrow \qquad\qquad\qquad\qquad \updownarrow$$

$$T_1 \circ T_2 : V_1 \xrightarrow{T_2} V_2 \xrightarrow{T_1} V_3 \qquad\qquad V_1 \xrightarrow{T} V_2$$

$$T(v_1) \in V_2, \quad v_1 \in V_1.$$

Note that the collection of linear maps is also a vector space; indeed,

$$M_n(\mathbb{F}) \cong \mathbb{F}^{n^2} \cong \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^n)$$

as vector spaces over $\mathbb{F}$.

But unlike $\mathbb{F}^{n^2}$, $M_n(\mathbb{F}) \cong \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^n)$ as rings (if you took 2070).

$\uparrow$ not a ring.

$\boxed{\text{Homework}}$ Prove or give a counterexample:

For a linear map $T: V \to W$, $S \subseteq V$ a subset.

$T(S)$ is a linearly independant set $\Rightarrow$ S is linearly independent set.

out of syllabus

<u>Lattice of subspace.</u> Let V be a v.s. $/\mathbb{F}$.

Let $\mathcal{S} = \{\text{subspaces of } V\}$.

Then $(\mathcal{S}, +, \cap)$ form a bounded modular lattice.

Def. (lattice) A partially ordered set $(\mathcal{L}, \leq)$ is a <u>lattice</u> if

(i). $\forall a, b \in \mathcal{L}$, $\exists c \in \mathcal{L}$ that is • $a \leq c$, $b \leq c$ (upper bound).

• $\forall d \in \mathcal{L}$ with $a \leq d$, $b \leq d$, we have $c \leq d$ (least).

denote such c as <u>$a \vee b$</u> called <u>join</u> or <u>least upper bound</u>. (Note that c is automatically unique)

(ii) $\forall a, b \in \mathcal{L}$, $\exists c' \in \mathcal{L}$ that is • $c' \leq a$, $c' \leq b$ (lower bound)

• $\forall d' \in \mathcal{L}$ with $d' \leq a$ and $d' \leq b$, we have $d' \leq c'$ (greatest).

denote such c' as <u>$a \wedge b$</u> called <u>meet</u> or <u>greatest lower bound</u>. (c' is also unique).

Eg.

$$\begin{tikzcd}[column sep=small] & W_1 + W_2 & \\ W_1 \arrow[ur, dash] & & W_2 \arrow[ul, dash] \\ & W_1 \cap W_2 \arrow[ul, dash] \arrow[ur, dash] & \end{tikzcd} \qquad \begin{tikzcd}[column sep=small] & \mathrm{lcm}(c,d) & \\ c \arrow[ur, dash] & & d \arrow[ul, dash] \\ & \gcd(c,d) \arrow[ul, dash] \arrow[ur, dash] & \end{tikzcd}$$

$(\mathcal{S}, \subseteq)$. $\qquad (\mathbb{Z}_+, \mid)$

divisible.

Def. (bounded, modular).

A lattice $(\mathcal{L}, \leq, \wedge, \vee)$ is

(i) <u>bounded</u> if it has greatest element and least element. (they are essentially unique).

(ii) <u>modular</u> if $\forall a, b, x \in \mathcal{L}$, $a \leq b \Rightarrow a \vee (x \wedge b) = (a \vee x) \wedge b$.

Eg. $(\mathcal{S}, \leq)$ is bounded as V is the greatest element and $\{0\}$ is the least element.

modular as $W_1 + (W_2 \cap W_3) = (W_1 + W_2) \cap W_3$ if $W_1 \leq W_3$. (Ex).

$(\mathbb{Z}_+, \mid)$ is not bounded as it has no greatest element.

is not modular as $\mathrm{lcm}(3, \gcd(2,4)) = 6 \neq 2 = \gcd(\mathrm{lcm}(3,2), 4)$

Ans to Homework in Tutorial 3.

Q. Prove or give a counterexample.

Let $T: V \to W$ be a linear map. $S \subseteq V$ is a subset.

$T(S)$ is linearly independent set $\Rightarrow$ S is linearly independent set.

Ans. False.

A possible false proof.

Take $\sum_i a_i s_i = 0$, $s_i \in S$. Applying T gives $0 = T(\sum_i a_i s_i) = \sum_i a_i \underbrace{T(s_i)}_{\text{in } T(S)}$

As $T(S)$ is linearly independent, we have $a_i = 0$ $\forall i$.

The proof is false as there may be $T(s_i) = T(s_j)$ for $s_i \neq s_j$. This term then becomes $0 = (a_i + a_j)\underbrace{T(s_i)}_{T(s_i)=T(s_j)} +$ other terms. but we can only conclude $a_i + a_j = 0$.

not $a_j = a_i = 0$. To fix this we need T is injective on S.

Cx. Consider $T: \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x$. Let $S = \{(1, x) : x \in \mathbb{R}\}$.

$T(S) = \{1\}$ is a linearly independent set, but S not as

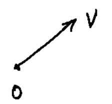
$$(1,4) - (1,3) + (1,1) - (1,2) = 0.$$ is a non-trivial linear combination of S. (Actually as $\dim_{\mathbb{R}} \mathbb{R}^2 = 2$, at most 2 elements in S can be linearly independent.)

§4 Matrix Representation of Linear Maps. Change of Basis

This section reduces 2040 to 1030 by a process called "choosing a basis".

<u>Coordinate free</u>

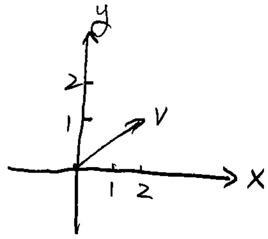
(i) Vector spaces



$v \in V$

Pick a coordinate β

Cartesian

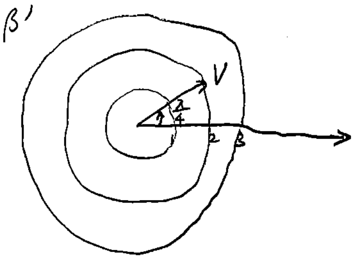


$$[v]_\beta = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \begin{matrix} \leftarrow x \\ \leftarrow y \end{matrix}$$

change of coordinate (basis).

Pick another coordinate β'

polar



$$[v]_{\beta'} = \begin{pmatrix} 2 \\ \frac{\pi}{4} \end{pmatrix} \begin{matrix} \leftarrow r \\ \leftarrow \theta \end{matrix}$$

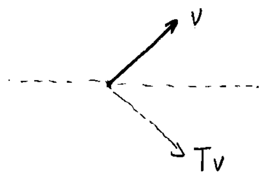
independent of coordinate ↑ <u>essential information</u> of an object
2040

<u>one</u> representation of an object.
1030

(ii) Linear maps

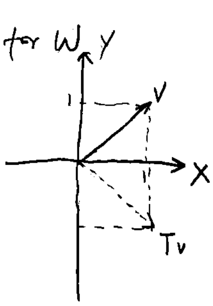
$$V \xrightarrow{T} W = V$$

v



Reflection by dotted line.

Pick a basis α for V, β for W



Choose $\alpha = \beta = \{\overset{(1,0)}{x}, \overset{(0,1)}{y}\}$.

$[v]_\alpha = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $[Tv]_\beta = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

T maps $\begin{pmatrix} 1 \\ 1 \end{pmatrix}_\alpha$ to $\begin{pmatrix} 1 \\ -1 \end{pmatrix}_\beta$

so $[T]_\alpha^\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

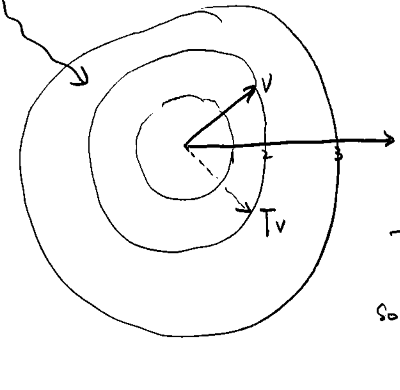
$$\Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$[T]_\alpha^\beta [v]_\alpha = [Tv]_\beta .$$

change of coordinate (basis)

Pick another basis α' for V and β' for W

Independent of coordinate.
↑
<u>essential information</u> of a linear map.
2040



Choose polar coordinate $\alpha' = \beta' = \{\overset{(1,0)}{r}, \overset{(0,1)}{\theta}\}$

$[v]_{\alpha'} = (2, \frac{\pi}{4})$ $\theta \in \mathbb{R}/2\pi\mathbb{Z}$

$[Tv]_{\beta'} = (2, -\frac{\pi}{4})$

$$-\frac{\pi}{4} = \frac{7\pi}{4}$$

T maps $\begin{pmatrix} 2 \\ \frac{\pi}{4} \end{pmatrix}$ to $\begin{pmatrix} 2 \\ \frac{3\pi}{4} \end{pmatrix}$.

so $[T]_{\alpha'}^{\beta'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ \frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} 2 \\ -\frac{\pi}{4} \end{pmatrix}.$$

$$[T]_{\alpha'}^{\beta'} [v]_{\alpha'} = [Tv]_{\beta'}$$

<u>one</u> representation of an object.
1030

Throughout this tutorial, we denote

V, W are $\underline{\text{finite}}$ dimension vector space over $\mathbb{F}$.

↑ we don't have $\infty \times \infty$ matrices.

$T: V \to W$ is a linear map:

⚠ By a basis we mean an $\underline{\text{ordered}}$ basis, i.e., the order matters. (eg. $\{e_1, e_2\} \neq \{e_2, e_1\}$ as basis).

Def. 4.1 (coordinate vector).

Choose a basis $\beta = \{e_1, \cdots, e_{\underset{\dim V}{n}}\}$ for V. Then for any $v \in V$.

$$v = \sum_{i=1}^{n} v_i e_i.$$

Write $[v]_\beta = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ called $\underline{\text{coordinate vector}}$ of v with respect to (w.r.t.) basis β.

Remark. (i) This is well-defined as the linear combination exist and unique (because of basis).

(ii). This defines a natural map $[\ ]_\beta : V \longrightarrow \mathbb{F}^n$

$$v \longmapsto [v]_\beta = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Every choice of basis gives rise to a map $V \to \mathbb{F}^n$.

Def. 4.2. (matrix representation).

For $T: V \to W$. Choose a basis $\alpha = \{e_1, \cdots, e_{\underset{\dim V}{n}}\}$ for V and a basis $\beta = \{e_1', \cdots, e_{\underset{\dim W}{m}}'\}$ for W. (n may not be equal to m).

Then we can represent T by a $\underline{\text{matrix}}$ $[T]_\alpha^\beta = A$ defined by.

$$A_{ij} = ([Te_j]_\beta)_i, \quad \text{or} \quad A = \begin{pmatrix} | & | & & | \\ [Te_1]_\beta & [Te_2]_\beta & \cdots & [Te_n] \\ | & | & & | \end{pmatrix} \quad \text{or}$$

$$Te_j = A_{1j} e_1' + A_{2j} e_2' + \cdots + A_{mj} e_m'.$$

⚠ Why want this def?

Eg. 4.3. $T: \mathbb{F}^2 \to \mathbb{F}^3$, $(x,y) \mapsto (x+3y, 2x+5y, 7x+9y)$.

(i) If we choose standard basis $\alpha = \{(1,0), (0,1)\}$ for $\mathbb{F}^2$, $\beta = \{(1,0,0), (0,1,0), (0,0,1)\}$ for $\mathbb{F}^3$.

then $T(1,0) = 1 \cdot (1,0,0) + 2(0,1,0) + 7(0,0,1) \in \mathbb{F}^3$, $[T(1,0)]_\beta = \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}$

$T(0,1) = 3 \cdot (1,0,0) + 5(0,1,0) + 9(0,0,1) \in \mathbb{F}^3$, $[T(0,1)]_\beta = \begin{pmatrix} 3 \\ 5 \\ 9 \end{pmatrix}$

so $[T]_\alpha^\beta = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}$.

(ii). If we choose a basis $\alpha' = \{(1,2),(2,-1)\}$ of $\mathbb{F}^2$ and $\beta' = \beta = \{(1,0,0),(0,1,0),(0,0,1)\}$ for $\mathbb{F}^3$.

Then

$$T(1,2) = 7(1,0,0) + 12(0,1,0) + 25(0,0,1). \rightsquigarrow [T(1,2)]_{\beta'} = \begin{pmatrix} 7 \\ 12 \\ 25 \end{pmatrix}.$$

$$T(2,-1) = (-1)(1,0,0) + (-1)(0,1,0) + 5(0,0,1). \rightsquigarrow [T(2,-1)]_{\beta'} = \begin{pmatrix} -1 \\ -1 \\ 5 \end{pmatrix}.$$

so $[T]_{\alpha'}^{\beta'} = \begin{pmatrix} 7 & -1 \\ 12 & -1 \\ 25 & 5 \end{pmatrix}$

(iii) What if switch order?

If we choose $\alpha'' = \alpha' = \{(1,2),(2,-1)\}$ for $\mathbb{F}^2$, and $\beta'' = \{(0,0,1),(0,1,0),(1,0,0)\}$ for $\mathbb{F}^3$.

Then

$$T(1,2) = 25(0,0,1) + 12(0,1,0) + 7(1,0,0). \rightsquigarrow [T(1,2)]_{\beta''} = \begin{pmatrix} 25 \\ 12 \\ 7 \end{pmatrix}.$$

$$T(2,-1) = 5(0,0,1) + (-1)(0,1,0) + (-1)(1,0,0). \rightsquigarrow [T(2,-1)]_{\beta''} = \begin{pmatrix} 5 \\ -1 \\ -1 \end{pmatrix}$$

so $[T]_{\alpha''}^{\beta''} = \begin{pmatrix} 25 & 5 \\ 12 & -1 \\ 7 & -1 \end{pmatrix}$. Note that $[T]_{\alpha''}^{\beta''} \neq [T]_{\alpha'}^{\beta'}$.

written vertically.

Ex: (i) $[T_1 \circ T_2]_\alpha^\beta = [T_1]_\gamma^\beta [T_2]_\alpha^\gamma$ (ii) $[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}$ if T is invertible. (iii) $[T]_\alpha^\beta [v]_\alpha = [Tv]_\beta$.

<u>Change of basis</u>

Q: Is there a systematic way to choose a basis so that $[T]_\alpha^\beta$ is $\left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right)$?

In particular, if $\dim V = \dim W$, is there a way to make $[T]_\alpha^\beta$ diagonal?

Ans: Consider the commutative diagram

$$\begin{tikzcd}
V \arrow[r, "I_V", "\cong"'] \arrow[d, "{[\ ]_{\alpha_0}}"] & V \arrow[r, "T"] \arrow[d, "{[\ ]_\alpha}"] & W \arrow[r, "I_W", "\cong"'] \arrow[d, "{[\ ]_\beta}"] & W \arrow[d, "{[\ ]_{\beta_0}}"] \\
\mathbb{F}^n \arrow[r, "{[I_V]_{\alpha_0}^{\alpha}}"'] & \mathbb{F}^n \arrow[r, "{[T]_\alpha^\beta}"'] & \mathbb{F}^m \arrow[r, "{[I_W]_\beta^{\beta_0}}"'] & \mathbb{F}^m
\end{tikzcd}$$

Eg 4.3. (Cont'd). $T(x,y)=(x+3y,\ 2x+5y,\ 7x+9y)$. $\alpha=\{(1,0),(0,1)\}$. $\beta=\{(1,0,0),(0,1,0),(0,0,1)\}$.

$$\begin{tikzcd}
V \arrow[r, "I_V"', "\cong"] \arrow[d, "{[\ ]_{\alpha_0}}"] \arrow[rrr, bend left, "\text{canonical embedding in first 2 coordinates.}"] & V \arrow[r, "T"] \arrow[d, "{[\ ]_{\alpha}}"] & W \arrow[r, "I_W"', "\cong"] \arrow[d, "{[\ ]_{\beta}}"] & W \arrow[d, "{[\ ]_{\beta_0}}"] \\
\mathbb{F}^n \arrow[r, "{[I_V]^{\alpha}_{\alpha_0}}"] \arrow[rrr, bend right, "{[T]^{\beta_0}_{\alpha_0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}}"'] & \mathbb{F}^n \arrow[r, "{[T]^{\beta}_{\alpha}}"] & \mathbb{F}^m \arrow[r, "{[I_W]^{\beta_0}_{\beta}}"] & \mathbb{F}^m
\end{tikzcd}$$

$$\begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 17 & -12 & 1 \end{pmatrix}$$

column operations.

First: $\begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}^{-1}}_{= \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -17 & 12 \end{pmatrix}$

row operations

Second: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 17 & -12 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -17 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$

So $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 17 & -12 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix} \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.

We try to find α_0 so that $[I_V]^{\alpha}_{\alpha_0} = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}$ Let $\alpha_0 = \{e_1, e_2\}$.

$$\underbrace{[I_V]^{\alpha}_{\alpha_0} \big([e_1]_{\alpha_0}\ [e_2]_{\alpha_0} \big)}_{\text{id !!!}} = \big(\underbrace{[I_V(e_1)]_{\alpha}}_{\overset{\text{def}}{=} [e_1]_\alpha}\ \underbrace{[I_V(e_2)]_{\alpha}}_{\overset{\text{def}}{=} [e_2]_\alpha} \big) = (e_1, e_2) \quad \text{as } \alpha \text{ is standard basis.}$$

Hence $e_1 = (-5, 2)$. $e_2 = (3, -1)$.

We try to find β_0 so that $[I_W]^{\beta_0}_{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 17 & -12 & 1 \end{pmatrix} \Rightarrow [I_W^{-1}]^{\beta}_{\beta_0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -17 & 12 & 1 \end{pmatrix} = \left([I_W]^{\beta_0}_{\beta} \right)^{-1}$. Let $\beta_0 = \{e_1', e_2', e_3'\}$

$$\underbrace{[I_W^{-1}]^{\beta}_{\beta_0} \big([e_1']_{\beta_0}\ [e_2']_{\beta_0}\ [e_3']_{\beta_0} \big)}_{\text{id !!!}} = \big([I_W(e_1')]_\beta\ [I_W(e_2')]_\beta\ [I_W(e_3')]_\beta \big) = (e_1'\ e_2'\ e_3')$$

as β is standard basis

Hence $e_1' = (1, 0, -17)$ $e_2' = (0, 1, 12)$. $e_3' = (0, 0, 1)$

You may check $[T]^{\beta_0}_{\alpha_0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ under these basis; our calculation is right.

Remark. Actually in definition of T, (x,y) means $x \cdot (1,0) + y(0,1)$ so we can view it to be coordinates w.r.t. standard basis.

Q. Why do we define

$$[T]_\alpha^\beta = \begin{pmatrix} | & & | \\ [Te_1]_\beta & & [Te_n]_\beta \\ | & & | \end{pmatrix} \ ? \qquad \alpha = \{e_1, \cdots, e_n\}. \ \beta = \{e_1', \cdots, e_m'\}.$$

Ans. We want

$$[Tv]_\beta = [T]_\alpha^\beta [v]_\alpha .$$

Indeed, $v = \sum_{i=1}^n v_i e_i = (e_1, \cdots, e_n) \underbrace{\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}}_{[v]_\alpha}$

$$Tv = \sum_{i=1}^n v_i (Te_i) = (Te_1 \cdots Te_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$$= (e_1' \cdots e_m') \underbrace{\begin{pmatrix} | & & | \\ [Te_1]_\beta & \cdots & [Te_n]_\beta \\ | & & | \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}}_{\begin{pmatrix} | \\ [Tv]_\beta \\ | \end{pmatrix}}.$$

Hence we define $\begin{pmatrix} | & & | \\ [Te]_\beta & \cdots & [Te_n]_\beta \\ | & & | \end{pmatrix}$ to be $[T]_\alpha^\beta$.

Note that here we write basis element horizontally, and coordinate vertically.

Midterm 1 Solution.

1. (a). Yes.

proof 1. First, $S_1 \neq \phi$ as $0 \in S_1$

(7')

- Let $x_i = (a_i+b_i,\ a_i-b_i+2c_i,\ b_i,\ c_i)$ $i=1,2$.

Then $x_1+x_2 = ((a_1+a_2)+(b_1+b_2),\ (a_1+a_2)-(b_1+b_2)+2(c_1+c_2),\ b_1+b_2,\ c_1+c_2) \in S_1$.

For any $\lambda \in \mathbb{R}$.

$$\lambda x_1 = (\lambda a_1 + \lambda b_1,\ \lambda a_1 - \lambda b_1 + 2(\lambda c_1),\ \lambda b_1,\ \lambda c_1) \in S_1.$$

Therefore by definition of subspace, S_1 is a vector subspace of $\mathbb{R}^4$.

- For the dimension, we need to give a basis. Take the standard basis of $\mathbb{R}^3$. $\{e_i\}_{i=1}^4$, where $e_i = (0,\cdots,1,\cdots,0)$ with 1 in the i^{th}-position.

Setting $(a,b,c) = e_i$, $i=1,2,3$ respectively, we get

$e_1' = (1,1,0,0)$.

$e_2' = (1,-1,1,0)$.

$e_3' = (0,2,0,1)$.

We claim $\{e_1', e_2', e_3'\}$ is a basis of S_1. Indeed, they are linearly independent as $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}$ has rank 3, as it can be reduced to

$\begin{pmatrix} * & 1 & 0 & 0 \\ * & 0 & 1 & 0 \\ * & 0 & 0 & 1 \end{pmatrix}$ by doing row operations. They span S_1 as

$$(a+b,\ a-b+2c,\ b,\ c) = a(1,1,0,0) + b(1,-1,1,0) + c(0,2,0,1) \in \operatorname{span}\{e_1', e_2', e_3'\}.$$

Hence $\dim_{\mathbb{R}} S_1 = 3$.

proof 2. Consider $T: \mathbb{R}^3 \to \mathbb{R}^4$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

$S_1 = \operatorname{Im} T$. Now T is linear as it is given by multiplication by a matrix. Its image is a subspace. $\ker T = \{0\}$ by solving $(a+b,\ a-b+2c,\ b,\ c) = (0,0,0,0)$. By rank-nullity theorem, $\dim \operatorname{Im} T = \dim \mathbb{R}^3 - \dim \ker T = 3$. Hence $\dim_{\mathbb{R}} S_1 = 3$.

b). No. as $1 \in S_2$ but $(-1)\cdot 1 = -1 \notin S_2$.

(3')

(c). Yes. First, $S_3 \neq \emptyset$ as $0 \in S_3$.

(7')

For $f_1, f_2 \in S_3$, $a \in \mathbb{R}$.

$(af_1+f_2)(1) = af_1(1) + f_2(1) = 0$.

$(af_1+f_2)(x) = af_1(x) + f_2(x) = af_1(-x) + f_2(-x) = (af_1+f_2)(-x)$

Hence $af_1+f_2 \in S_3$. Hence it is a vector subspace of $P_n(\mathbb{R})$.

Suppose $a_0 + a_1x + \cdots + a_nx^n \in S_3$.

$f(1)=0$ gives $a_0 + a_1 + \cdots + a_n = 0$.

$f(x) = f(-x)$ gives $a_1 = a_3 = \cdots = \underbrace{a_{2k+1}}_{\text{largest odd} \le n} = 0$ where $k = \begin{cases} \frac{n-1}{2} & n \text{ odd} \\ \frac{n}{2}-1 & n \text{ even} \end{cases}$

Consider the map $T: P_n(\mathbb{R}) \to \mathbb{R}^{k+2}$

$$a_0 + a_1x + \cdots + a_nx^n \mapsto (a_1, a_3, \cdots, a_{2k+1}, a_0 + a_2 + \cdots + a_n).$$

It is easy to check that this is linear. T is also surjective, as

$$(c_0, c_1, \cdots, c_{k+1}) = T(c_0x + c_1x^3 + c_2x^5 + \cdots + c_kx^{2k+1} + c_{k+1})$$

Observe that $S_3 = \ker T$. (so no need to show S_3 is a subspace above)

By rank-nullity theorem, $\dim \ker T = \dim P_n(\mathbb{R}) - \dim \mathbb{R}^{k+1}$

$= (n+1) - (k+2) = n-k-1$

Hence $\dim S_3 = n-k-1 = \lfloor \frac{n}{2} \rfloor$.

(3') (d) No. As $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in S_4$ but $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \notin S_4$.

[Yes/No. 1' (a), (c) proof of subspace 2', proof of dim 4'. (b) (d) Reason 2'].

Remark. Of course, there are many other ways to do it, for example, one may show that

$\{x^2-1, x^4-1, \cdots x^r-1\}$ is a basis for S_3, r is the largest even number $\le n$.

So $\dim S_3 = \lfloor \frac{n}{2} \rfloor$.

2. (a). Yes.

proof. $T_1(f_1+f_2) = x(f_1'+f_2') = xf_1' + xf_2' = T(f_1)+T(f_2)$

$T_1(af) = x(af)' = axf' = aT(f)$. $\quad a \in \mathbb{R}.\ f, f_1, f_2 \in P_n(\mathbb{R})$.

(b) No. For $f_1 = 1$, $f_2 = -1$, the constant function.

$T_2(f_1+f_2) = 0 \neq 2 = T_2(f_1) + T_2(f_2)$.

(c) Yes. Indeed, rotation counter clockwise by $\theta = \frac{\pi}{2}$ is given by $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

$T_3: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+1 \\ y \end{pmatrix} \xrightarrow{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}} \begin{pmatrix} -y \\ x+1 \end{pmatrix} \longmapsto \begin{pmatrix} -y \\ x \end{pmatrix}$

Hence, $T_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (Indeed, the translations cancelled).

It can be seen this is linear as it is given by a matrix.

(of course, one may directly verify that)

(d). Yes. $T_4 \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$

Let $A = (a_{ij})$. $B = (b_{ij})$. $\quad i, j = 1, \cdots, n$.

Looking entriwise, $T_4(A+B) = (a_{ij}+b_{ij}) = T_4(A) + T_4(b)$. $\quad 1 \le i,j \le n-1$

$T_4(\lambda A) = (\lambda a_{ij}) = \lambda T_4(A) \qquad 1 \le i,j \le n-1,\ \lambda \in \mathbb{F}$.

Hence T_4 is linear.

(Indeed, projections are linear).

[Yes/No 1'. Each question 5']

3. (a). (ii) is not true, as $R(T)$ is a vector subspace of $P_2(R)$.

(b). (ii) is not true.

proof. $\{1+x, x+x^2, \dots, x^3+x^4\}$ linearly independent set. To see that, assume there is a_i not all zero, such that

$$\sum_{i=0}^{3} a_i (x^i + x^{i+1}) = 0.$$

Let j be the largest i such that $a_i \neq 0$. Then

$$x^j + x^{j+1} = \frac{1}{a_j} \sum_{k=0}^{j-1} a_k (x^k + x^{k+1})$$

But this is impossible as degree on LHS is $j+1$, degree of RHS is at most j.
Hence the set is linearly independent, which implies

$$\dim R(S) \geq 4.$$

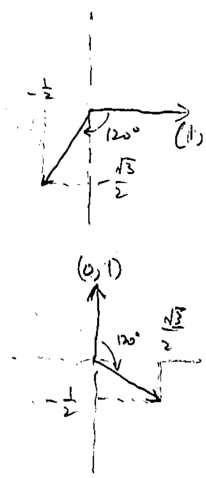
By rank nullity, $\dim P_2(R) \geq \dim R(S)$ but $\dim P_2(R) = 3 < 4$. $\Rightarrow \Leftarrow$.

[(a) 10'. (b) 10'].

4. (a). $T_1(1,0) = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) = -\frac{1}{2}(1,0) - \frac{\sqrt{3}}{2}(0,1)$

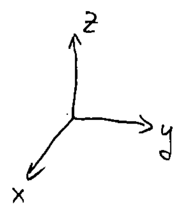
$T_1(0,1) = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{\sqrt{3}}{2}(1,0) - \frac{1}{2}(0,1)$.

Hence $[T_1]_\beta = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$ where β is the standard basis of R^2.



(b). $T_2(1,0,0) = (-1,0,0) = (-1)(1,0,0) + 0(0,1,0) + 0(0,0,1)$.

$(1,0,0) \xrightarrow{r_{yz}} (-1,0,0) \xrightarrow{r_{xy}} (-1,0,0) \xrightarrow{r_{xz}} (-1,0,0)$.



$T_2(0,1,0) = (0,-1,0) = 0(1,0,0) + (-1)(0,1,0) + 0(0,0,1)$.

$(0,1,0) \xrightarrow{r_{yz}} (0,1,0) \xrightarrow{r_{xy}} (0,1,0) \xrightarrow{r_{xz}} (0,-1,0)$.

$T_2(0,0,1) = (0,0,-1) = 0(1,0,0) + 0(0,1,0) + (-1)(0,0,1)$

$(0,0,1) \xrightarrow{r_{yz}} (0,0,1) \xrightarrow{r_{xy}} (0,0,-1) \xrightarrow{r_{xz}} (0,0,-1)$.

Hence $[T_2]_{\beta'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

where β' is the standard basis of R^3.

[Each (a), (b) worth 10'].

Q5. Define $\oplus: \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ by

$$(x, y) \mapsto xy$$

$$*: \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0} \text{ by}$$

$$(a, x) \mapsto x^a$$

Now we check this is isomorphic to $\mathbb{R}$ as a $\mathbb{R}$-vector space.

Consider $f_1: \mathbb{R} \to \mathbb{R}_{>0}$

$$x \mapsto e^x$$

$$g: \mathbb{R}_{>0} \mapsto \mathbb{R}$$

$$a \mapsto \ln a.$$

First, we show that they are linear. $a, b \in \mathbb{R}$. $x_1, x_2, x \in \mathbb{R}$, $y_1, y_2, y \in \mathbb{R}_{>0}$.

$$f(x_1+x_2) = e^{x_1+x_2} = e^{x_1} \cdot e^{x_2} = f(x_1) f(x_2) = f_1(x) \oplus f(x_2)$$

$$f(ax) = e^{ax} = (e^x)^a = f(x)^a = a * f(x)$$

$$g(y_1 \oplus y_2) = g(y_1 y_2) = \ln(y_1 y_2) = \ln(y_1) + \ln(y_2) = g(y_1) + g(y_2).$$

$$g(b * y) = g(y^b) = \ln y^b = b \ln y = b\, g(y).$$

Hence they are both linear maps and obviously inverse to each other. This gives $\mathbb{R}_{>0} \cong \mathbb{R}$ as $\mathbb{R}$-vector spaces.

[Definition 5'. Two maps f, g 5'. proof of linearly 10'].

Tutorial 5. Eigenvalues, Eigenspaces, Diagonalizability, Invariant Subspaces and Cayley-Hamilton Theorem.

We will focus on some insightful problems in this tutorial.

Q1. If a linear map $T: V \to V$ is <u>nilpotent</u>, i.e., $T^n = O$ (zero linear map) for some $n \in \mathbb{N}$, then all eigenvalues of T are 0.

pf. Suppose λ is its eigenvalue, with eigenvector $v \neq 0$. then

$$0 = T^n v = \lambda^n v$$

This implies $\lambda^n = 0$. Hence $\lambda = 0$. $\square$

Q2 (i) Find two 2×2 matrices which have the same characteristic polynomial but not similar.

(ii) How about 4×4 matrices? ($\mathbb{F} = \mathbb{R}$)

Ans. (i). $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$\chi_{A_1} = \chi_{A_2} = (1-x)^2$. They are definitely not similar as $PIP^{-1} = I$ $\forall P \in M_{2\times 2}(\mathbb{R})$.

(ii). $A_1 = \left(\begin{array}{cc|cc} 1 & 1 & & \\ 0 & 1 & \multicolumn{2}{c}{0} \\ \hline & & 1 & 0 \\ \multicolumn{2}{c|}{0} & 0 & 1 \end{array}\right)$ $A_2 = \left(\begin{array}{cc|cc} 1 & 1 & & \\ 0 & 1 & \multicolumn{2}{c}{0} \\ \hline & & 1 & 1 \\ \multicolumn{2}{c|}{0} & 0 & 1 \end{array}\right)$

$\chi_{A_1} = \chi_{A_2} = (1-x)^4$.

But A_1 and A_2 are not similar as the <u>geometric multiplicities</u> of eigenvalue 1 are different.

$\dim\ker(A_1 - I) = 3$. $\dim\ker(A_2 - I) = 2$.

Def 5.1. The <u>geometric multiplicity</u> of an eigenvalue λ of a linear map $T: V \to V$ is

$$\dim\ker(T - \lambda\, Id).$$

The dimension of eigenspace V_λ.

Prop. 5.2. The geometric multiplicity is an invariant for similar matrices, i.e., for invertible P,

$$\dim\ker(A - \lambda I) = \dim\ker(PAP^{-1} - \lambda I).$$

pf. $v \in \ker(A - \lambda I) \Leftrightarrow Av = \lambda v \Leftrightarrow PAP^{-1}(Pv) = PAv = \lambda Pv \Leftrightarrow Pv \in \ker(PAP^{-1} - \lambda I)$.

As P invertible, dim is preserved. $\square$

[<u>Algebraic multiplicity</u> is the exponent m_i of λ_i in $\chi_T(x) = (x - \lambda_i)^{m_i}(\cdots)$, where χ_T is the characteristic polynomial.]

Q3. Determine the formula for Fibonacci number x_n by $x_{n+2} = x_{n+1} + x_n$, $x_0 = 0$, $x_1 = 1$, i.e., Find x_n in terms of n.

Ans. $X_{n+3} = X_{n+2} + X_{n+1} = 2X_{n+1} + X_n$

Written in matrix, $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} X_n \\ X_{n+1} \end{pmatrix} = \begin{pmatrix} X_{n+2} \\ X_{n+3} \end{pmatrix}$.

let $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. For $n \geq 0$, we have

$$A^n \begin{pmatrix} X_0 \\ X_1 \end{pmatrix} = \begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix}.$$

To compute A^n, we factorize A. The characteristic polynomial

$$\chi_A(x) = x^2 - 3x + 1$$

so A has eigenvalues

$$\lambda_1 = \frac{3+\sqrt{5}}{2} = \varphi^2 \quad \lambda_2 = \frac{3-\sqrt{5}}{2} = \varphi^{-2}$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ is the Golden ratio.

and eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ \varphi \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ -\varphi^{-1} \end{pmatrix}.$$

So $\exists P$ invertible s.t. $P^{-1}AP = \begin{pmatrix} \varphi^2 & \\ & \varphi^{-2} \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 \\ \varphi & -\varphi^{-1} \end{pmatrix}$, $P^{-1} = \frac{1}{\sqrt{5}}\begin{pmatrix} \varphi^{-1} & 1 \\ \varphi & -1 \end{pmatrix}$.

Therefore

$$A^n = P(P^{-1}AP)^n P^{-1} = P\begin{pmatrix} \varphi^{2n} & \\ & \varphi^{-2n} \end{pmatrix} P^{-1} = \begin{pmatrix} \varphi^{2n-1} + \varphi^{-2n+1} & \varphi^{2n} - \varphi^{-2n} \\ \varphi^{2n} - \varphi^{-2n} & \varphi^{2n+1} + \varphi^{-2n-1} \end{pmatrix}.$$

So $\begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} \varphi^{2n} - \varphi^{-2n} \\ \varphi^{2n+1} + \varphi^{-2n-1} \end{pmatrix}$

$$X_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Remark. One may also do $\begin{pmatrix} X_{n+1} \\ X_{n+2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} X_n \\ X_{n+1} \end{pmatrix}$ significantly simplifying the process.

Q4. Condition 1: $\mathbb{F}=\mathbb{C}$ or algebraically closed field. ← every non-constant polynomial has a root in the field.

condition 2: $\dim V < \infty$. $V \neq 0$.

(i) If both conditions satisfied, any linear map $T: V \to V$ has a non-zero eigenvector.

(ii) Give counterexamples when either condition is omitted.

pf. (i). Let $n = \dim V < \infty$. Consider the set of $n+1$ vectors.

$$\{v, Tv, \cdots, T^n v\}$$

This must be linearly dependent as $n+1 > \dim V$, so $\exists\, a_i \in \mathbb{C}$ s.t.

$$\sum_{i=0}^{n} a_i T^i v = 0.$$

Consider $f(x) = \sum_{i=0}^{n} a_i x^i$. By <u>fundamental theorem of algebra</u>,

$$f(x) = a_0 (x - x_1)(x - x_2) \cdots (x - x_n) \quad \text{where } x_1, \cdots, x_n \text{ are roots of } f(x).$$

As $f(T) = 0$ by above,

$$a_0 (T - x_1 Id)(T - x_2 Id) \cdots (T - x_n Id) = 0.$$

This means at least one of $T - x_i Id$ is not invertible (by taking determinant for example).

Therefore we have a corresponding eigenvector for this eigenvalue x_i.

(ii) If $\mathbb{F} = \mathbb{R}$. $T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ for $\theta \notin 2\pi\mathbb{Z}$ has no eigenvectors, as it rotates all vectors in $\mathbb{R}^2$.

If $\mathbb{F} = \mathbb{C}$ and V is infinite dimensional, take

$$T(x_1, x_2, x_3, \cdots) = (0, x_1, x_2, x_3, \cdots) \quad \text{right-shift.}$$

Then we claim T has no eigenvectors.

Suppose $T(v_1, v_2, \cdots) = \lambda (v_1, v_2, \cdots)$ $\lambda \in \mathbb{C}$. $v = (v_1, v_2, \cdots)$ is an eigenvector.

$$\underset{(0, v_1, v_2, \cdots)}{\underset{\shortparallel}{T(v_1, v_2, \cdots)}} = \underset{(\lambda v_1, \lambda v_2, \cdots)}{\underset{\shortparallel}{\lambda (v_1, v_2, \cdots)}}$$

Then $\lambda v_1 = 0$. $v_1 = \lambda v_2$, $v_2 = \lambda v_3$ etc.

If $\lambda \neq 0$, then $v = 0$. $\Rightarrow \Leftarrow$.

If $\lambda = 0$ then $v_1 = \lambda v_2 = 0$, $v_2 = \lambda v_3 = 0$ etc, $v = 0$ $\Rightarrow \Leftarrow$.

Diagonalizability.

Def 5.3. Let V be a finite dimensional vector space. $T: V \to V$ a linear map.

(i) T is <u>diagonalizable</u> iff $\exists$ a basis β of V s.t. $[T]_\beta$ is diagonal.

(ii) β is called <u>eigenbasis</u> of T consisting of eigenvectors of T.

(iii) $E_{\lambda_i} = \ker(T - \lambda_i Id)$ is called <u>eigenspaces</u> of T with respect to eigenvalue λ_i.

Eg. $T: \mathbb{R}^5 \to \mathbb{R}^5$, $(x_1, x_2, x_3, x_4, x_5) \longmapsto (2x_1, 2x_2, 2x_3, 3x_4, 3x_5)$

Then just choose standard basis $\beta = \{e_1, e_2, e_3, e_4, e_5\}$ of $\mathbb{R}^5$.

$$[T]_\beta = \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 3 & \\ & & & & 3 \end{pmatrix} \quad \text{so } \lambda_1 = 2, \ \lambda_2 = 3.$$

$V = E_{\lambda_1} \oplus E_{\lambda_2}$

$= \text{span}\{e_1, e_2, e_3\} \oplus \text{span}\{e_4, e_5\}$

$$[T]_\beta = \left(\begin{array}{c|c} [T|_{E_{\lambda_1}}]_\beta & \\ \hline & T|_{E_{\lambda_2}} \end{array}\right) \text{ is block diagonal.}$$

Prop. 5.4 Let $T: V \to V$ and $\dim V < \infty$. Suppose $\lambda_1, \dots, \lambda_m$ are <u>distinct</u> eigenvalues of T with eigenspaces E_{λ_i}.

(i) Each E_{λ_i} is T-invariant.

(ii) T is diagonalizable iff $V = \bigoplus_{i=1}^{m} E_{\lambda_i}$. ← span of corresponding eigenbasis of λ_i.

(iii) If $V = \bigoplus_{i=1}^{m} V_i$ where V_i are all T-invariant, then

T is diagonalizable $\Leftrightarrow$ $T|_{V_i}$ are diagonalizable for all i.

Remark. (iii) is to look at $[T]_\beta$ as blocks $[T|_{V_i}]_\beta$. It is obvious to see it in matrices.

Def. 5.5. (simultaneously diagonalizability).

Let V be a finite dimensional vector space and $T, S: V \to V$ two linear maps.

We say T, S are <u>simultaneously diagonalizable</u> if $\exists$ a basis β for V s.t. $[T]_\beta$ and $[S]_\beta$ are diagonal matrices.

same basis

[We say A, B are " " " " if $\exists P$ invertible s.t. $P^{-1}AP$ and $P^{-1}BP$ are diagonal]

Q5. In the setting above,

(i). If T, S are simultaneously diagonalizable, then $TS = ST$.

(ii) If T, S are diagonalizable and $TS = ST$, then T, S are simultaneously diagonalizable.

Pf. (i) As $[T]_\beta, [S]_\beta$ are diagonal matrices, $[T \circ S]_\beta = [T]_\beta [S]_\beta = [S]_\beta [T]_\beta = [S \circ T]_\beta$ so $TS = ST$.

(ii). By change of basis formula, T, S are simultaneously diagonalizable if we can find an eigenbasis β common to T and S. (P is the matrix consists of eigenbasis β)

Let $V = \bigoplus_{i=1}^{m} E_{\lambda_i}$ where $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T, following from prop 5.4 (ii) and T being diagonalizable. Now we claim $S(E_{\lambda_i}) \subseteq E_{\lambda_i}$. As for every $v \in E_{\lambda_i}$,

$$TSv \overset{\text{commutative}}{=} STv = S\lambda_i v = \lambda_i (Sv).$$

so $Sv \in \ker(T - \lambda_i I) = E_{\lambda_i}$. Now by Prop 5.4 (iii), $S|_{E_{\lambda_i}}$ is diagonalizable.

Take an S-basis / β_i for E_{λ_i} and $\coprod_{i=1}^{m} \beta_i$ is the required basis β. Indeed, $[S]_\beta$ is diagonal by construction of β_i and prop 5.4(iii). $[T]_\beta$ is diagonal because E_{λ_i} are eigenspaces of T, for $i = 1, \dots, m$.

<u>Characteristic polynomial</u>

Def. For $A \in M_{n\times n}(\mathbb{R})$, the characteristic polynomial $\chi_A = \det(A - xI)$.

For $T: V \to V$, the characteristic polynomial $\chi_T = \det(A - xI)$ where $[T]_\beta = A$ for some basis.

Remark. 1). χ_T is independent of choice of basis, so $\chi_A = \chi_{PAP^{-1}}$.

Moreover, for any polynomial f, $f(PAP^{-1}) = P f(A) P^{-1}$. To see this, say $f = a_0 + a_1 x + \dots + a_n x^n$

$$\begin{aligned} f(PAP^{-1}) &= a_0 + a_1 PAP^{-1} + a_2 PAP^{-1}PAP^{-1} + \dots + a_n (PAP^{-1})^n \\ &= a_0 PP^{-1} + a_1 PAP^{-1} + a_2 PA^2P^{-1} + \dots + a_n PA^n P^{-1} \\ &= P(a_0 + a_1 A + a_2 A^2 + \dots + a_n A^n) P^{-1} = P f(A) P^{-1} \end{aligned}$$

In particular, $\chi_A(PAP^{-1}) = P\chi_A(A)P^{-1} = 0$ by Cayley-Hamilton. Also $\chi_{PAP^{-1}}(A) = 0$.

2). $\chi_A = \det(A - xI) = (-1)^n x^n + (-1)^{n-1} \underline{\underline{tr(A)}} x^{n-1} + \dots + \underline{\underline{\det A}}$.

Invariant subspaces

Q6. Let $T: V \to V$ and $\dim V < \infty$. W is an invariant subspace under T. If $v_1, \cdots, v_n$ are eigenvectors to T corresponding to distinct eigenvalues $\lambda_1, \cdots, \lambda_n$ s.t. $v_1 + \cdots + v_n \in W$, then $v_i \in W$ for all i.

pf. (Try to learn from the proof "$v_1, \cdots, v_n$ are linearly independent").

First, we have

$$v_1 + \cdots + v_n \in W \quad ①$$

As W is T-invariant, applying T to ① gives

$$T(v_1 + \cdots + v_n) = Tv_1 + \cdots + Tv_n = \lambda_1 v_1 + \cdots + \lambda_n v_n \in W. \quad ②$$

Let $\lambda_1 \times ① - ②$ gives

$$(\lambda_1 - \lambda_2) v_2 + \cdots + (\lambda_1 - \lambda_n) v_n \in W.$$

But then let $v_i' = (\lambda_1 - \lambda_i) v_i$.

$$Tv_i' = T(\lambda_1 - \lambda_i) v_i = \lambda_i (\lambda_1 - \lambda_i) v_i = \lambda_i v_i'.$$

so $v_2', \cdots, v_n'$ are now in the situation of $n-1$ case. Induction gives the result. $\square$

Q7. Let $\mathbb{F} = \mathbb{C}$. Let $T: V \to V$ be a linear map. $\dim V < \infty$.

(i). Show that there exist $r \leq \dim V$

$$\{0\} \subsetneq \ker T \subsetneq \ker T^2 \subsetneq \cdots \subsetneq \ker T^r = \ker T^{r+1} = \cdots$$

(ii) Show that

$$V \cong \ker T^r \oplus \underbrace{T^r V}_{\text{range } T^r}$$

Now suppose the only eigenvalues for T are 0 and $\lambda \neq 0$. Let $W = \text{range } T^r$.

(iii). Show that W is a T-invariant and $T|_W$ has only eigenvalue λ. hint: use (ii) and Q4.

(iv). Let $S = (T - \lambda \text{Id})|_W$. Show that 0 is the only eigenvalue of S and $S^m = 0$ for some m.

pf. (i) First, it is obvious $\ker T^i \subseteq \ker T^{i+1}$ as for any $v \in \ker T^i \Rightarrow T^i v = 0 \Rightarrow T^{i+1} v = 0 \Rightarrow v \in \ker T^{i+1}$

Second, we want to show there exist r s.t. $\ker T^r = \ker T^{r+1}$

Suppose not, $\{0\} \subsetneq \ker T \subsetneq \cdots \subsetneq \ker T^r \subsetneq \ker T^{r+1} \subsetneq \cdots$ continuous indefinitely, then $\exists x_i \in \ker T^i \setminus \ker T^{i-1}$ for all $i \in \mathbb{N}$. We claim $\{x_1, x_2, \cdots\}$ is a linearly independent set. Indeed, $x_i \notin \text{span}\{x_1, \cdots, x_{i-1}\}$ as $x_i = \sum_{j=1}^{i-1} a_j x_j \Rightarrow T^{i-1} x_i = \sum_{j=1}^{i-1} a_j \underbrace{T^{i-1} x_j}_{0} = 0$ as $x_j \in \ker T^j \subseteq \ker T^{i-1}$ for all $j \leq i-1$. $\Rightarrow x_i \in \ker T^{i-1} \Rightarrow \Leftarrow$.

Hence $\{x_i\}_{i \in \mathbb{N}}$ is a linearly independent set in V. But V is finite dimensional, this is impossible, so there exist $r \leq \dim V$, $\ker T^r = \ker T^{r+1}$.

Third, we show $\ker T^{r+k+1} = \ker T^{r+k}$ for all $k \in \mathbb{Z}_{\geq 0}$. Indeed, $v \in \ker T^{r+k+1} \Rightarrow T^{r+1}(T^k v) = 0$
$\Rightarrow T^k v \in \ker T^{r+1} = \ker T^r \Rightarrow T^r(T^k v) = 0 \Rightarrow v \in \ker T^{r+k}$

(ii). Pick a basis $\{e_i'\}$ for T^rV. Let $e_i \in V$ s.t. $T^r e_i = e_i'$. Then $e_i' \mapsto e_i$ extends to a linear map $\varphi: T^rV \longrightarrow V$, $e_i' \mapsto e_i$. As $T^r: V \to T^rV$, $e_i \mapsto e_i'$, we have $T^r \circ \varphi = id_{T^rV}$.

Claim: $V = \ker T^r \oplus \varphi(T^rV)$.

First, we show $V = \ker T^r + \varphi(T^rV)$.

Indeed, for any $v \in V$, $v = (v - \varphi(T^rv)) + \varphi T^r v$. It suffices to show $v - \varphi(T^rv) \in \ker T^r$.

But this is clear as $T^r(v - \varphi T^r v) = T^r v - \underbrace{T^r\varphi}_{id} T^r v = T^r v - T^r v = 0$.

Second, we show $\ker T^r \cap \varphi(T^rV) = \{0\}$.

For $v \in \varphi(T^rv) \cap \ker T^r$, say $v = \varphi T^r u$ for some $u \in V$. Then

$$\underset{\substack{\uparrow \\ v \in \ker T^r}}{0} = T^r v = \underbrace{T^r\varphi}_{id} T^r u = T^r u \underset{\substack{\uparrow \\ \text{apply } \varphi}}{\Rightarrow} 0 = \varphi(0) = \varphi T^r u = v. \Rightarrow v = 0.$$

This concludes the proof of claim, and hence (ii).

(iii). First, $T(W) \subseteq W$ as $T(T^r v) = T^r(\underbrace{Tv}_{\text{in } V}) \in T^rV = W$.

Second, for any T-eigenvector $w = T^r u \in W$ with eigenvalue $\alpha \in \{0, \lambda\}$. If $\alpha = 0$.

$T^{r+1} u = Tw = \alpha w = 0 \Rightarrow u \in \ker T^{r+1} = \ker T^r \Rightarrow w = T^r u = 0 \Rightarrow \Leftarrow$. So $\alpha = \lambda$.

(iv). As $T|_W$ has only eigenvalue λ, $(T - \lambda I)|_W$ has only eigenvalue 0.

Indeed, if $w = T^r v \in W$ is an eigenvector of S with eigenvalue α. Then

$$\alpha w = Sw = Tw - \lambda w. \Rightarrow Tw = (\lambda + \alpha) w.$$

By (iii), $\lambda + \alpha = \lambda$ so $\alpha = 0$.

For the claim $S^m = 0$ for some m, replacing V, T by W, S in (ii) we have

$$W = \ker S^m \oplus S^m W \quad \text{for some } m$$

As $\mathbb{F} = \mathbb{C}$, by Q4(i), if $S^m W \neq 0$, then there is a eigenvector $0 \neq w \in S^m W$.

But the only eigenvalue is zero, so $S^m w = 0$. Hence $w \in \ker S^m \cap S^m W = \{0\} \Rightarrow \Leftarrow$.

Hence $S^m W = 0$.

$\square$.

Q8. Let T be a linear operator on V. Suppose V is T-cyclic, i.e.,

$$V = \text{span}\{v, Tv, T^2v, \cdots\}$$

for some generator $v \in V$.

For another linear operator U on V, show that

$$TU = UT \Leftrightarrow U = g(T) \text{ for some polynomial } g(t).$$

proof. $(\Leftarrow)$ is easy. As T commutes with any polynomial in T,

$$TU = Tg(T) = g(T)T = UT.$$

$(\Rightarrow)$. Let $v \in V$ be a generator of V. Then every $w \in V$ is written as

$$w = f(T)v \qquad \text{for some polynomial } f(t).$$

In particular, for $w = U(v)$,

$$U(v) = g(T)v \qquad \text{for some polynomial } g(t).$$

We claim $U = g(T)$. Indeed, for any $x \in V$, $x = h(T)v$ for some polynomial $h(t)$.

$$Ux = U(h(T)v) \underset{UT=TU}{=} h(T)Uv = h(T)g(T)v = g(T)h(T)v = g(T)x.$$

Since x is arbitrary, $U = g(T)$ □.

<u>Cayley-Hamilton Theorem</u>

Q9. Let A be a 2×2 matrix with eigenvalue $-1, 2$. Find the inverse of $B = A - I$ in terms of A and I.

Ans. Since eigenvalue of A is $-1, 2$, eigenvalue of $B = A - I$ is $-2, 1$ as

$$Av = -v \Rightarrow (A-I)v = -v - v = -2v.$$
$$Av' = 2v' \Rightarrow (A-I)v' = 2v' - v' = v'.$$

Then $\chi_B = (x+2)(x-1)$. By Cayley-Hamilton theorem,

$$0 = (B+2I)(B-I) = B^2 + B - 2I.$$

$$\Rightarrow B(B+I) = 2I. \Rightarrow B\cdot\left(\frac{A}{2}\right) = I$$

$$\Rightarrow B^{-1} = \frac{A}{2}.$$

□.

Eigenspace and generalized eigenspace.

Def. 5.6. Let T be a linear operator on V and λ be an eigenvalue.

(i). The <u>eigenspace</u> for λ is
$$E_\lambda := \ker(T-\lambda I) = \{x \in V \mid Tx = \lambda x\}.$$

(ii). The <u>generalized eigenspace</u> for λ is
$$K_\lambda := \ker(T-\lambda I)^n = \{x \in V \mid (T-\lambda I)^n x = 0 \text{ for some } n\}.$$

for some $n \in \mathbb{N}$. If V is finite dimensional, we may take $n = \dim V$.

Remark. We have composite series
$$\{0\} \subsetneq \ker(T-\lambda I) \subsetneq \ker(T-\lambda I)^2 \subsetneq \cdots \subsetneq \ker(T-\lambda I)^r = \ker(T-\lambda I)^{r+1} = \cdots$$
where r is the first place the chain stablizes, and we have $r \le \dim V$.

In the definition we take $n = \dim V$ will be the biggest possible vector space of this form.

Eg. The most important example:
$$V = \{\mathbb{R} \to \mathbb{C} \text{ differentiable}\} \qquad D = \frac{d}{dt} : V \to V \quad \text{is a linear operator.}$$

- For any $\lambda \in \mathbb{C}$, the λ-eigenspace is
$$E_\lambda = \{f \in V \mid \frac{df}{dt} = \lambda f\}$$

  This is a first order differential equation.
$$E_\lambda = \{Ae^{\lambda t} \mid A \in \mathbb{C}\}.$$

  It is one dimensional, with geometric multiplicity 1.

- For the same λ, the generalized λ-eigenspace appear when you solve higher order DE's.

  e.g.: $$f'' - 2f' + f = 0$$
  is written as $$(D-I)^2 f = 0$$

  One solution is e^t, the other is te^t. This is because
$$D(te^t) = Dt(e^t) + t(De^t) = e^t + te^t$$
$$\Rightarrow \quad (D-I)(te^t) = e^t \in \ker(D-I)$$
$$\Rightarrow \quad te^t \in \ker(D-I)^2.$$

Hence, the generalized λ- eigenspace for D is

$$K_\lambda = \{p(t)e^{\lambda t} \mid p \in \underset{\uparrow}{\mathbb{C}[t]}\}.$$

polynomial with coefficients in $\mathbb{C}$.

This is infinite dimensional, as $\mathbb{C}[t]$ is infinite dimensional. The geometric multiplicity for K_λ is infinite.

We want to study generalized eigenspace because for T not diagonalizable, we may not have enough eigenvector to form a basis for V. For example, $V = \mathbb{R}^2$.

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \underset{\uparrow}{\chi_A} = (x-1)^2. \quad E_1 = \ker(A-I) = \ker\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \operatorname{span}\{e_1\}.$$

characteristic polynomial.

but V is 2-dimensional! $\quad K_1 = \ker(A-I)^2 = \ker\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \ker\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \operatorname{span}\{e_1, e_2\}$

Now by considering generalized eigenspace, we have a eigenbasis for V.

In general, we have the following theorem.

(Primary decomposition theorem / spectral decomposition theorem).

Theorem 5.7. Let V be a vector space with $\dim V < \infty$. T be a linear operator on V. $\lambda_1, \cdots, \lambda_r$ are distinct eigenvalue of T. Then

$$V = \bigoplus_{i=1}^{r} K_{\lambda_i}$$

Moreover, each K_{λ_i} is T-invariant

- $\chi_T = (x-\lambda_1)^{q_1} \cdots (x-\lambda_r)^{q_r}$ the characteristic polynomial.
- $\chi_{T|_{K_{\lambda_i}}} = (x-\lambda_i)^{q_i}$ for $i = 1, \cdots, r$, $\sum_{i=1}^{\lambda} q_i = \dim V$.
- $(\text{algebraic multiplicity of } \lambda_i) \overset{\text{def}}{=} q_i = \dim K_{\lambda_i}$.

Eigendecomposition. $T \in \mathcal{L}(V)$. V finite dimensional. $n = \dim V$.

Question: Is it possible to decompose V into T-eigenspaces?

Ans : Four cases ①-1 $\Rightarrow$ ①-2 $\Rightarrow$ ② $\Rightarrow$ ③.

① Best case ; $V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_n}$

①-1. λ_i's are distinct, with eigenvectors v_i, $i = 1, \ldots, n$.

$\Leftrightarrow$ $\dim E_{\lambda_i} = 1$, and $E_{\lambda_i} = \operatorname{span}\{v_i\}$ and λ_i distinct.

$\Leftrightarrow$ $\{v_i\}_{i=1}^{n}$ are linearly independent

$\Leftrightarrow$ $\{v_i\}_{i=1}^{n}$ form an eigenbasis of V.

$\Leftrightarrow$ $\chi_T(x) = c(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ splits into distinct factors.

Eg. $T = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{pmatrix}$. $V = \mathbb{R}^4$.

①-2. Some of λ_i are repeated, but we still have an eigen basis.

$\Leftrightarrow$ T is diagonalizable.

$\Leftrightarrow$ every algebraic multiplicity = geometric multiplicity, for all λ_i.

(Some of E_{λ_i} can be more than one dimensional.)

Eg. $T = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 2 & & \\ & & & 3 & \\ & & & & 3 \end{pmatrix}$ $V = \mathbb{R}^5$. is ①-2 but not ①-1.

② Not so good : $V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots \oplus K_{\lambda_m}$ λ_i distinct, K_{λ_i} is the generalized eigenspaces, $m \leq n$.

$\Leftrightarrow$ $\chi_T(x) = c(x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_m)^{\lambda_m}$ splits but may have repeated factors.

- $\dim K_{\lambda_i} = m_i$ = algebraic multiplicity.
- $\dim E_{\lambda_i}$ = geometric multiplicity $\leq$ $\dim K_{\lambda_i}$ = algebraic multiplicity.

Eg. $T = \begin{array}{c} \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix} \\ \begin{matrix} e_1 & \; & \; & e_4 & e_5 & \; \end{matrix} \end{array}$

$\chi_T(x) = (x-0)^4(x-1)^2$

$\lambda_1 = 0$. $E_0 = \operatorname{span}\{e_1, e_4\}$. $\subseteq K_0 = \operatorname{span}\{e_1, e_2, e_3, e_4\}$

$\lambda_2 = 1$. $E_1 = \operatorname{span}\{e_5\}$. $\subseteq K_1 = \operatorname{span}\{e_5, e_6\}$.

is ② but not ①-2 or ①-1.

If $\mathbb{F} = \mathbb{C}$, we always have $\chi_T(x)$ splits.

Thm. Whenever $\chi_T(x)$ splits, there exist basis β s.t. $[T]_\beta$ is in <u>Jordan normal form</u>:

Eg. $$[T]_\beta = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{matrix}} & & & \\ & \boxed{\begin{matrix} \lambda_1 & 1 \\ & \lambda_1 \end{matrix}} & & \\ & & \boxed{\lambda_2} & \\ & & & \boxed{\begin{matrix} \lambda_2 & 1 & & \\ & \lambda_2 & 1 & \\ & & \lambda_2 & 1 \\ & & & \lambda_2 \end{matrix}} \end{pmatrix}$$

size 3

size 2

size 1

size 4.

each of blocks is called Jordan block.

$\dim E_{\lambda_i} = \#$ Jordan blocks with eigenvalue λ_i $= \#$ lin indep eigenvectors of λ_i = geometric multiplicity.

$\dim K_{\lambda_i} = \Sigma$ size of all Jordan blocks with eigenvalue λ_i = algebraic multiplicity.

$=$ largest exponent m_i s.t. $(x-\lambda_i)^{m_i} \mid \chi_T(x)$.

Each Jordan block generates a T-cyclic space, for example,

$$T = \begin{pmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{pmatrix} \quad \text{has generator } e_3 \text{ as}$$

$$\{ e_3, \underset{\substack{\parallel \\ e_2}}{Te_3}, \underset{\substack{\parallel \\ e_1}}{T^2e_3} \} \text{ is a basis of } \mathbb{R}^3$$

③ Worst case. $\chi_T(x) = c \cdot \prod_{i=1}^{k} f_i^{q_i}$ f_i irreducible, but may not be linear, i.e., $\chi_T(x)$ may not splits.

Eg. $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ when $\theta \notin \pi\mathbb{Z}$ $\quad \mathbb{F} = \mathbb{R}$. $\quad V = \mathbb{R}^2$.

$\chi_T(x) = x^2 - 2\cos\theta x + 1$, discriminant $\Delta = 4(\cos^2\theta - 1) < 0$ for $\theta \notin \pi\mathbb{Z}$.

has no solution in $\mathbb{R}$! so this example is ③ but not ① or ②.

$K_\lambda, E_\lambda = \{0\}$ for all $\lambda \in \mathbb{R}$. Therefore, we cannot decompose $V = K_{\lambda_1} \oplus \cdots \oplus K_{\lambda_m}$.

(*) A more general version of <u>primary decomposition</u> is available.

$$V = \bigoplus_{i=1}^{k} (\ker f_i)^{q_i}$$

Ref. "https://math.mit.edu/~dav/generalized.pdf" notes on "generalized eigenspaces", 2019. (Thm 6.7)
Artin, Algebra, Prentice Hall Inc, 1991

Tutorial 6 Inner Product Spaces and their Operators

Def. 6.1. An <u>inner product</u> is a function $\langle\ ,\ \rangle : V\times V \to \mathbb{F}$, where $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, s.t.

(IP1). Sesquilinear $\langle ax+y, z\rangle = a\langle x,z\rangle + \langle y,z\rangle$. $\forall x,y,z\in V,\ a\in\mathbb{F}$.

Linear in first and conjugate linear in second

$\langle x, ay+z\rangle = \bar{a}\langle x,y\rangle + \langle x,z\rangle$

conjugate : $\overline{a+bi} = a-bi$

(IP2). Conjugate symmetric $\langle x,y\rangle = \overline{\langle y,x\rangle}$ $\forall x,y\in V$.

(IP3). positive definite $\langle x,x\rangle \geq 0$ with $=$ iff $x=0$. $\forall x\in V$.

Remark. (i) (IP2) implies $\langle x,x\rangle\in\mathbb{R}$, so (IP3) make sense as $\mathbb{R}$ has a natural order $\geq$.

(ii) When $\mathbb{F}=\mathbb{R}$, a function $\langle\ ,\ \rangle$ satisfying (IP1) is called a <u>bilinear form</u>.

a function satisfying (IP1)+(IP2) is called a <u>symmetric bilinear form</u>.

So an inner product is a positive definite symmetric bilinear form.

(a function $\langle\ ,\ \rangle : V\times V\to\mathbb{F}$ is called a <u>form</u>)

They are all very important objects in linear algebra, and appear everywhere in geometry.

When $\mathbb{F}=\mathbb{C}$, we also study <u>Hermitian form</u> : A function $H(\ ,\ ): V\times V\to\mathbb{F}$ satisfying (IP1)+(IP2). They are heavily studied in analysis, especially functional analysis.

(iii). The dictionary :

$\mathbb{F}=\mathbb{R}$ Symmetric bilinear form $\langle Tv,w\rangle=\langle v,Tw\rangle \longleftrightarrow$ symmetry matrix $A^T=A \longleftarrow A=[T]_\beta$.

$\mathbb{F}=\mathbb{C}$ Hermitian form $\langle Tv,w\rangle=\langle v,Tw\rangle \longleftrightarrow$ Hermitian matrix $A^H=A$.

orthonormal basis

(iv) For every symmetry bilinear form B we can associate a <u>quadratic form</u> $q: V\to\mathbb{F}$ defined by $q(x)=B(x,x)$. A <u>quadratic form</u> is a function $q: V\to\mathbb{F}$ satisfying

(QF1). $q(ax)=a^2q(x)$ $\forall x\in V$, $a\in\mathbb{F}$.

(QF2). $q(x+y)-q(x)-q(y)$ is bilinear $\forall x,y\in V$.

Eg. $q(x)=x^TQx$, where Q is a matrix corresponding to q if a basis is chosen.

Eg. $q(x,y)=4x^2+2xy-3y^2 = \begin{pmatrix} x & y\end{pmatrix}\begin{pmatrix}4 & 1\\ 1 & -3\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$ so we may just study $Q=\begin{pmatrix}4 & 1\\ 1 & -3\end{pmatrix}$.

Classification of quadratic form gives classification of quadrics in $\mathbb{R}^n$.

When $n=3$, we may classify all <u>conic sections</u> (圓錐曲線) in high school!

Eg. $q(x,y) = x^2+4xy+y^2$

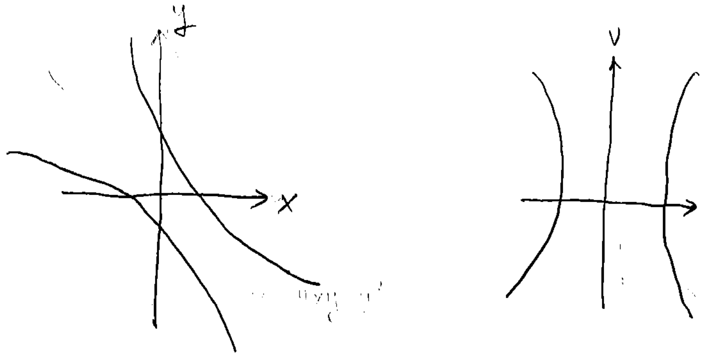
$$= (x\ y)\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

change of basis.

$$= (x\ y)\,P\begin{pmatrix}3 & 0\\ 0 & -1\end{pmatrix}P^{-1}\begin{pmatrix}x\\ y\end{pmatrix} \qquad P = \frac{1}{2}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} = \text{rotation by } \frac{\pi}{4}$$

$$= (w\ v)\begin{pmatrix}3 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}w\\ v\end{pmatrix} \qquad \text{where } (w,v) = (x,y)P$$

$$= 3w^2 - v^2 \qquad\qquad = \left(\tfrac{1}{\sqrt{2}}(x+y),\ \tfrac{1}{\sqrt{2}}(y-x)\right)$$



(v) A bilinear form $B: V\times V \to \mathbb{F}$ is equivalent to a linear map $\tilde{B}: V\otimes V \to \mathbb{F}$.

tensor product.

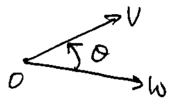
In this language, we can study multilinear maps: $f: V^{\otimes n} \to \mathbb{F}$.

This is another generalization of this topic.

* all remarks above are non-examinable except (i).

Def. 6.2. An <u>inner product space</u> is a vector space with a inner product on it.

Remark (i) Recall $\cos\theta = \dfrac{\langle v,w\rangle}{\sqrt{\langle v,v\rangle\langle w,w\rangle}}$



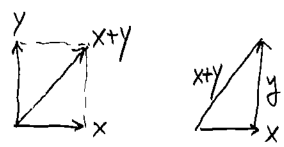
, $\langle v,w\rangle = v^T w$ here.

$\|v\| = \sqrt{\langle v,v\rangle}$ length of v.

So on inner product space we may talk above "angle" and "length".

(ii) Pythagorean theorem (勾股定理)

$\|x\|^2 + \|y\|^2 = \|x+y\|^2$ where $x \perp y$ $\quad \langle x,y\rangle = 0$



This works for arbitrary dimension!

$$\sum_{i=1}^n \|x_i\|^2 = \Big\|\sum_{i=1}^n x_i\Big\|^2 \quad \text{if } \langle x_i, x_j\rangle = 0 \text{ for all } i\neq j.$$

(iii). Cauchy inequality (柯西不等式)

$$(x_1^2+x_2^2)(y_1^2+y_2^2) \geq (x_1y_1+x_2y_2)^2 \iff \|x\|\|y\| \geq \|\langle x,y\rangle\| \quad x=(x_1,x_2) \quad y=(y_1,y_2).$$

Again this works for arbitrary dimension! Cauchy-Schwarz inequality.

$$\|x\|\,\|y\| \geq \|\langle x,y\rangle\| \qquad x=(x_1,x_2,\cdots,x_n) \quad y=(y_1,\cdots,y_n)$$

$$\Updownarrow$$

$$\left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right) = \left(\sum_{i=1}^n x_iy_i\right)^2$$

(iv). matrix transpose. Let $T: V\to V$ be a linear operator.

when we define $\langle v,w\rangle = v^*w$ the standard inner product,

$$\langle Tv,w\rangle = \langle v,T^*w\rangle \iff (Tv)^*w = v^*T^*w$$

where $*$ is the transpose when $\mathbb{F}=\mathbb{R}$

hermitian where $\mathbb{F}=\mathbb{C}$.
$\quad$ " conjugate + transpose

T^* is called the adjoint operator of T.

So we have notion of matrix transpose and matrix Hermitian in our language.

$*$ not all remarks above are examinable, depends on the progress of lecture.

Eg. Frobenius inner product: $\langle A,B\rangle = tr(AB^H)$. gives a inner product structure on $M_n(\mathbb{C})$.

Euclidean space: $\mathbb{R}^n$ with an inner product, $\mathbb{F}=\mathbb{R}$.

Inner product gives additional structure to a vector space, and such restrictions give many astonishing consequences.

Thm 6.3. Let V be an inner product space, $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Suppose $\dim V<\infty$. Then

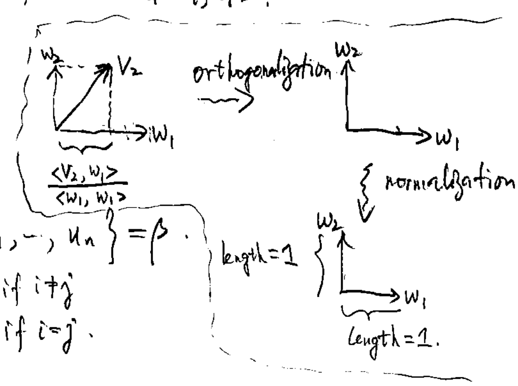
(i). There exist an orthonormal basis, via Gram-Schmidt orthonormalisation process.

$$w_1 = v_1$$
$$w_2 = v_2 - \frac{\langle w_1,v_2\rangle}{\langle w_1,w_1\rangle}w_1$$
$$\vdots$$
$$w_k = v_k - \sum_{i=1}^{k-1}\frac{\langle w_i,v_k\rangle}{\langle w_i,w_i\rangle}w_i$$

orthogonalization

$$u_i = \frac{w_i}{\|w_i\|}, \quad \|w_i\| = \sqrt{\langle w_i,w_i\rangle}.$$

normalization



Any basis $\{v_1,\cdots,v_n\} \xrightarrow{\text{orthonormalization}}$ Orthonormal basis $\{u_1,\cdots,u_n\}=\beta$.

$$\langle u_i,u_j\rangle = \delta_{ij} = \begin{cases}0 & \text{if } i\neq j\\ 1 & \text{if } i=j.\end{cases}$$

orthonormal basis

$$v = \sum_{i=1}^n \underbrace{\langle v,u_i\rangle}_{\text{Scalar}} u_i$$ is the linear combination for any $v\in V$.

$$\Rightarrow ([T]_\beta)_{ij} = \langle u_i, T(u_j)\rangle.$$

(ii). Every subspace $U \subseteq V$ has an <u>orthogonal complement</u>

$$U^{\perp} := \{ v \in V \mid \langle v, u \rangle = 0 \text{ for all } u \in U \}.$$

So that $\quad V = U \oplus U^{\perp}$

This is direct result from (i); the existence of orthonormal basis.

Moreover, ~~one may~~ extend an orthonormal basis in U to a orthonormal basis in V. This is done by an arbitrary extension of basis, and then do Gram-Schmidt. As the basis in U is already orthonormal, this will do nothing to them.

(iii) The adjoint T^* always exists and linear. (T^* always unique if exist).

eigenspace

<u>Spectral Theorem.</u> Let $T: V \to V$ be a linear operator on a <u>finite dimensional</u> v.s. V.

$$\begin{array}{ccc} \text{Diagonalizable} & \xrightarrow[???]{\text{Inner product}} & \text{orthonormally diagonalizable ?} \\ \Updownarrow & & \Updownarrow \\ \exists \text{ eigenbasis} & \xrightarrow{???} & \exists \text{ orthonormal eigenbasis} \end{array}$$

The spectral theorem gives the condition when $T: V \to V$ is orthonormally diagonalizable.

self adjoint

(spectral) Thm 6.4 (i) $\mathbb{F} = \mathbb{R}$. $\quad T = T^* \iff \exists$ orthonormal eigenbasis $\iff T$ is orthonormally diagonalizable.

(ii). $\mathbb{F} = \mathbb{C}$. $\quad TT^* = T^*T \iff \exists$ orthonormal eigenbasis $\iff T$ is unitarily diagonalizable.

normal

Thm 6.5 (i). $T = T^* \iff \langle Tv, v \rangle \in \mathbb{R} \;\; \forall v \in V$. $\quad \mathbb{F} = \mathbb{C}$.

(ii) $T = T^* \Rightarrow$ all eigenvalues are real. $\quad \mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

(iii) $TT^* = T^*T$ and all eigenvalues are real $\Rightarrow T = T^*$ $\quad \mathbb{F} = \mathbb{C}$.

(iv) $TT^* = T^*T \iff \|Tv\| = \|T^*v\| \iff \exists$ orthonormal eigenbasis $\quad \mathbb{F} = \mathbb{C}$.

(v) $TT^* = T^*T \iff \exists$ orthonormal basis β s.t. $[T]_\beta$ is block diagonal with each block is a scalar or $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ where $b > 0$. $\quad \mathbb{F} = \mathbb{R}$.

Isometry Let $T: V \to V$ be a linear operator on a finite dimensional vector space V.

We want to study distance preserving symmetries of space V, this is the notion of isometry. Examples include rotations and reflection. [cf. Artin Algebra].

Def. 6.6. T is an isometry if $\|Tv\| = \|v\|$ for all $v \in V$.

Lemma 6.7. $\|Tv\| = \|v\| \Leftrightarrow \langle Tv, Tw \rangle = \langle v, w \rangle \Leftrightarrow T^*T = Id \Leftrightarrow \begin{cases} A^tA = I & \text{orthogonal } \mathbb{F} = \mathbb{R}. \\ A^HA = I & \text{unitary } \mathbb{F} = \mathbb{C}. \end{cases}$

 $= TT^*$

$\Rightarrow$ all eigenvalues $|\lambda_i| = 1$. where $A = [T]_\beta$ ← orthonormal basis

Thm. 6.8. (i) $\mathbb{F} = \mathbb{R}$. $T^*T = Id \Leftrightarrow \exists$ orthonormal basis β s.t. $[T]_\beta = \begin{pmatrix} I & & & & \\ & -I & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_\ell} \end{pmatrix}$ $\theta_i \neq 0, \pi$.

↙ corresponding to $\lambda, \bar{\lambda}$ in $\mathbb{C}$.

where $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

(ii) $\mathbb{F} = \mathbb{C}$. $T^*T = id \Leftrightarrow \exists$ orthonormal eigenbasis s.t. $[T]_\beta = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ with $|\lambda_i| = 1$.

← eigenvalues

$\Leftrightarrow$ T is unitarily diagonalizable with eigenvalues $|\lambda_i| = 1$. ↖ unitrary.

(note that $T^*T = Id$ implies $T^*T = TT^*$ so spectral thm is applicable).

Q. (Unitary matrix v.s. unitarily diagonalizable:

Unitary matrix means $A^HA = I = AA^H$. Unitarily diagonalizable means $\exists$ unitary P s.t. $A = P^{-1}DP$ for diagonal D. By 6.8 (ii), unitary matrix is unitarily diagonalizable with eigenvalues $|\lambda_i| = 1$.

Q. normal v.s. unitary. Over $\mathbb{C}$.

A is normal if $AA^* = A^*A$. By 6.4 (ii) and 6.8 (ii) we see that

A is normal iff A is unitary with all eigenvalues $|\lambda_i| = 1$.

[Ref. Axler Linear Algebra Done Right.]