Review Functions & Equivalence

- Injective / One-to-One: Every output is only hit once
- Surjective / Onto: Range = Codomain
- Bijective / Invertible: Injective & Surjective $\Leftrightarrow$
  $\Leftrightarrow f: X \to Y,\ f^{-1}: Y \to X \quad f \circ f^{-1} = id_X \ \&\ f^{-1} \circ f = id_Y$
- Restriction: For $X' \subseteq X$ $f: X \to Y$, $f|_{X'}: X' \to Y$ s.t. $f|_{X'}(x') = f(x')$ for $x' \in X'$
- Equivalence Relation $(\sim)$ is symmetric, reflexive, & transitive.
- Equivalence class of $x \in X$ is $E_x = \{y \in X \mid y \sim x\}$
- Equivalence classes partition the set X.

An equivalence relation for example is $x \sim y \Leftrightarrow x \equiv y \bmod N$ for $x, y \in \mathbb{Z}$

Fields

A field is a set w/ 0 & 1 & operations $a+b$, $a-b$, ab, a/b $(b \neq 0)$.

Examples: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subset \mathbb{R}$, $\mathbb{Q}(i) \subset \mathbb{C}$

i) $\mathbb{Q}, \mathbb{R}, \mathbb{C}$

ii) $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subset \mathbb{R}$, $\mathbb{Q}(i) \subset \mathbb{C}$

iii) $\mathbb{Z}_p = \{0, 1, \dots, p-1\}$ w/ $\cdot, +$ mod & p is prime — division inverse of a — Galois Field!

really $\mathbb{Z}_{p^n}$ is also a finite field

A characteristic $n > 0$ is the smallest integer s.t. $\underbrace{1 + 1 + \dots + 1}_{n} = 0$. If there is no such n, then the characteristic is 0.

(i) & (ii) have characteristic 0 & (iii) has characteristic p

Thm: For some field F, $\mathrm{char}(F) = 0 \Leftrightarrow \mathbb{Q} \subseteq F$

Systems of Linear Equations

To solve linear systems we want to simplify the system using operations which do not change the solutions to the system. We simplify by eliminating variables

$$\begin{aligned} 2x_1 - x_2 + x_3 &= 0 \quad -2eq_2 \\ x_1 + 3x_2 + 4x_3 &= 0 \end{aligned} \Rightarrow \begin{aligned} -7x_2 - 7x_3 &= 0 \quad \cdot -1/7 \\ x_1 + 3x_2 + 4x_3 &= 0 \end{aligned}$$

$$\Rightarrow \begin{aligned} x_2 + x_3 &= 0 \\ x_1 + 3x_2 + 4x_3 &= 0 \quad -3eq_1 \end{aligned} \Rightarrow \begin{aligned} x_2 + x_3 &= 0 \\ x_1 \qquad + x_3 &= 0 \end{aligned}$$

$\Rightarrow x_1 = x_2 = -x_3$ or the solution is $(c, c, -c)$

... formalize this system by reducing matrices to REF.

Def: Equivalence of Linear Systems
Two systems of linear eqns are equivalent iff each eqn of the 1st system is a l.c. of the 2nd & vice versa.

This is an equivalence relation.

Also if they have the same RREF

Thm:
Equivalent systems have the same set of solutions.

Pf:
Showing that 2 systems are l.c.'s of each other uses operations which do not change the solution set.

Matrices

$$\begin{cases} 2x_1 - x_2 + x_3 = 0 \\ x_1 + 3x_2 + 4x_3 = 0 \end{cases} \Leftrightarrow \begin{array}{c} \text{(1,1) entry} \\ \begin{bmatrix} 2 & -1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \\ \text{(2,3) entry} \end{array} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We can do the same transformations (elementary row operations) on the augmented coefficient matrix (coefficient matrix w/ solutions

$$\left[\begin{array}{ccc|c} 2 & -1 & 1 & 0 \\ 1 & 3 & 4 & 0 \end{array}\right]$$

Elementary Row Operations

i) Multiply a row by a scalar $c \neq 0$
ii) Multiply a row by a scalar $c \neq 0$ & add it to another row
iii) Switch two rows

These operations are invertible.

Def: Row Equivalence
A matrix A is row equivalent to a matrix B iff B can be obtained from A by a sequence of elementary row operations

This is an equivalence relation

Thm:
Row-equivalent matrices correspond to equivalent (homogeneous w/ RHS = 0) systems of linear eqns, meaning they have the same solutions.

Row Reduced Form

Def:
A $m \times n$ matrix R is row-reduced iff

i) 1st $\neq 0$ entry (called pivot) in each $\neq 0$ row is 1
ii) In a column containing a pivot, all other entries $= 0$.

Def: RREF

R is a row-reduced echelon matrix iff

i) & ii) it is row-reduced
iii) All zero rows are below any non-zero row
iv) If a pivot is below another pivot, then it is to it's right

Any row-reduced matrix can be made a RREF matrix by reordering the rows

Thm:

All matrices are row-equivalent to a row-reduced echelon matrix

Cor: Corollary

All undetermined homogenous system (# eqns < # unknowns) always has non-trivial solutions

Cor:

A homogereous system w/ n eqns & n unknowns ($A\bar{x}=\bar{0}$ where A square) has only the trival solution $\bar{x}=\bar{0}$ $\Leftrightarrow$ A's RREF is equivalent to the identity matrix ($A=I$)

Matrix Multiplication

For $A \in M_{m \times n}$ & $B \in M_{n \times p}$, $AB \in M_{m \times p}$.

To find c_{ij} of $AB = C$

$c_{ij} = \text{row}_i(A) \times \text{col}_j(B)$

Properties:

i) $I_m A = A$ & $A I_n = A$ $\forall A \in M_{m \times n}$
ii) $(AB)C = A(BC)$
iii) $AB \neq BA$ in general

Def: Elementary Matrix

An elementary matrix is a matrix obtained from the identity matrix I by a single row operation

Examples:

i) $\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}$, $\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$ $c \neq 0$

iii) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

ii) $\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$

Let e be an elementary row operation & $E = e(I_m)$ be an elementary matrix.

Then $\forall A \in M_{m \times n}$, $e(A) = EA$

Cor:

B is row-equivalent to $A \Leftrightarrow B = PA$ where P = product of elementary matrices

Examples:

i) $A = \begin{bmatrix} x & z \\ y & t \end{bmatrix} \xrightarrow{cR_1} \begin{bmatrix} cx & cz \\ y & t \end{bmatrix} = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & z \\ y & t \end{bmatrix} = EA$

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{cR_1} E = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$

ii) $A = \begin{bmatrix} x & z \\ y & t \end{bmatrix} \xrightarrow{cR_2 + R_1} \begin{bmatrix} x+cy & z+ct \\ y & t \end{bmatrix} = \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & z \\ y & t \end{bmatrix} = EA$

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{cR_2 + R_1} \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} = E$

iii) $A = \begin{bmatrix} x & z \\ y & t \end{bmatrix} \xrightarrow{R_1 \Leftrightarrow R_2} \begin{bmatrix} y & t \\ x & z \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x & z \\ y & t \end{bmatrix} = EA$

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 \Leftrightarrow R_2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = E$

Invertible Matrices

Consider $n \times n$ (square) matrices

Def:

A left inverse of A is B s.t. $BA = I$

A right inverse of A is C s.t. $AC = I$

If $BA = I = AC$ then $B = C$

Def:

If $\exists A^{-1}$, then A^{-1} is a unique matrix s.t. $AA^{-1} = A^{-1}A = I$

A is called invertible iff $\exists$ such A^{-1}.

Prop:

$(A^{-1})^{-1} = A$

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

Thm:

$R \equiv I$ row-equivalent iff R is a product of elementary operations

pf:

$B = PA$ where P is the product of elementary matrices.

When $A = I$, $R = PI = P$ means R is row-equivalent to I

Elementary Matrices & Their Inverses

Every elementary operation $e(A)$ is invertible, that is $\exists e^{-1}$ s.t. $e(e^{-1}(A)) = e^{-1}(e(A)) = A$

Likewise every elementary matrix E is invertible that is
$\exists E^{-1}$ s.t. $EE^{-1} = E^{-1}E = I$

We can expand this to products of elementary matrices, so
Let $P = E_1E_2 \cdots E_n$
$\exists P^{-1}$ s.t. $PP^{-1} = P^{-1}P = I$
In particular $P^{-1} = E_n^{-1}E_{n-1}^{-1} \cdots E_1^{-1}$. Intuitively, this means you are apply the inverses of the elementary operations in the reverse order to what you originally did.

Vector Spaces

Def:
A vector space over a field F is a set V w/ operations
$+ : V \times V \to V$, $(u,v) \mapsto u+v$, vector addition
$\cdot : F \times V \to V$, $(\alpha, v) \mapsto \alpha v$, scalar multiplication

Subject to the axioms
i) $\bar{u} + \bar{v} = \bar{v} + \bar{u}$ commutativity
ii) $(\bar{u}+\bar{v}) + \bar{w} = \bar{u} + (\bar{v}+\bar{w})$ associativity
iii) $\exists\ \bar{0} \in V$ s.t. $\bar{v} + \bar{0} = \bar{v}$ $\forall \bar{v} \in V$ additive identity
iv) $\forall v \in V$ $\exists -v \in V$ s.t. $v + (-v) = 0$
v) $1v = v$ for $1 \in F$ $\quad p = 2 \to \mathbb{Z}_4$
vi) $(\alpha\beta)v = \alpha(\beta v)$ $\quad n = 2$
vii) $\alpha(u+v) = \alpha u + \alpha v$ viii) $(\alpha+\beta)v = \alpha v + \beta v$

v	$-v$	$(p^n-1)v = 3v$
0	0	0
1	3	3
2	2	$6 \% 4 = 2$
3	1	$9 \% 4 = 1$

Rem: Remark
i) $\alpha\bar{0} = \bar{0} = 0\bar{v}$ $\forall \alpha \in F$, $v \in V$
ii) $(-1)v = -v$ $\quad \alpha \neq 0$ $\quad$ In the context of a finite field $\mathbb{Z}_p$, $-1 = p-1$
iii) $\alpha \neq 0$, $v \neq 0 \Rightarrow \alpha v \neq \bar{0}$

Examples of Vector Spaces:
i) $F^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \middle| \ x_i \in F \right\}$
ii) $F^{m \times n}$ = space of $m \times n$ matrices w/ entries in F ← generalization of above

iii) Let S be a set. ← This is a generalization of earlier ones where the function maps indices to entries

$V = \{ f : S \to F \}$ w/ pointwise operations, that is

$(f+g)(s) = f(s) + g(s)$ & $(cf)(s) = c f(s)$ for $c \in F$ $f, g \in V$

iv) $F[x]$ = polynomials w/ coefficients in F

$F[x]_{\deg \leq n} = \{ \sum_{k=0}^{n} a_k x^k \mid a_k \in F \}$

v) $F \subset K$ - field extension $\Rightarrow$ K is a vector space over F

Def: Linear Combination

A linear combination of $v_1, \dots, v_n \in V$ is

$v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n \in V$ where $c_i \in F$

Let $\text{span}\{v_1, \dots, v_n\}$ = set of all linear combinations of $v_1, \dots, v_n$.

We can extend this to work w/ infinite sets:

$\forall$ set $S \subseteq V$, $\text{span}(S)$ = set of all linear combinations of elements in S

← should be fairly intuitive

Thm:

$\text{span}(S)$ is closed under vector addition & scalar multiplication

$(a_1 v_1 + \dots + a_n v_n) + (b_1 v_1 + \dots + b_n v_n) = (a_1 + b_1) v_1 + \dots + (a_n + b_n) v_n$

$c(a_1 v_1 + \dots + a_n v_n) = (ca_1) v_1 + \dots + (ca_n) v_n$

Subspaces

Def: Sub

A subspace of a vector space V is the subset U which is itself a vector space under the operations of V. (after they've been restricted)

In other words, U is closed under vector addition & scalar multiplication

i) $u + v \in U$ $\forall u, v \in U$

ii) $cv \in U$ $\forall v \in U, c \in F$

Equivalently, $+|_U : U \times U \to U$ & $\cdot|_U : F \times U \to U$ are well-formed.

Thm:

A subset is a subspace $\Leftrightarrow$ it is closed under all linear combinations.

In essence, linear combinations are the "essence" of $+$ & $\cdot$.

Rem:
All subspaces must contain the zero vector (by the axioms) & b/c $0v = \vec{0}$ $\forall v \in V$ where V is any vector space. By the field axioms $0 \in F$ for all fields F.

Examples of Subspaces:

i) $\{0\} \subseteq V$ is a trivial subspace *TODO: he had more notes

ii) $F^{n-1} \subset F^n$ subspace $\Rightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \\ 0 \end{bmatrix}$

iii) $\forall A \in F^{m \times n}$, it's nullspace $null(A) = \{x \in F^n \mid Ax = 0\}$ is a subspace of F^n. This is b/c $Ax + Ay = 0 + 0 = 0$ & $cAx = c0 = 0$

iv) $span(S)$ where $S \subseteq V$ is a subspace.

Thm:
$\forall$ subspaces $W_i \subseteq V$, then $\bigcap W_i$ is a subspace.

It doesn't work for union!

This falls out b/c if you're in W_i, then you are closed under W_i. This means that if you're in the intersection, you are closed under all W_i, meaning you can't escape $\bigcap W_i$.

Rem:
For a subset $S \subseteq V$, $span(S) =$ min subspace of V containing $S = \bigcap_{S \subseteq W} W$

Notation:
$span(S) = \langle S \rangle$

Take subspaces $W_1, W_2 \subset V$ & consider $\langle W_1 \cup W_2 \rangle$

$$\langle W_1 \cup W_2 \rangle = \{\underbrace{c_1 v_1 + \ldots + c_k v_k}_{W_1} + \underbrace{c_{k+1} v_{k+1} + \ldots + c_n v_n}_{W_2}\} \quad v_1, \ldots, v_k \in W_1 \text{ \& } v_{k+1}, \ldots, v_n \in W_2$$
$$= \{w_1 + w_2 \mid w_1 \in W_1 \;\; w_2 \in W_2\}.$$

We use this to define addition of subspaces. (Space of subspaces!)

Def:
Trivially expanded to $\sum W_i$

$W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1 \text{ \& } w_2 \in W_2\} = \langle W_1 \cup W_2 \rangle$

Example:
In $\mathbb{R}^2$, take W_1 & W_2 to be two lines through 0 (where $W_1 \neq W_2$).
Then $W_1 \cap W_2 = \{0\}$ & $W_1 + W_2 = \mathbb{R}^2$

Basis & Dimension

Def: Basis
A subset $B \subset V$ is a basis for V iff $\forall v \in V$ $\exists!$ $v =$ l.c. of elements & B.

This is split in two parts, existence (i) & uniqueness (ii)

i) $\text{span}(B) = V \Leftrightarrow \forall v \in V, v \in \text{span}(B)$ ← literally $V \subseteq B$, but we know $B \subseteq V$ by definition
$\Leftrightarrow B$ spans V

ii) No vector $v \in B$ can be expressed as a l.c. of other vectors in V.

In particular, we know $0 = 0v \in \text{span}(B)$ by the axioms.
Suppose also $0 = \sum c_i v_i$ $v_i \in B$. This is unique iff all coefficients are zero ($\forall c_i = 0$). This means uniqueness is satisfied iff $\forall c_i = 0$ is the only solution to $\sum c_i v_i = 0$

Def: Linearly Independent (l.i.)
Vectors are linearly independent iff there are no non-trivial solution to $c_1v_1 + \dots + c_nv_n = 0$. (Trivial where $c_1 = \dots = c_n = 0$)

A basis B of a vector space V is a linearly independent spanning set of V.

Def: Linearly Dependent
Not l.i. or there exist non-trivial solution to $c_1v_1 + \dots + c_nv_n = 0$

Def: Dimension
The dimension of v.s. V is the cardinality of some basis B of V.

$\dim(V) = |B|$

We say $\dim(V)$ is infinite if there is no finite basis.

We need to show that the dimension of V doesn't depend on the specific basis B chosen.

Example:
$\mathbb{R}^n$ has a standard basis $B = \{ e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \}$

Standard basis of a v.s. V where $\dim(V) = n$ is $\{e_1, \dots, e_n\}$

b/c $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1e_1 + \dots + x_ne_n$, so every $v \in \mathbb{R}^n$ is a unique l.c. of B.

Example:
The space of polynomials $F[x]$ has a standard basis $B = \{1, x, \dots, x^n\}$.
$\dim(F[x]) = \infty$.

Example:

$\mathbb{R} \subset \mathbb{C}$ subfield $\Rightarrow$ $\mathbb{C}$ is a v.s. over $\mathbb{R}$

$\dim_{\mathbb{C}} \mathbb{C} = 1 \quad B = \{1\}$

$\dim_{\mathbb{R}} \mathbb{C} = 2 \quad B = \{1, i\}$

$\dim_{\mathbb{Q}} \mathbb{C} = \infty$

Rem:

The basis of the trivial space $V = \{0\}$ is $B = \emptyset$.
By definition, $\dim(V) = \dim(\{0\}) = |B| = |\emptyset| = 0$.

Now it's time to prove that choice of basis doesn't matter for dimension. (We'll actually go more general w/ the fundamental theorem of ______)

Thm:

Let V be a vector space spanned by m vectors.
Any set of $n > m$ vectors is linearly dependent

Pf:

$V = \text{span}\{w_1, \ldots, w_m\}$. Consider $\{v_1, \ldots, v_n\}$.
Since $v_j \in V$, $v_j = \sum_{i=1}^{m} a_{ij} w_i \quad \forall j \quad 1 \le j \le n$.

For $x_1, \ldots, x_n \in F$

TODO: Take from notes

Cor:

If V is a finite v.s., then any two basis has the same cardinality.

Pf:

Assume v.s. V has basis $B_1 = \{w_1, \ldots, w_m\}$.
Take any other basis B_2.
If $|B_2| > m$, then B is linearly dependent, so $|B_2| \le m$.

Swapping the roles of B_1 & B_2 gives you $|B_1| \leq m$. ← Also if $|B_2| < m$, then B_2 doesn't span V.

Therefore $|B_1| = |B_2|$

Thm/Cor:
Let $\dim(V) = m$ where m is finite.
Let $S \subseteq V$, $|S| = n$.

If $n > m \Rightarrow S$ is linearly dependent (from above)

If $n < m \Rightarrow S$ is not a spanning set of V. $\langle S \rangle \neq V$.

This means $\dim(V)$ = max # of l.i. vectors = min # vectors in a spanning set.

Thm:
Let S be a l.i. subset of v.s. V where $v \in V$ & $v \notin \langle S \rangle$.
Then $S \cup \{v\}$ is linearly independent.

Pf:
Suppose $x_0 v + \sum_i x_i v_i = 0$ for $v_i \in S$.
Suppose $x_0 \neq 0$ for contradiction. Then

$$v = -\frac{1}{x_0} \sum_i x_i v_i \in \langle S \rangle$$

This contradicts $v \notin S$, so $x_0 = 0$ must be true. Since S is l.i.

$$\sum_i x_i v_i = 0$$

only when $\forall x_i = 0$.

By repeated use of the above theorem, we can prove.

Thm:
Let V be a finite dimensional v.s.
Any l.i. subset of V can be extended to a basis.

Cor:
Let W be a subspace of a finite dimensional v.s. V.

Any basis of W can be extended to a basis for V (since the basis of W is a l.i. subset of V).

Cor:
$\forall$ subspaces $W \subseteq V$, $\dim(W) \leq \dim(V)$ w/ $\dim(W) = \dim(V)$ iff $W = V$.

Thm:
Let V be a v.s. where $\dim(V) = n < \infty$.
Any n l.i. vectors form a basis. they imply each other.
Any n spanning vectors form a basis.

This falls from the theorem that we can add vectors to l.i. set to get a basis & you can remove vectors from a spanning set to get a basis.

Example:

Consider $v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $v_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \in \mathbb{R}^3$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = x_1 v_1 + x_2 v_2 + x_3 v_3 = \begin{bmatrix} x_1 + x_2 + x_3 \\ 2x_1 + x_2 + 3x_3 \\ 2x_1 \quad + 2x_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

This is true for some b & x iff $Ax = 0$ has only one solution x, which is true iff A is invertible, which is true iff $\det(A) = 0$.

Thm:

Let W_1 & W_2 be subspaces of V where $\dim(V) < \infty$.

Then $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$

Recall: $W_1 + W_2 = \langle W_1 \cup W_2 \rangle$

Like inclusion exclusion principle for sets

pf:

Let $\hat{B}$ be a basis for $W_1 \cap W_2$.
Since $W_1 \subseteq W_1 \cap W_2$, we can extend $\hat{B}$ to be a basis B_1 for W_1.
Likewise we get B_2 basis for W_2.

We claim $B_1 \cup B_2$ is a basis for $W_1 + W_2$. In that case
$|B_1 \cup B_2| = |B_1| + |B_2| - |B_1 \cap B_2 = \hat{B}| = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$
$\Leftrightarrow \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$

Now we show $B_1 \cup B_2$ is a basis for $W_1 + W_2$.
Let

$\hat{B} = \{v_1, \dots, v_k\}$

$B_1 = \{v_1, \dots, v_k, u_1, \dots, u_m\}$

$B_2 = \{v_1, \dots, v_k, w_1, \dots, w_n\}$

$B_1 \cup B_2 = \{v_1, \dots, v_k, u_1, \dots, u_m, w_1, \dots, w_n\}$

$\forall$ vector $= x + y \quad x \in V_1, y \in W_2$
$x = \sum a_i v_i + \sum b_i u_i$
$y = \sum c_i v_i + \sum d_i w_i$
$x + y = \sum (a_i + c_i) v_i + \sum b_i u_i + \sum d_i w_i$
Therefore, $B_1 \cup B_2$ spans $W_1 + W_2$.

$\sum a_i v_i + \sum b_i v_i + \sum c_i w_i = 0$

$\underbrace{\sum a_i v_i + \sum b_i v_i}_{\in W_1} = \underbrace{-\sum c_i w_i}_{\in W_2} \in W_1 \cap W_2$

So $\forall_i\, b_i = 0$ & $\forall c_i = 0$.

Since $\hat{B}$ is a basis for $W_1 \cap W_2$, $\sum a_i v_i = 0$ iff $\forall_i a_i = 0$.
Therefore $B_1 \cup B_2$ is l.i.

Thus $B_1 \cup B_2$ is a basis for $W_1 + W_2$, so
$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$

Coordinates

Ex:

Let $V = \mathbb{R}^2$ w/ standard basis $B = \{e_1 = \begin{bmatrix}1\\0\end{bmatrix}, e_2 = \begin{bmatrix}0\\1\end{bmatrix}\}$.
$\forall v \in V$, $x e_1 + y e_2 = \begin{bmatrix}x\\y\end{bmatrix} = v$.

Let's consider another basis $B' = \{v_1, v_2\}$.
$v = x v_1 + y v_2$.

This gives you another set of coordinates x & y.

This is the idea of using any basis to create a coordinate system for your v.s. Really, you need an ordered basis b/c otherwise you wouldn't know which coordinate goes to which vector $v \in B$.

Thm: Coordinate Vector

Let V be a v.s. over field F where $\dim(V) = n$.

Fix some ordered basis $B = \{v_1, \dots, v_n\}$ for V.
Then $\forall v \in V$ $\exists!$ $v = x_1 v_1 + \dots + x_n v_n$ where $\forall_i x_i \in F$

This means there is some unique coordinate vector w.r.t. B

$$[v]_B = \begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix} \in F^n$$

We say $x_1, \dots, x_n$ are the coordinates of v w.r.t. B

Recall: Isomorphisms are bijections b/w vector spaces that preserves the properties of + & ·

Rem:

The map $V \to F^n$, $v \mapsto [v]_B$ is a bijection b/c you can freely go back & forth. This map is an isomorphism.

This means all n-dimension vector spaces V over F are isomorphic w/ F^n.

Change of Basis

Let $B = \{v_1, \dots, v_n\}$ & $B' = \{v_1', \dots, v_n'\}$ be 2 ordered bases for V.

$$(i)\ v = \sum_i x_i v_i \iff [v]_B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = X$$

$$(ii)\ v = \sum_j x_j' v_j' \iff [v]_{B'} = \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} = X'$$

We want algorithm that can transform b/w $[v]_B$ & $[v]_{B'}$ for any basis B & B' of V & any vector $v \in V$. ~~We do this by defining a linear transformation from B to B'.~~ For this it is sufficient to write every element of B' as a linear combination of B. We can then be clever & make this a linear transformation from B' to B & thus as a matrix $P \in F^{n \times n}$ such that $\forall v \in V$ $[v]_B = P[v]_{B'}$. We call P the transition matrix.

Since $\forall_j\ v_j' \in V$, we know

$$(iii)\ \forall_j\ v_j' = \sum_i p_{ij} v_i \text{ for some scalar } p_{ij} \in F \iff v_j' = \begin{bmatrix} p_{1,j} \\ \vdots \\ p_{n,j} \end{bmatrix} \text{ j-th column of } P$$

We then plug (iii) into (ii)

$$v = \sum_j x_j' v_j' = \sum_j \left(x_j' \sum_i (p_{ij} v_i) \right) \iff$$

TODO: Look at lecture notes

Conversely, we have the following theorem.

Thm:

Let $B = \{v_1, \dots, v_n\}$ be a basis for V.
Let $P \in F^{n\times n}$ be an invertible matrix. $P = (p_{ij})$.

$v_j' = \sum_{i=1}^{n} p_{ij} v_i \quad \forall j \ 1 \le j \le n$

Then $B' = \{v_1', \dots, v_n'\}$ is a basis.

Example:

Let $V = \mathbb{R}^3$, $v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $v_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$

Let $x \in V$

$$V = x_1 v_1 + x_2 v_2 + x_3 v_3$$

$$= [v_1 \ v_2 \ v_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= AX$$

To solve $AX = v$, we make an augmented matrix $[A|v]$ & row-reduce to $[A|w]$.

$AX = v \Leftrightarrow RX = w$. If $R = I$, then solution is $v = w$.

In general, we can make our work more general by solving $[A|I]$ into $[I|A^{-1}]$. That way we can do

$$AX = v \Leftrightarrow X = A^{-1}v \Leftrightarrow X = A^{-1}$$

In this way we transform any vector V into a coordinate vector w.r.t. $B = \{v_1, v_2, v_3\}$.

In particular $[e_1]_B$ is the 1st col of A^{-1}, $[e_2]_B$ the 2nd, $[e_3]_B$ the 3rd.

#Functions/ Mappings TODO: Move to MA 407

Def:

Let $f: A \to B$ be a function from A to B.
$A = \text{Dom}(f) =$ domain of f
$B = \text{Codom}(f) =$ codomain of f
$\text{Rng}(f) \subseteq \text{Codom}(f)$ is the range of f.

Examples:

Let $\mathbb{Q}^+ = \{\frac{m}{n} : m, n \in \mathbb{Z}\}$ be the set of positive rational numbers.

We define a function $f: \mathbb{N} \times \mathbb{N} \to \mathbb{Q}^+$ as
$$f((m,n)) = m/n$$

Suppose $(m_1, n_1) = (m_2, n_2)$. We show $f((m_1, n_1)) = f((m_2, n_2))$
$$f((m_1, n_1)) = \frac{m_1}{n_1} = \frac{m_2}{n_2} = f((m_2, n_2))$$

Now define $g: \mathbb{Q}^+ \to \mathbb{N} \times \mathbb{N}$, let's define $g(\frac{m}{n}) = (m, n)$.
Is g a function? No. g tries to be f^{-1} but f^{-1} is not injective.

We show that $a = b \Rightarrow g(a) = g(b)$ does not hold.
$$3/2 = 6/4 \text{ but } g(3/2) \neq g(6/4) \Leftrightarrow (3,2) \neq (6,4)$$

Def: Image

Let $f: A \to B$.

Let $C \subseteq A$.
$f(C) = \{f(c) : c \in C\}$ is called the image of C under f.

Let $D \subseteq B$. We're not assuming f^{-1} exists!
$f^{-1}(D) = \{a \in A : f(a) \in D\}$ is called the inverse image of D under f.

The image of A under f is $\text{Rng}(f)$ TODO: are these right?

The inverse image of B under f is $\text{Dom}(f)$

Row Space

Def: Row Space
For $A \in F^{m \times n}$ over field F, its <u>row space</u> row(A) is
row(A) = span({row vectors of A}) $\subseteq F^{m \times 1}$ (subspace of possible row vectors)

Rem:
Let $A \in F^{m \times n}$. The transpose matrix $A^T \in F^{n \times m}$ where
$(A^T)_{ij} = A_{ji}$ (switch rows & col's)

In this case row(A) = col(A^T) & col(A) = row(A^T)

Def: Rank
The rank of $A \in F^{m \times n}$ is rank(A) = dim(row(A)), that is the dimension of the row space.

you can actually replace row & col & it is equivalent, meaning rank(A) $\leq$ min(m, n)

Rem:
Since row(A) $\subseteq F^{m \times 1}$, rank(A) $\leq$ m.

Note:
For $A \in F^{n \times n}$, A is invertible iff rank(A) = n. (Proof later)

Thm:
i) If $A, B \in F^{m \times n}$ are row equivalent, then row(A) = row(B).

ii) If R is a row-reduced matrix, then its non-zero rows form a basis for row(R).

Pf:
i) Let $A \in F^{m \times n}$. $A \xrightarrow{e} A_1 = e(A)$ where e is elem. operation. All rows of A_1 are l.c.'s of A. Thus, the rows are in row(A), so row(A_1) $\subseteq$ row(A).
We do the same in the reverse direction (elem. ops. are invertible) to show row(A) $\subseteq$ row(A_1).
Thus row(A) = row(A_1). We can apply this transitively to get the theorem.

ii) Let $\rho_1, \ldots, \rho_\ell$ be the non-zero rows of R.
row(R) = span{$\rho_1, \ldots, \rho_\ell$} by def.
We want to show l.i.

Basically, eliminate the redundant vectors

$\forall \rho_i = (0, 0, \ldots, \underset{\text{pivot}}{1}, *, *, \ldots, *)$

$$\begin{array}{cc} R = \begin{bmatrix} 1 & & 0 & & \\ 0 & \ldots & 1 & \ldots & * \\ & & 0 & & \end{bmatrix} & \leftarrow \rho_i \\ \quad\quad\quad\quad k\text{-th column} & \end{array}$$

$$\sum_{j=1}^{\ell} c_i \rho_i = (*, \ldots, *, \underset{k\text{-th entry}}{c_i}, *, \ldots, *)$$

b/c it is c_i times 1 in the k-th column

This sum is only zero if all c_i are zero. Thus there are no non-trivial solutions & R is l.i.

Example:

Let

$v_1 = (1, 1, 1, 1)$ ← could write vertically
$v_2 = (2, 3, -1, -1)$
$v_3 = (3, 2, 1, 1)$
$v_4 = (3, 6, -1, -1)$

be vectors in $\mathbb{R}^4$.

We form a basis for V span$\{v_1, v_2, v_3, v_4\}$ by defining a matrix

$$A = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

might be proper subset

Note here row(A) = span$\{v_1, v_2, v_3, v_4\} = V \subseteq \mathbb{R}^4$

To find a basis for V, we row-reduce A.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & -1 & -1 \\ 3 & 2 & 1 & 1 \\ 3 & 6 & -1 & -1 \end{bmatrix} \begin{matrix} \\ R_2 - 2R_1 \\ R_3 - 3R_1 \\ R_4 - 3R_1 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -3 & -3 \\ 0 & -1 & -2 & -2 \\ 0 & 3 & -4 & -4 \end{bmatrix} \begin{matrix} R_1 - R_2 \\ \\ R_3 + R_2 \\ R_4 - 3R_2 \end{matrix} \begin{bmatrix} 1 & 0 & 4 & 4 \\ 0 & 1 & -3 & -3 \\ 0 & 0 & -5 & -5 \\ 0 & 0 & 5 & 5 \end{bmatrix} R_3 \div -5$$

$$\begin{bmatrix} 1 & 0 & 4 & 4 \\ 0 & 1 & -3 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 5 & 5 \end{bmatrix} \begin{matrix} R_1 - 4R_3 \\ R_2 + 3R_3 \\ \\ R_4 - 5R_3 \end{matrix} \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix} \end{array}.$$

By the theorem, a basis for V is

$w_1 = (1, 0, 0, 0)$
$w_2 = (0, 1, 0, 0)$ } non-zero entries of the row-reduced matrix
$w_3 = (0, 0, 1, 1)$

By def, $\{w_1, w_2, w_3\}$ spans V & trivially they are l.i.

We can use $\{w_1, w_2, w_3\}$ to rewrite V as

$V = \{(x_1, x_2, x_3, x_4) \mid x_3 - x_4 = 0\}$

Linear Transformations

Linear transformations are functions/maps that preserve linearity. That is they preserve the properties of linear combinations in a sense.

Def:

Let V & W be vector spaces over F.

A linear transformation/map from V to W is a function $T: V \to W$ s.t.

$T: V \to W$ s.t.

$T(v_1 + v_2) = T(v_1) + T(v_2)$ $\forall v_1, v_2 \in V$

$T(cv_1) = cT(v_1)$ $\forall v_1 \in V$ $c \in F$.

When $V = W$, we say T is a linear operator.

Equivalently, image of a linear combination ~~is the linear~~ combination of images

$$T(\sum c_i v_i) = \sum c_i T(v_i)$$

(We use Hom b/c linear transformations are homomorphisms on vector spaces)

Notation:

$L(V, W) = \text{Hom}(V, W) = \text{Hom}_F(V, W)$ is the set of linear transformations from V to W over field F.

$L(V, V) = \text{End}(V) = \text{End}_F(V)$ is the set of all linear maps from V to V over field F, or equivalently an endomorphism on V.

These are both actually vector spaces themselves.

Examples: Linear Maps

i) $\frac{d}{dx}$ (derivative) is a linear map. ← & actually a linear operator

$(f+g)' = f' + g'$ & $(cf)' = cf'$ $c \in F$ $f, g \in C^\infty(F)$ or $f, g \in F[x]$ or $f, g \in F[x]_{\deg \leq n}$

continuous functions on field F

polynomials

ii) ODE $a_n(x)\left(\frac{d}{dx}\right)^n f + a_{n-1}(x)\left(\frac{d}{dx}\right)^{n-1} f + \dots + a_1(x)\frac{d}{dx} f + a_0(x) f = T(f) = 0$

where T is a linear operator is a linear map. (extension of (i))

iii) Definite integrals are linear maps. ← Indefinite integrals aren't b/c they aren't functions b/c of $+C$

$$(T(f))(x) = \int_0^x f(y)\,dy$$

iv) Let $A \in F^{m \times n}$, $X \in F^n$. $T: F^n \to F^m$ $T(X) = AX$ is a linear map

$$T(X+Y) = A(X+Y) = AX + AY \quad \& \quad T(cX) = A(cX) = c(AX) \quad \forall X, Y \in F^n, c \in F$$

We will later show F^n is the canonical form of any v.s. V where $n = \dim(V)$. Likewise $F^{m \times n}$ is the canonical form of any linear transformation $T \in L(V, W)$ where $n = \dim(V)$ & $m = \dim(W)$. We do this by defining a bijection b/w them that is actually itself a linear map.

<u>Thm:</u>

Let V & W be vector spaces & B be a basis for V.

For any map $f: B \to W$, there is a unique linear map $T: V \to W$ that is an extension of f (i.e. $T|_B = f$). In other words, every linear transformation is uniquely determined by its action on the basis.

<u>Pf:</u>

Let $v \in V$. Since B is a basis,
$\exists! \; v = \sum x_i v_i \quad x_i \in F, \; v_i \in B$.

We define the linear map extension as T (that is $T|_B = f$)
$T(v) = T(\sum x_i v_i) = \sum x_i T(v_i) = \sum x_i f(v_i)$
We have thus shown T exists & is unique.

In order to apply T, we break the vector into coordinates/ a l.c of the basis, apply T (or f) to the basis, & then do the l.c.

<u>Rem:</u> $c \in F$

Let $v, w \in V$ where $B = \{b_1, \dots, b_n\}$ is an ordered basis for V.
Then
$$[v+w]_B = [v]_B + [w]_B \quad \& \quad [cv]_B = c[v]_B$$

This means the map $V \to F^n$, $v \to [v]_B$ is a <u>linear map</u>.
In fact, it is a bijection.

Since $v \to [v]_B$ is a bijective linear map, it is an isomorphism.

<u>Def:</u> Isomorphism
An <u>isomorphism</u> (of vector spaces) is a bijective linear map. That means we can freely switch b/w the two spaces.

That is V & F^n are "equivalent" in that we can rename them to be equivalent.

We write $V \cong W$ iff $\exists$ isomorphism $V \to W$.

<u>Rem:</u>
i) $\cong$ ($\exists$ isomorphism) is an equivalence relation

ii) $V \cong W \Leftrightarrow \dim(V) = \dim(W)$ ← yes! both ways
This is b/c (as shown above) $V \cong F^n$, $n = \dim(V)$.

How do we identify an isomorphism? It needs to be a linear map, surjective, & injective.

Def: Image of Map

Let $T: V \to W$ be a linear map. The image of T (or range) is

$$\mathrm{Im}(T) = T(V) = \{T(v) : v \in V\} \subseteq W.$$

also called range

T is surjective iff $T(V) = W$.

Rem:

$\forall T \in \mathrm{Hom}(V,W)$ (or $L(V,W)$), we have $T(0) = 0$. If this wasn't the case, T wouldn't be a linear map.

Thm:

Let $T \in \mathrm{Hom}(V,W)$. T is injective iff $T(v) = 0$ only when $v = 0$.

Pf:

Let $v_1, v_2 \in V$ where $v_1 \neq v_2$ (so $v_1 - v_2 \neq 0$).

$$T(v_1) = T(v_2) \Leftrightarrow T(v_1) - T(v_2) = 0 \Leftrightarrow T(v_1 - v_2) = 0$$

Thus $T(v') = 0$ where $v' \neq 0$ implies T is injective & vice versa.

Def:

The kernel (or null space) of $T: V \to W$ is

$$\mathrm{Ker}(T) = \mathrm{Null}(T) = \{v \in V : T(v) = 0\} \subseteq V.$$

Prop:

$\forall T \in \mathrm{Hom}(V, W)$, we have

i) $\mathrm{Ker}(T) \subseteq V$ (subspace)

ii) $\mathrm{Img}(T) \subseteq V$ (subspace)

iii) T is injective iff $\mathrm{Ker}(T) = \{0\}$. ← from above proof

iv) T is surjective iff $\mathrm{Im}(T) = W$. ← def of surjective

We can thus measure how close a linear map is to being injective ($\dim(\mathrm{Ker}(T)) = 0$) & how close to being surjective ($\dim(\mathrm{Img}(T)) = \dim(W)$)

Def:

Let $T: V \to W$ be a linear map.
The nullity of T is $\mathrm{nullity}(T) = \dim(\mathrm{Ker}(T))$
The rank of T is $\mathrm{rank}(T) = \dim(\mathrm{Im}(T))$

Note Ae_i is i-th column of A

Example:

Let $A \in F^{m \times n}$. We define a linear map $T_A: F^n \to F^m$ is $T_A(X) = AX$. Then

$\mathrm{Ker}(A) = \{X \in F^n \mid AX = 0\}$ = null space of A
$\mathrm{Im}(A) = \{AX \mid X \in F^n\}$ = column space of A

Thm: Rank-Nullity Theorem

Let V & W be v.s. where $\dim(V) < \infty$.
Let $T \in \mathrm{Hom}(V, W)$.
Then

$$\underbrace{\mathrm{rank}(T)}_{\dim(\mathrm{Im}(T))} + \underbrace{\mathrm{nullity}(T)}_{\dim(\mathrm{Ker}(T))} = \dim(V)$$

Pf:

We know $\mathrm{Ker}(T) \subseteq V$ subspace.
Since $\dim(V) < \infty$, we know $\dim(\mathrm{Ker}(T)) = \mathrm{nullity}(T) < \infty$.

Pick a basis $\{v_1, \ldots, v_k\}$ be a basis for $\mathrm{Ker}(T)$
& extend it to be a basis for V $\{v_1, \ldots, v_k, \ldots, v_n\}$.
Since $\mathrm{Ker}(T) \subseteq V$, $k \leq n$.

$\forall v \in V \; \exists! \; v = \sum_{i=1}^{n} x_i v_i \quad (x_i \in F)$
$T(v) = T\left(\sum_{i=1}^{n} x_i v_i\right) = \sum_{i=1}^{n} x_i T(v_i)$

We split this sum to be using $v_1, \ldots, v_k$ & $v_{k+1}, \ldots, v_n$.

$$T(v) = \sum_{i=1}^{k} x_i T(v_i) + \sum_{i=k+1}^{n} x_i T(v_i) = 0 + \sum_{i=k+1}^{n} x_i T(v_i)$$

$v_1, \ldots, v_k$ basis for $\mathrm{Ker}(T)$

Thus $\mathrm{Im}(T) = \{T(v) \mid v \in V\} = \mathrm{span}\{v_{k+1}, \ldots, v_n\}$

We show $\{v_{k+1}, \ldots, v_n\}$ is l.i., meaning it's a basis for $\mathrm{Im}(T)$.

Suppose $\sum_{i=k+1}^{n} x_i T(v_i) = 0 = T\left(\sum_{i=k+1}^{n} x_i v_i\right)$ for some x_i.

$\Rightarrow \sum_{i=k+1}^{n} x_i v_i \in \mathrm{Ker}(T)$

This would mean $v_1, \ldots, v_k, \ldots, v_n$ would be linearly dependent, which violates the assumption that they are a basis.

This implies $\{T(v_{k+1}), \ldots, T(v_n)\}$ is a basis for $\mathrm{Img}(T)$, meaning

All this means
$\dim(\mathrm{Ker}(T)) = \mathrm{nullity}(T) = k$
$\dim(\mathrm{Im}(T)) = \mathrm{rank}(T) = n-k$
$\dim(V) = n$
This is the equality we want. $\square$

Cor:
$\forall A \in F^{m \times n}$, rowrank$(A)$ = column rank(A).

Pf:

Let $T_A : F^n \to F^m$, $T_A(X) = AX$ be a linear map.

nullity(A) = nullity(T_A) & columnrank(A) = rank(T_A) is known from earlier work.

n = nullity(T_A) + rank(T_A) by the rank-nullity theorem.

On the other hand, nullity(A) & rowrank(A) aren't changed by elementary row operations.

We can apply a series of row operations to row-reduce A, we can see

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & * & * \\ 0 & 1 & 0 & * & * \\ 0 & 0 & 1 & * & * \end{bmatrix}}_{n} \Bigg\} m$$

&

nullity(A) = # free variables = n − # pivots
rowrank(A) = # non-zero rows = # pivots = columnrank(T)

We have thus shown row rank & column rank are equivalent. We call it simply rank.

Linear Transformations $\simeq$ Matrices

If we have $A \in F^{m \times n}$, we can define $T_A : F^n \to F^m$ $T_A(X) = AX$ easily. What about the reverse direction?

Given a linear map $T : F^n \to F^m$, we find $A \in F^{m \times n}$ ~~such that $T(X) = AX$.~~

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i e_i \in F^n$$

standard basis vectors

This process can be applied to any basis, not just the standard basis.

$$T(X) = T\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i T(e_i) = \sum_{i=1}^{n} A_i x_i = AX$$

where $A = [A_1 \dots A_n] \in F^{m \times n}$ & $A_i = T(e_i)$. That is, create a matrix where the columns are the results of T applied to the basis vectors

We have thus shown any linear transformation can be made into a matrix & vice versa.

Algebra of Linear Maps

Here we will show the set of linear maps is a vector space, making them nice, consistent, & meta.

Def:
Let V, W be vector spaces.
Let $T, U \in \mathrm{Hom}(V, W)$ = set of linear maps $T : V \to W$ & $c \in F$.

$(T + U)(v) = T(v) + U(v)$
$(cT)(v) = cT(v)$

We won't show it now but here vector addition & scalar multiplication follow the vector space axioms so
Hom(V, W) is a vector space.

Thm:

Let V, W, Z be vector spaces.
Let $T \in Hom(V, W)$ & $S \in Hom(W, Z)$.
The composition $S \circ T = ST : V \to Z$ is a linear map (i.e. $ST \in Hom(V, Z)$)

Pf:

Let $u, v \in V$ & $c \in F$.

$$
\begin{aligned}
(ST)(u+v) &= S(T(u+v)) \\
&= S(T(u) + T(v)) \\
&= S(T(u)) + S(T(v)) \\
&= (ST)(u) + (ST)(v)
\end{aligned}
$$

$$
\begin{aligned}
(ST)(cv) &= S(T(cv)) \\
&= S(cT(v)) \\
&= cS(T(v)) \\
&= c(ST)(v) \quad \square
\end{aligned}
$$

Def:

The identity operator $I : V \to V$ on vector space V
$I(v) = v \ \forall v \in V$

endomorphism

$I \in Hom(V, V) = Endo(V)$ trivially. (not proved here)

Prop:

i) $I_W T = T$ & $T I_V = T$

ii) $S(T_1 + T_2) = ST_1 + ST_2$ & $(S_1 + S_2)T = S_1 T + S_2 T$

iii) $c(ST) = (cS)T = S(cT) \ \forall c \in F$

iv) $T_1(T_2 T_3) = (T_1 T_2) T_3$

All these properties combined means that Endo(V) is an associative algebra over field F. (this idea isn't important here)

Recall:

A function $T: V \to W$ is invertible iff $\exists! S = T^{-1}: W \to V$ such that

$ST = I_V$ & $TS = I_W$ (or $ST = TS = I$ as abuse of notation or generic composition)

We already know T is invertible iff T is bijective.

Thm:

If $T \in Hom(V, W)$ & T is invertible, then $T^{-1} \in Hom(W, V)$ ← T^{-1} linear transformation

Def:

An <u>isomorphism</u> $T: V \to W$ is an invertible/bijective linear map.

Prop:

i) The inverse of an isomorphism is an isomorphism. (symmetric)

ii) The composition of an isomorphism is an isomorphism (transitive)

iii) All vector spaces are isomorphic to themselves by the identity (reflexive)

Def:

We say <u>V isomorphic to W</u> iff $\exists T: V \to W$ s.t. T is an isomorphism.

By the above properties, we can see this defines an equivalence relation.

Thus, if V is isomorphic to W, we write

$V \cong W$.

Def:

Let $T \in Hom(V, W)$.

We say T is non-singular iff $Ker(T) = Null(T) = \{0\}$, that is its kernel is trivial.

T non-singular
$\Leftrightarrow$ T injective
$\Leftrightarrow$ $nullity(T) = 0$

Prop:

i) T injective
$\Leftrightarrow Ker(T) = \{0\}$
$\Leftrightarrow T$ is non-singular
$\Leftrightarrow nullity(T) = 0$.

ii) T surjective
$\Leftrightarrow Img(T) = W$
$\Leftrightarrow rank(T) = dim(W)$

Recall:

By the rank-nullity theorem, $nullity(T) + rank(T) = dim(V)$

Thm:

Let $T \in Hom(V, W)$ where $dim(V) = dim(W) < \infty$.

T isomorphism
$\Leftrightarrow \exists T^{-1}$
$\Leftrightarrow T$ bijective
$\Leftrightarrow T$ injective $\Leftrightarrow T$ non-singular $\Leftrightarrow$ nullity$(T) = 0$
$\Leftrightarrow T$ surjective $\Leftrightarrow Img(T) = W \Leftrightarrow$ rank$(T) = W$

Thm:

Let $T \in Hom(V, W)$.

T is non-singular iff T sends any linearly independent subset of V to a linear independent subset W.

Pf: $\Rightarrow$

Assume T is non-singular.
Let $A \subseteq V$ be a linearly-independent subset of V.

Suppose for contradiction that $T(A) = B$ is lin. dep.
This means for some
$$c_1 T(v_1) + \dots + c_k T(v_k) = 0$$
for some non-trivial $c_i \in F$, $v_i \in A$.

$$\Rightarrow T(c_1 v_1 + \dots + c_k v_k) = 0$$
$$\Rightarrow c_1 v_1 + \dots + c_k v_k = 0$$
Since A is linearly independent
$$c_1 = \dots = c_k = 0.$$
This contradicts our assumption that c_i is trivial, so
$B = T(A)$ is lin. ind.

$\square$

Pf: $\Leftarrow$
Take $v \in V$ $v \neq 0$. Thus $\{v\}$ is lin. ind.
This means $\{T(v)\}$ is linearly independent.
This means $T(v) \neq 0$, so
$$Ker(T) = Null(T) = 0.$$

Thm:
$T \in Hom(V, W)$ is an isomorphism iff T sends a basis to a basis.

This is an extension of the above theorem.

Coro:
$V \cong W$ iff $\dim(V) = \dim(W)$

Recall:
A lin. map $T: V \to W$, where V has basis B, is uniquely determined by the restriction of T w/ B $T|_B$. ← You pick $T|_B : B \to W$ arbitrarily

Linear Transformation & Matrixes

This is review & extension of earlier work.

Let $B = \{v_1, \ldots, v_n\}$ be an ordered basis for V.

Since B is a basis
$\forall v \in V \quad \exists!\, v = \sum_{j=1} x_j v_j$

We can rewrite this as a coordinate vector

$$[v]_B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in F^n$$

We earlier showed the map $C_B : V \to F^n$ s.t.
$C_B(v) \doteq [v]_B$
is an isomorphism.

Let $B' = \{w_1, \ldots, w_m\}$ be an ordered basis for W.
Let $T \in \mathrm{Hom}(V, W)$.

$$\begin{tikzcd} V \arrow[r, leftrightarrow, "T"] \arrow[d, leftrightarrow, "C_B"'] & W \arrow[d, leftrightarrow, "C_{B'}"] \\ F^n \arrow[r, leftrightarrow, "T_A"', "C_{B'} \circ T \circ C_B^{-1}"'] & F^m \end{tikzcd}$$

View of isomormphisms b/w vector spaces

This means $T_A(X) = AX$ for some matrix $A \in F^{m \times n}$.

For $v \in V$,

$\boxed{[T(v)]_{B'} = A[v]_B}$ ← summarizes graph as equality

Summarizing everything,

$$[v]_B = X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \iff v = \sum_{j=1}^{n} x_j v_j$$

Then

$$\underset{\in W}{T(v)} = T\Big(\sum_{j=1}^{n} x_j v_j\Big) = \sum_{j=1}^{n} x_j \underset{\in W}{T(v_j)} \implies [T(v)]_{B'} = \sum_{j=1}^{n} x_j \underbrace{[T v_j]_{B'}}_{A_j \in F^m} = [A_1 \cdots A_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = AX$$

where $A \in F^{m \times n}$ where $A_j = [T(v_j)]_{B'}$

Def:

use all the above items,
A is the <u>matrix of T</u> relative to basis B, B'.

We write
$$\boxed{A = [T]_B^{B'}}$$
(map B to B')

sometimes we elide B & B' if they're understood

Thm:

Let $T \in Hom(V, W)$.
$\forall v \in V$
$$\boxed{[T(v)]_{B'} = [T]_B^{B'} [v]_B}$$
where the j-th column of $[T]_B^{B'}$ is $[T(v_j)]_{B'}$ where $B = \{v_1, \ldots, v_n\}$.

Thm:

If we fix basis B for V & B' for W,
then the map
$$Hom(V, W) \to F^{m \times n} \cong F^n \to F^m$$
$$T \mapsto [T]_B^{B'}$$
is a vector space isomorphism.

That is we can transform lin. transforms to other lin. transforms.

Coro:

$$\dim(Hom(V, W)) = \dim(V)\dim(W) = mn$$

The basis for $Hom(V, W)$ is $\{E_{ij}\}$ where $1 \le i \le m$ & $1 \le j \le n$.
$$E_{ij}(v_j) = w_i$$
where
$B = \{v_1, \ldots, v_n\}$ basis for V &
$B' = \{w_1, \ldots, w_m\}$ basis for W.

Let us define vector addition, scalar multiplication, & composition of linear maps in terms of how they relate linear maps & matrices

Thm: Vector Addition & Scalar Multiplication are same

Let $T, S \in Hom(V, W)$ & $c \in F$. B basis V & B' basis W.

$$T + S = T(v) + S(v) \iff [T]_B^{B'} + [S]_B^{B'}$$

$$cT = cT(v) \iff c[T]_B^{B'}$$

Thm: Composition is Matrix Multiplication

Let $T \in \text{Hom}(V, W)$, $S \in \text{Hom}(W, Z)$.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[rr, bend right, "ST = S \circ T"'] & W \arrow[r, "S"] & Z \end{tikzcd}$$

where $ST \in \text{Hom}(V, Z)$.

Choose ordered basis B, B', B'' for V, W, Z resp.

Then

note we elide basis decorations

$$[ST] = [S][T]$$

Pf:

$\forall v \in V$ we have

$$\begin{aligned} &[ST][v] \\ &= [(ST)(v)] \\ &= [S(T(v))] \\ &= [S][T(v)] \\ &= [S][T][v] \end{aligned}$$

Thus $\underline{[ST] = [S][T]}$.

Recall:

endomorphism

The identity operator $I \in \text{End}(V) = \text{Hom}(V, V)$.

Then

$$[I]_B^B = \text{identity matrix} = I = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

for all basis B for V. ← must be same basis for this to hold.

Now consider two ordered basis B, B' for V.

$$[v]_{B'} = [I]_B^{B'} [v]_B \quad \forall v \in V$$

Here $P = [I]_B^{B'}$ is the transition matrix from B to B'.

We know P is invertible & $P^{-1} = [I]_{B'}^B$.

$[v]_{B'}$ & $[v]_B$ are non-zero for non-zero v.

Example:

$V = F[x]_{\deg \leq 2}$ has standard basis $B = \{\overset{e_1}{1}, \overset{e_2}{x}, \overset{e_3}{x^2}\}$.

Consider $B' = \{\overset{v_1}{x+1}, \overset{v_2}{x^2+1}, \overset{v_3}{x^2+x}\}$.

We find the transition matrix $P = [I]_{B'}^B$,

$$[v]_B = [I]_{B'}^B [v]_{B'} = P[v]_{B'}$$

We now do this for v_1, v_2, v_3

In general

$$[v_j]_B = P[v_j]_{B'} = Pe_j = j\text{-th column in } P$$

v_j basis vector in B'

Here $[v_j]_B$ is the coordinates of v_j in standard basis, which is easy to find.

$$[v_1]_B = [x+1]_B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$[v_2]_B = [x^2+1]_B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$[v_3]_B = [x^2+x]_B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

This means P is matrix w/ above 3 columns

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

Cool! So $P = [I]^B_{B'}$ is trivial to find.

To find the opposite transition, we do

$$[I]^{B'}_B = P^{-1}$$

which is easy (enough) since P is trivial.

Consider vector space V w/ basis B.
For all linear operators $T \in End(V) \Leftrightarrow F^{n \times n}$ ($n \times n$ matrixes)

$$[T]^B_B = [T]_B = \text{matrix of } T \text{ relative to } B$$

$$[T(v)]_B = [T]_B [v]_B \quad \forall v \in V$$

Now consider basis B'

$$[T(v)]_{B'} = [T]_{B'} [v]_{B'} \quad \forall v \in V$$

Let $P = [I]^{B'}_B$ be the transition matrix from B to B'

$$[v]_{B'} = P[v]_B \quad \forall v \in V.$$

We can relate $[T]_B$ & $[T]_{B'}$ using P. (See next page)

using P we get

$$[T(v)]_{B'} = P[T(v)]_B = P[T]_B[v]_B = [T]_{B'}[v_{B'}] = [T]_{B'}P[v]_B$$

Cancelling $[v]_B$ we get

$$P[T]_B = [T]_{B'}P$$

Thm:

Let $T \in End(V)$ where V has basis B & B'.
Let $P = [I]_B^{B'}$

$$\boxed{[T]_B = P^{-1}[T]_{B'}P}$$

We say $[T]_B$ & $[T]_{B'}$ are conjugate or similar matrices.

Def:

Two $n \times n$ matrices A & B are similar iff $\exists$ invertible matrix P s.t.

$\boxed{B = P^{-1}AP}$ ← Symmetric relation should be obvious

This relation is

i) symmetric $B = P^{-1}AP \Rightarrow A = PBP^{-1} = P'^{-1}BP'$

ii) reflexive $A = I^{-1}AI$

iii) transitive $B = P^{-1}AP$ $C = Q^{-1}BQ \Rightarrow C = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ)$

Therefore similarity is an equivalence relation.

Example:

$V = F[x]_{deg \le 2}$, $B = \{\overset{e_1}{1}, \overset{e_2}{x}, \overset{e_3}{x^2}\}$, $B' = \{\overset{v_1}{1+x}, \overset{v_2}{1+x^2}, \overset{v_3}{x+x^2}\}$

Let $D \in End(V)$ where $D(v) = \frac{d}{dx}v$

We apply D to B

$D(e_1) = D(1) = 0$, $D(e_2) = D(x) = 1$, $D(e_3) = 2x$

$$[D]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

We apply D to B'

$D(v_1) = D(1+x) = 1$ $= 1/2(v_1 + v_2 - v_3)$
$D(v_2) = D(1+x^2) = 2x$, $= v_1 - v_2 + v_3$
$D(v_3) = D(x + x^2) = 1 + 2x$ $= 3/2\, v_1 - 1/2\, v_2 + 1/2\, v_3$

$$[D]_{B'} = \begin{bmatrix} 1/2 & 1 & 3/2 \\ 1/2 & -1 & -1/2 \\ -1/2 & 1 & 1/2 \end{bmatrix}$$

Linear Functionals

it's a function on functions, so functional!

Def:
A linear functional on V is a linear map $V \to F$
where V is a vector space over field F.
The space $Hom(V,F) = V^*$ is called the dual space of V.

Examples:

i) Let $V = C([a,b])$ = space of continuous functions $f: [a,b] \to \mathbb{R}$.

Let $L(f) = \int_a^b f(x)dx \ \forall f \in V$.

L is a linear functional on V, that is $L \in V^*$.

ii) Let $V = F^{n \times n}$.
The trace, the sum of the diagonal,
$tr(A) = \sum_{i=1}^{n} a_{ii}$ where $A = [a_{ij}] \ \forall A \in V$
is a linear functional.

iii) Let $V = F^n$ w/ the standard basis $\{e_1, \dots, e_n\}$. That is

Key: All linear functionals isomorphic to F^n

$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i e_i$.

Like all linear maps

Any $f \in V^* = Hom(V,F)$ is uniquely determined by
$a_i = f(e_i) \in F$ where $a_1, \dots, a_n$ arbitrary.

This generalizes to any finite dimensional space

To find $f(X)$, we find the linearity of f

$$f(X) = f\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} x_i a_i = \underset{F^{1\times n}}{A}\ \underset{F^{n\times 1}}{X}$$

Then $f \mapsto A = [a_1 \dots a_n] \in F^{1\times n}$ is an isomorphism.

Let $B = \{v_1, \dots, v_n\}$ be a basis for V. (finite dimensional)
Any linear functional $f \in V^*$ is uniquely determined by its action on B
$a_i = f(v_i) \in F \ \forall v_i \in B$ & $a_1, \dots, a_n$ can be arbitrary

B/c B is a basis
$v = \sum_{i=1}^{n} x_i v_i \ \forall v \in V$

$\Rightarrow f(v) = f(\sum x_i v_i) = \sum x_i f(x_i) = \sum x_i a_i$

Let $f_i \in V^*$ be such that

$f_i(v) = x_i$ ← take i-th coordinate

$\Leftrightarrow f_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{else} \end{cases}$

Kronecker delta. It's named b/c it's super common

These $f_1, \dots, f_n$ actually give a basis for V^* called the <u>dual basis</u> as

$f(v) = \sum x_i f(v_i) = \sum x_i a_i = \sum \delta_i(v) a_i$

Thus

$f = \sum a_i f_i$

<u>Thm:</u>

Let $S = \{v_1, \dots, v_n\} \subseteq V$ be a subset of V.
Let $S^* = \{f_1, \dots, f_n\} \subseteq V^*$ be a subset of V^* s.t.
$f_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i=j. \\ 0 & \text{o/w} \end{cases}$

The sets S & S^* are linearly independent.

<u>Pf:</u>

Suppose $v = \sum_{j=1}^{n} x_j v_j \in \text{span}(S)$.

$f_i(v) = \sum_{j=1}^{n} x_j f_i(v_j) = x_i$

$x_i = f_i(v) = 0$ iff $v = 0$. Thus S is lin. independent.

Similarly, suppose $f = \sum_{j=1}^{n} c_i f_i \in \text{span}(S^*)$.

$f(v_j) = \sum_{i=1}^{n} c_i f_i(v_j) = c_j$

$c_j = f(v_j) = 0$ iff $f = 0$. Thus S^* is lin. independent.

<u>Thm:</u>

Let V be a vector space $\dim(V) = n < \infty$.
Let $B = \{v_1, \dots, v_n\} \subseteq V$ & $B^* = \{f_1, \dots, f_n\} \subseteq V^*$ s.t.
$f_i(v_j) = \delta_{ij}$.

Then

i) B is a basis for V

ii) B^* is a basis for V^* (<u>dual basis</u> of V)

iii) $v = \sum_{j=1}^{n} x_j v_j = \sum_{j=1}^{n} f_j(v) v_j \quad \forall v \in V$ ← Find v from evaluating B^* on v

iv) $f = \sum_{i=1}^{n} c_i f_i = \sum_{i=1}^{n} f(v_i) f_i \quad \forall f \in V^*$ ← Find f from evaluating f on B

These fall as an extension of the earlier theorem.

Pf:

i) By the earlier theorem B is linearly independent. Since $|B| = \dim(V) = n$ B is a basis.

ii) By the earlier theorem B^* is linearly independent. Since $|B^*| = \dim(V^*) = n = \dim(V)\dim(F)$ B^* is a basis.

Alternatively, every $f \in V^*$ is uniquely determined by $f(v_1), \dots, f(v_n)$ & use iv to see $f \in \text{span}(B^*) = \langle B^* \rangle$

iii) & iv) Done in the earlier theorem.

Example:

$V = F[x]_{\deg \leq n-1}$ w/ (standard) basis $B = \{\overset{v_0}{1}, \overset{v_1}{x}, \dots, \overset{v_{n-1}}{x^{n-1}}\}$. ($n-1$ so $\dim(V) = n$)

We find $B^* = \{f_0, f_1, \dots, f_{n-1}\}$ dual basis.

By definition,
$f_i(v_j) = \delta_{ij}$

evaluation is linear map

$f_0(a_0 + a_1 x + \dots + a_{n-1}x^{n-1}) = a_0 \iff f_0(p(x)) = p(0) \quad \forall p(x) \in V$

derivative & evaluation lin. map

$f_1(a_0 + a_1 x + \dots + a_{n-1}x^{n-1}) = a_1 \iff f_1(p(x)) = \frac{dp}{dx}(0) = p'(0) \quad \forall p(x) \in V$

cancel out power rule exponents

$f_i(p(x)) = \frac{1}{i!}\left(\frac{d}{dx}\right)^i p(0) = \frac{1}{i!} p^{(i)}(0) \quad \forall p(x) \in V$ is the general rule.

Thus $B^* = \left\{ \begin{array}{l} f_i : i \in \{0, \dots, n-1\}, \\ f_i(p(x)) = \frac{1}{i!} p^{(i)}(0) \end{array} \right\}$

Thus $\forall p(x) \in V$
$$p(x) = \sum_{i=0}^{n-1} f_i(p(x)) x^i = \sum_{i=0}^{n-1} \frac{1}{i!} p^{(i)}(0) x^i \leftarrow \text{Taylor Formula}$$

You can expand this to all polynomials $F[x]$ to get Taylor Series!

Example:

We can do the same as above but for
$B = \{1, x-t, \dots, (x-t)^{n-1}\}$
where $t \in F$ fixed.

This gives us

$$B^* = \{f_i : i \in \{0, \dots, n-1\},\ f_i(p(x)) = \frac{1}{i!} p^{(i)}(t)\}$$

Then $\forall p(x) \in V$

$$p(x) = \sum_{i=0}^{n-1} \frac{1}{i!} p^{(i)}(t)\, x^i$$

(Here we get into duality & dual basis. More on that later.)

We call this the evaluation of p at t_i

__Example:__

Now fix $t_1, \dots, t_n \in F$, $t_i \neq t_j\ \forall i \neq j$ & consider $L_i(p(x)) = p(t_i)$.

$B^* = \{L_1, \dots, L_n\} \subseteq V^*$.

$V = F[x]_{\deg \leq n-1}$ (same as above)

We want to find $B = \{p_1, \dots, p_n\} \subseteq V$ s.t.

$L_i(p_j) = p_j(t_i) = \delta_{ij}$.

The fact that these are polynomials & not any function is unimportant.

We define p_j as

$$p_j(x) = \prod_{\substack{1 \leq k \leq n \\ k \neq j}} \frac{x - t_k}{t_j - t_k}$$

Here it is easy to see $p_j(t_i) = \begin{cases} 1, & i = j \\ 0 & o/w \end{cases}$

Using this B we get

$$p(x) = \sum_{j=1}^{n} L_j(p)\, p_j(x) = \sum_{j=1}^{n} p(t_j)\, p_j(x) \leftarrow \text{Lagrange Interpolation Formula}$$

This says every polynomial of degree strictly less than n can be uniquely determined by evaluating it on n unique values.

__Thm:__

Fix $f \in V^*$. $Img(f) \subseteq F$ by definition of $V^* = Hom(V, F)$.

$$rank(f) = dim(Img(f)) = \begin{cases} 1 & \text{if } f \neq 0 \\ 0 & \text{if } f = 0 \end{cases}$$

b/c f is 1D, any vector output being non-zero means whole space covered.

By the rank-nullity theorem

$nullity(f) = dim(Ker(f)) = dim(V) - rank(F)$ ← or $dim(V) = nullity(f) + rank(f)$.

If $f = 0$, then $Ker(f) = V$.

If $f \neq 0$, then $Ker(f) \subset V$ subspace where $dim(Ker(f)) = dim(V) - 1 \Rightarrow codim(Ker(f)) = 1$

__Def:__

Let $W \subseteq V$ be a subspace of V. The __codimension__ of W is
$codim(W) = dim(V) - dim(W)$.

Essentially how much the subspace loses.

Def:

A hyperspace of vector space V w/ subspace W where $\operatorname{codim}(W) = 1$.

This is the maximal proper subspace.

Thm:

i) $\forall f \in V^*$, $f \neq 0$ $\operatorname{Ker}(f) = N_f$ is a hyperspace.

ii) $\forall W \subset V$ where $\operatorname{codim}(W) = 1$ (W hyperspace) $\exists f \in V^*$ s.t. $W = \operatorname{Ker}(f)$

Pf:

i) We proved this in an earlier "theorem."

ii) Pick a basis $\{v_1, \ldots, v_{n-1}\}$ for W & extend it to a basis for V $\{v_1, \ldots, v_n\}$.

Define $f \in V^*$ by $f(v_i) = 0$ $1 \leq i \leq n-1$, $f(v_n) = 1$. ← actually f_n of dual basis!

Here f is an $f \in V^*$ s.t. $\operatorname{Ker}(f) = W$.

Def:

For subset $S \subseteq V$, its <u>annihilator</u> is $S^0 = \{ f \in V^* \mid f(v) = 0 \ \forall v \in S \}$.

If $S = \{0\}$, then $S^0 = V$

Note that S^0 is a subspace of V^*. (

Closed Under Addition: $f(v) + g(v) = 0 + 0 = 0$

Closed Under Multiplication: $cf(v) = c(0) = 0$

Thm:

Let $W \subseteq V$ be a subspace of V where $\dim(V) < \infty$.

Then $\dim(W^0) = \operatorname{codim}(W) = \dim(V) - \dim(W)$

This is b/c $\operatorname{codim}(W) = \dim(\operatorname{Ker}(f)) = \operatorname{nullity}(f)$.

Pf:

Pick a basis $\{v_1, \ldots, v_k\}$ for W & extend it to basis $B = \{v_1, \ldots, v_n\}$ for V.

The dual basis B^* for V^* is $B^* = \{f_1, \ldots, f_n\}$.

By the definition of dual basis

$f_i(v_j) = \delta_{ij}$

Take $1 \le j \le k$ & $k+1 \le i \le n$. Then $i \neq j$. Thus $f_i(v_j) = 0$.
Then $f_i \in W^0$.

f_i are elements in dual basis that vanish $\forall v \in W$

Recall $\forall f \in V^*$
$$f = \sum_{j=1}^{n} f(v_j) f_j.$$

Then $f \in W^0 \iff f(v_j) = 0 \ \forall 1 \le j \le k \iff \mathrm{span}\{f_{k+1}, \dots, f_n\}$.

This gives us
$$W^0 = \mathrm{span}\{f_{k+1}, \dots, f_n\} \quad (\& \ \{f_{k+1}, \dots, f_n\} \text{ is a basis for } W).$$

Coro:
$$W = N_{f_{k+1}} \cap \dots \cap N_{f_n}.$$

Pf:
Let $v \in V$. $v = \sum_{i=1}^{n} x_i v_i$, then $v \in N_{f_j} \iff x_j = 0$ & $x_j = f_j(v)$.

Double Dual

Def:
Let V be a vector space over F.
$V^* = \mathrm{Hom}(V, F)$ is the dual space.
$V^{**} = (V^*)^* = \mathrm{Hom}(V^*, F)$ is the double dual (space).

We want to find symmetry b/w V & V^* via double dual.

$\forall v \in V$, there is a linear functional $L_v : V^* \to F$ i.e. $L_v \in (V^*)^* = V^{**}$

$L_v(f) = f(v)$ for $f \in V^*$. ← This is called the evaluation of f on V.
Also works for $\dim(V) = \infty$:)

Lemma:
The map $v \to L_v$ where $v \in V$ & $L_v \in V^{**}$ is a linear map.

Pf:
We first assert L_v is closed under addition & multiplication.

$$L_v(f+g) = L_v(f) + L_v(g) \quad \& \quad L_v(cf) = cL_v(f) \qquad f, g \in V^* \ \& \ c \in F.$$ ← you can check yourself

Thus $L_v \in V^{**}$

Further

$$L_{v+w} = L_v + L_w \quad \& \quad L_{cv} = cL_v \qquad v, w \in V, \ c \in F.$$

Thus $v \to L_v$ is linear.

Suppose $v \in V$ s.t. $L_v = 0$. Then
$L_v(f) = f(v) = 0 \quad \forall f \in V^*$

If $v \neq 0$, then $\exists$ basis for V containg v, so then there is a linear functional $f \in V^*$ s.t. $f(v) \neq 0$.

In this case where $L_v = 0$, $v = 0$. Thus the map is injective.

In general, $\dim(\text{Hom}(V,W)) = \dim(V)\dim(W)$.

Thus $\underline{\dim(V^*)} = \dim(\text{Hom}(V,F)) = \dim(V)\dim(F) = \underline{\dim(V)}$.
Likewise $\dim(V^{**}) = \dim(V)$.

Thm:
The map $v \in V \to L_v \in V^{**}$, where V is a finite-dimensional vector space, is a vector space isomorphism $V \to V^{**}$.

Example:
This theorem breaks down for infinite dimension V's.

Let $V = F[x]$ w/ basis $\{1, x, x^2, \dots\}$.
Let $f \in V^*$ be defined as
$a_i = f(x^i) \in F$ where $i \in \{0, 1, 2, \dots\}$

For an arbitrary vector $x \in F[x]$,
$f(x) = \sum_{i=0}^{\infty} a_i x^i$.
Since polynomials must be finite $f(x) \notin F[x]$ in general.

Instead $f(x)$ can be infinite, so we say
$f(x) \in F[[x]]$, ← power series

Thus $F[x]^* \cong F[[x]]$

Notation:
For all linear functionals $f \in V^*$ & vectors $v \in V$, we write the evaluation of f & v as
$\langle f, v \rangle = f(v) \in F$

This gives a map from $V^* \times V \to F$. We call this map bilinear b/c it is linear in both arguments when one is fixed.

This notation is nice b/c it makes the relationship more symmetric seeming, giving intuition for L_v.
$\langle L_v, f \rangle = \langle f, v \rangle = \langle v, f \rangle$
if we identify $v \in V$ w/ $L_v \in V^{**}$

Rem:

Let $B = \{v_1, \dots, v_n\}$ basis for V.
Let $B^* = \{f_1, \dots, f_n\}$ basis for V^*, dual basis for V.

Then the dual basis of V^* (i.e. B^{**} for V^{**}) is
$B^{**} = \{L_{v_1}, \dots, L_{v_n}\}$ basis for V^{**}, dual basis for V^*.

Rem:

The isomorphism $V \to V^{**}$ is natural, unlike the isomorphism $V \to V^*$.

By natural we mean
i) it doesn't depend on a choice of basis
ii) it behaves well wrt linear maps.

Informally,
The mapping $f: S \to S^{**}$ transforms to $f: T \to T^{**}$ where all properties hold.

In other words, its a functor on the category of vector spaces.

Recall:

The annihilator W^0 of $W \subseteq V$ is
$W^0 = \{f \in V^* \mid f(w) = 0 \;\; \forall w \in W\} \subseteq V^*$ subspace

Thm:

Let $W \subseteq V$ subspace w/ $\dim(V) < \infty$.

Then $W^{00} = (W^0)^0 \subseteq V^{**}$ corresponds to W under $V^{**} \cong V$.

isomorphism b/w V & V^{**}

Pf:

Recall $V^{**} \cong V$ is given by $v \to L_v : V \to V^{**}$, where $L_v(f) = f(v) \;\; \forall f \in V^*$.

1st way

Take $v \in W$ & $f \in W^0$. Then $f(v) = 0$ by definition.
By the definition of L_v
$L_v(f) = f(v) = 0$
Then $L_v \in W^{00}$.

Therefore W is sent to a subspace of W^{00} under $V \xrightarrow{\cong} V^{**}$.

We want to show W & W^{00} are equal. We'll go thru two ways one w/ dimension & the other w/ basis.

First, let's use dimension.
$\dim(W^0) = \operatorname{codim}(W) = \dim(V) - \dim(W)$
$\dim(W^{00}) = \dim(V^*) - \dim(W^0) = \dim(V) - (\dim(V) - \dim(W)) = \dim(W)$

We have $W \subseteq W^{00}$ but $\dim(W) = \dim(W^{00})$, so $W = W^{00}$. □

Next, let's use basis.
Pick a basis $\{v_1, \dots, v_k\}$ for W.
Extend it to a basis $\{v_1, \dots, v_n\}$ where $k \le n$ for V.
Take the dual basis $\{f_1, \dots, f_n\}$ for V^*.

The basis for W^0 is $\{f_{k+1}, \dots, f_n\}$ b/c
$$f_i(v_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$
b/c $\forall i, j \in \mathbb{Z}$ $k+1 \le j \le n$ $1 \le i \le k$ $i \neq j$.

This (& next page) is a full proof.

Now we do the same for $W^\circ \subseteq V^{**}$.
Start w/ basis $\{f_{k+1}, \dots, f_n\}$ for W°.
Extend it to basis $\{f_1, \dots, f_n\}$ for V^* (from earlier).
Take the basis $\{L_{v_1}, \dots, L_{v_n}\}$ for V^{**} (dual basis for V).
By $V^{**} \cong V$, $\{v_1, \dots, v_n\}$ is a basis for V.
Hence a basis for $W^{\circ\circ}$ is $\{L_{v_1}, \dots, L_{v_k}\}$ by similar logic.
This is the basis for W identified by $V^{**} \cong V$.

Therefore $W^{\circ\circ} = W$. □

Matrix Coefficients of a Linear Map

Let $T \in \text{Hom}(V, W)$ where V, W are vector spaces.
In order to represent T by a matrix, we need basis
$B = \{v_1, \dots, v_n\}$ for V & $B' = \{w_1, \dots, w_m\}$ for W.
Then $A = [T]_B^{B'} \in F^{m \times n}$ is the matrix of T relative to B, B'.

we elide parens

The j-th column of $A = [T] = (a_{ij})$ is $[Tv_j]_{B'}$

elide parens → $Tv_j = \sum_{i=1}^{m} a_{ij} w_i \quad \forall j \in \{1, \dots, n\}$

Consider the dual basis $(B')^*$ of B'.
$(B')^* = \{g_1, \dots, g_m\}$ basis for W^*
where $g_k(w_i) = \delta_{ki}$
Then
Then
$$a_{ij} = g_i(Tv_j) = \langle g_i, Tv_j \rangle$$

These are called the <u>matrix coefficients of the linear map T</u>.

Transpose of Linear Transforms

Let $T \in \text{Hom}(V, W)$.
Take $g \in W^*$ & $v \in V$.
$Tv \in W$, so it makes sense to apply G. That is
$$g(Tv) = \underbrace{\langle g, Tv \rangle}_{} \in F$$
We can view this as a bilinear map $W^* \times V \to F$. That is it is linear in g $\forall$ fixed v & linear in v $\forall$ fixed g.

This gives us a linear functional $S_g : V \to F$ where
$S_g v = g(Tv)$ where $S_g \in V^*$.

This means $\forall g \in W^*$ there is a linear functional $S_g \in V^*$.

this depends linearly on g, so we can think of S (not applied to any g) as
$S: W^* \to V^*$
where
$Sg = S_g$ where $S_g \in$
So...
$(S_g)(v) = g(Tv) \iff \langle S_g, v \rangle = \langle g, Tv \rangle$
where $S_g \in V^*$ & $v \in V$.

I like to think of this as $S_\square$ where we have a "slot" for $g \in W^*$ which returns V^*.

We call S the transpose of T
or adjoint of T

Def:
$S = T^t$ is the transpose of $T: V \to W$ where $S: W^* \to V^*$ defined by

this is called a contravariant function

$$\underbrace{(T^t g)}_{V^*}\underbrace{(v)}_{V} = \underbrace{g}_{W^*}\underbrace{(Tv)}_{W} \iff \langle T^t g, v \rangle = \langle g, Tv \rangle$$

We can visualize this w/ the following diagram

$$\begin{tikzcd} V \arrow[rr, "T"] \arrow[dr, "T^t g"'] & & W \arrow[dl, "g"] \\ & F & \end{tikzcd}$$

That is $T^t g = g \circ T$

Thm:
Let $T \in \mathrm{Hom}(V, W)$. Then $T^t \in \mathrm{Hom}(W^*, V^*)$.
$[T^t]$ corresponds to $[T]^t$.

Pf:
Let the following be basis
$B = \{v_1, \dots, v_n\}$ for V,
$B' = \{w_1, \dots, w_m\}$ for W,
$B^* = \{f_1, \dots, f_n\}$ for V^* dual basis to B where $f_i(v_j) = \delta_{ij}$,
$B'^* = \{g_1, \dots, g_m\}$ for W^* dual basis to B'

$[T]_B^{B'} \in F^{m \times n}$ matrix T wrt B, B'.

$[T^t]_{B'^*}^{B^*} \in F^{n \times m}$ matrix T^t wrt B'^*, B^*.

We show $[T]_B^{B'}$ & $[T^t]_{B'^*}^{B^*}$ are transpose of each other.

To do this we use our matrix coefficients from earlier.
Last time we found the (i,j)-th entry of $[T]$ is $g_i(Tv_j)$.

By the definition of T^t
$g_i(Tv_j) = (T^t g_i)(v_j)$

If we identify $V \cong V^{**}$ we get
$g_i(Tv_j) = (T^t g_i)(v_j) = L_{v_j}(T^t g_i)$

This is the (j, i)-th entry of $[T^t]$.

Thus $[T]^t = [T^t]$. $\square$

Prop:

Many of these come directly from the properties of matrix transpose.

i) $(ST)^t = T^t S^t$ where $T \in Hom(V,W)$, $S \in Hom(W,U)$

ii) $(T^t)^t = T$ if we identify $V \cong V^{**}$ & $W \cong W^{**}$

Thm: not assuming $\dim(V)<\infty$ or $\dim(W)<\infty$

Let $T \in Hom(V,W)$. Then

i) $\underbrace{Ker(T^t)}_{\subseteq W^*} = \underbrace{(\underbrace{Img(T)}_{\subseteq W})^\circ}_{\subseteq W^*}$

ii) $\underbrace{Img(T^t)}_{V^*} \subseteq \underbrace{(\underbrace{Ker(T)}_{\subseteq V})^\circ}_{\subseteq V^*}$. This is equal if $\dim(V)<\infty$ & $\dim(W)<\infty$

Pf:

i) We use the definition of T^t
$(T^t g)v = g(Tv) \quad g \in W^*, v \in V$

$g \in Ker(T^t)$ iff $T^t g = 0 \in V^*$.

$0 \in V^*$ is the functional such that $0(v) = 0 \quad \forall v \in V$. Thus

$$\begin{aligned} g \in Ker(T^t) &\Leftrightarrow T^t g = 0 \in V^* \\ &\Leftrightarrow T^t g = 0 \in V^* \\ &\Leftrightarrow (T^t g)(v) = 0 \quad \forall v \in V \\ &\Leftrightarrow g(Tv) = 0 \quad \forall v \in V \\ &\Leftrightarrow g(w) = 0 \quad \forall w \in Img(T) \\ &\Leftrightarrow g \in [Img\, T]^\circ \end{aligned}$$

ii) We try to do something similar, but if we go both directions we fail (if non-finite).

We want to show $\forall f \in Img(T^t)$, $f \in (Ker(T))^\circ$.

Take any $f = T^t g \in Img\, T^t$ for some $g \in W^*$. We show f "vanishes" on the kernel of T. Let $v \in Ker(T)$

$$f(v) = (T^t g)(v) = g(Tv) = g(0) = 0.$$

Thus $f \in (Ker\, T)^\circ$, so $Img(T^t) \subseteq (Ker(T))^\circ$.

We now take on the finite dimensional case. To do that we show their dimensions are equal. We do this w/ the rank-nullity theorem.

$$\mathrm{rank}(T^t) + \mathrm{nullity}(T^t) = \dim(\mathrm{Dom}(T^t))$$
$$\Rightarrow \mathrm{rank}(T^t) = \dim(\mathrm{Dom}(T^t)) - \mathrm{nullity}(T^t)$$
$$\Rightarrow \dim(\mathrm{Img}(T^t)) = \dim(W^*) - \dim(\mathrm{Ker}(T^t))$$

From part (i), $\mathrm{Ker}(T^t) = (\mathrm{Img}(T))^0$. $\dim((\mathrm{Img}(T))^0) = \mathrm{codim}(\mathrm{Img}(T)) = \dim(W) - \dim(\mathrm{Img}(T))$

$$\dim(\mathrm{Img}(T^t)) = \dim(W) - \dim((\mathrm{Img}(T))^0)$$
$$\Rightarrow \dim(\mathrm{Img}(T^t)) = \dim(\mathrm{Img}(T)) = \mathrm{rank}(T)$$
$$= \mathrm{rank}(T)$$
$$= \dim(V) - \dim(\mathrm{Ker}(T)) \leftarrow$$
$$= \mathrm{codim}(\mathrm{Ker}(T))$$
$$= \dim((\mathrm{Ker}(T))^0)$$

Thus $\dim(\mathrm{Img}(T^t)) = \dim((\mathrm{Ker}(T))^0)$ & $\mathrm{Img}(T^t) \subseteq (\mathrm{Ker}(T))^0$, so
$\mathrm{Img}(T^t) = (\mathrm{Ker}\, T)^0$
when $\dim(V) < \infty$ & $\dim(W) < \infty$.

Coro:
$\forall T \in \mathrm{Hom}(V, W)$ where $\dim(V) < \infty$ & $\dim(W) < \infty$,
$\mathrm{rank}(T) = \mathrm{rank}(T^t)$

This is the same as $\forall A \in M$
$\mathrm{rowrank}(A) = \mathrm{colrank}(A)$

Chapter 5: Determinants

The previous section was chapters 1-3

Recall:

i) $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \leftarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

ii) $\forall A \in M,\ \det(A) \neq 0 \iff A$ invertible

Def:

The determinant $\det: F^{m \times n} \to F$ is a function from $m \times n$ matrixes to their field.

We can decompose or view the determinant as $D: \underbrace{F^n \times \dots \times F^n}_{n \text{ times}} \to F$.

also works for rows

That is a function of the columns of the matrix

Properties:

i) The determinant D is n-linear, that is linear in each column if we keep the other $n-1$ columns fixed.

same for columns

$\therefore D(A_1, \dots, B_i + C_i, \dots, A_n)$ (that is make $A_i = B_i + C_i$)
$= D(A_1, \dots, A_{i-1}, B_i, A_{i+1}, \dots, A_n)$
$+ D(A_1, \dots, A_{i-1}, C_i, A_{i+1}, \dots, A_n)$

$\therefore D(A_1, \dots, \lambda A_i, \dots, A_n) = \lambda D(A_1, \dots, A_i, \dots, A_n) \quad \forall \lambda \in F$

$\therefore D(\lambda A) = \lambda^n D(A)$

ii) D is an alternating function. That is it changes sign when we swap two columns

$D(A_1, \dots, A_i, \dots, A_j, \dots, A_n) \quad i<j$
$= -D(A_1, \dots, A_j, \dots, A_i, \dots, A_n)$

$\therefore A_i = A_j$ for any $i \neq j \Rightarrow D(A) = 0$ Given characteristic$(F) \neq 2$.

iii) $D(I) = 1$ where I = identity matrix $\leftarrow$ This is required to uniquely determine determinant b/c o/w you'd have scalar multiples.

Normalization

Thm: $\exists!$ n-linear alternating function $\det: F^{n\times n} \to F$ where $\det(I) = 1$.

That is any n-linear alternating function $D: F^{n\times n} \to F$ can be written as
$D(A) = D(I)\det(A) \quad \forall A \in F^{n\times n}$. ← Shows det is unique

Pf:

We construct such a $\det: F^{n\times n} \to F$ where $\det(I)$ exists & then we show it is unique.

We define det as

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1,\sigma(1)}\, a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \qquad \forall A = (a_{ij}) \in F^{n\times n}$$

permutations of $1..n$

sign (± 1) of σ, depending on σ

where

$\sigma: \{1, \dots, n\} \to \{1, \dots, n\}$ bijection (thus permutation) b/c $S_n = \{f: 1..n \to 1..n \mid f \text{ bijective}\}$

$\operatorname{sgn}(\sigma) = (-1)^m$ if σ = product of m transpositions

well defined b/c $m \pmod 2$ invariant b/c any extra swap is pair

sgn(σ) makes the alternating clear. Every entry appears exactly once in sum/products & every column only appears once in a product so it is n-linear. det is normal b/c only the diagonal entries are non-zero, so only one entry (where σ = identity) is non-zero & in particular it is 1, so $\det(I) = 1$.

We now show uniqueness by showing any n-linear, alternating function $D: F^{n\times n} \to F$ can be written as

$D(A) = D(I)\det(A) \quad \forall A \in F^{n\times n}$

To do this we use elementary column operations. (or similarly row operations)

i) multiply column by scalar & add it to another

$i \neq j \to D(\dots, \underbrace{cA_i}_{\text{add}} + \underbrace{A_j}_{\text{old}}, \dots) = c\,D(\dots, A_i, \dots) + D(\dots, A_j, \dots) = D(A_1, \dots, A_n)$.

@ idx = j — has A_i @ idx = i, so $D(\dots) = 0$ w/ duplicate columns

$\therefore D$ stays the same

ii) Multiply column by non-zero scalar

$D(A_1, \dots, cA_i, \dots, A_n) = c\,D(A_1, \dots, A_n)$ (i-th col)

$\therefore D$ multiplied by same scalar $c \neq 0$

iii) Swap two columns

$D(A_1, \dots, A_j, \dots, A_i, \dots, A_n) = -D(A_1, \dots, A_n)$ ← from axioms/def of D

$\therefore D$ changes sign

Note that for each elementary column operation, D is mulitplied by a non-zero scalar.

like row reduce

With these column operations, we column reduce the matrix

$A \to \dots \to R \leftarrow$ reduced matrix

When reduced, R will either be the identity matrix or have a zero column.

When $R = I$, $D(R) = D(I)$. When has zero column, $D(R) = 0$
det has the same property but $\det(I) = 1$.

Hence

$$D(A) = c\,D(A) \quad \& \quad \det(A) = c\det(R) \quad \text{for same } c \in F$$

So

$$D(A) = D(I)\det(A)$$

This comes from division (case where $R = I$)

$$\frac{D(A)}{\det(A)} = \frac{c\,D(R)}{c\det(R)} = \frac{\cancel{c}D(I)}{\cancel{c}\cancel{\det(I)}}$$

$$\Rightarrow D(A) = D(I)\det(A)$$

Properties of Determinant:

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

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

iii) $\det(P^{-1}AP) = \det(A)$
(Determinant of two similar matrices is equal)

iv) $\det\begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \det(A)\det(C)$

block triangular matrix

Coro:

i) For linear operator $T \in \operatorname{End}(V)$ we can define
$\det(T) = \det(A)$ where $A = [T]_B$
wrt some ordered basis B.

This is well defined b/c if we change basis B, we get a similar matrix $P^{-1}AP$ & $\det(P^{-1}AP) = \det(A)$.

ii) $\det\begin{bmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{bmatrix} = \prod_{i=1}^{n} a_i = a_1 \cdots a_n$ ← nice for computation along w/ elementary operations b/c you can get a string of things to multiply until you get triangular matrix

upper triangular matrix

also works for lower triangular

Def: ← can be infinite dimensional

Let V be a vector space w/ $W \subseteq V$ subspace. The quotient space/ factor space V/W is the set of all cosets W in V.

$V/W = \{v + W \mid v \in V\}$

Recall:

A coset of V in W is the following for some $v \in V$.

$v + W = \{v + w \mid w \in W\} \subseteq V$ subset.

We call $v \in V$ the representative of the coset $v + W$.

The representative is not unique.

Two cosets w/ representatives $v, v' \in V$ are equal iff $v - v' \in W$

$v + W = v' + W \Leftrightarrow v - v' \in W$ ← you could reach v' from v using only elements of W.

In other words

$(v + w) + W = v + W \quad \forall w \in W$.

For V/W to be a vector space, we need to define operations on cosets

Vector Operations on Cosets:

Let $u, v \in V$ & $c \in F$.

i) $(u + W) + (v + W) = (u + v) + W$

ii) $c(v + W) = (cv) + W$

The zero vector in V/W is $0 + W = W$ (& $w + W = 0 + W \quad \forall w \in W$). b/c $w - 0 = w \in W$

All vector space properties fall from V.

Thus V/W is a vector space.

Thm:

Let $\{v_1, \ldots, v_m\}$ be a basis for W & extend it to a basis $\{v_1, \ldots, v_n\}$ for V.

Then $\{v_{m+1} + W, \ldots, v_n + W\}$ is a basis for V/W. ($v_1 + W, \ldots, v_m + W$ are all $0 + W$ b/c $v_1, \ldots, v_m \in W$.)

Rem:

$\dim(V/W) = \dim(V) - \dim(W)$

Thm: 1st Isomorphism Theorem

For all linear transformation $T \in Hom(V,W)$, we have a vector space isomorphism

$V/Ker(T) \cong Img(T)$. This gives another way to prove rank-nullity!

This is given by $v + Ker(T) \to Tv$ b/c $\forall v \in Ker(T)$ $Tv = 0$.

Tensor Products

Informally, tensor products are what happen when you try to multiply two vector spaces. We could do cartesian product of two vector spaces but that is no longer a vector space & lacks nice properties like distributivity. The tensor product is similar to the cartesian product but "fixes" these issues.

Let V, W be vector spaces (can be infinite) over F. We want the following properties for the tensor product $\otimes$

$$\left.\begin{aligned} (cv)\otimes w &= c(v\otimes w) = v\otimes(cw) \\ (v+v')\otimes w &= v\otimes w + v'\otimes w \\ v\otimes(w+w') &= v\otimes w + v\otimes w' \end{aligned}\right\} \star$$

That is we want the tensor product $\otimes$ to be bilinear.

Def:

Let V, W be vector spaces. We construct the tensor product $\otimes$.

We start w/ the cartesian product $V \times W$ & construct a vector space S consisting of all formal linear combinations such that $V \times W$ is a basis. (This is huge b/c $V \times W$ is huge & the basis)

We impose relations on S by taking the quotient space S/K where K is the vector space spanned by the properties mentioned in $\star$. That is K's elements have the following

$c(v,w) - (cv,w) = 0$

$c(v,w) - (v,cw) = 0$

$(v+v',w) - (v,w) - (v',w) = 0$

$(v, w+w') - (v,w) - (v,w') = 0$

$\forall c \in F$ $v, v' \in V$ $w, w' \in W$.

The tensor product of V & W is

$V \otimes W = S/K$ where $S = V \times W$ & K subspace of S spanned by linearity relations

& for $v \in V$ $w \in W$, their tensor product is

$v \otimes w = (v,w) + K$

More Quotient Spaces

Thm:

Let V be a vector space w/ $W \leq V$ subspace. Let U be another vector space.

Then there is a natural isomorphism

$$Hom(V/W, U) \cong \{T \in Hom(V,U) \mid \underset{\substack{\uparrow \\ \text{or } W \subseteq Ker(T) = Null(T)}}{Tw = 0} \ \forall w \in W\}$$

connected to annihilator W^0 where U is field

Pf: ~~x~~ where $Tv=0$ ← critical

Given $T \in Hom(V, U)$, we get an induced map $\bar{T}: V/W \to U$, where
$\bar{T}(v+W) = Tv \quad \forall v \in V$

This is an isomorphism b/c we can define T given $\bar{T}$.

This is well defined b/c if we choose v' where $v-v' = w \in W$
Then by the definition of cosets
$\bar{T}(v+W) = \bar{T}(v'+W)$

We can do the same in reverse

This is true b/c if we rewrite v' as $v' = v-w$, we get
$T(v) = T(v'-w) = T(v') - T(w) = T(v')$.

Thus $\bar{T}(v+W) = \bar{T}(v'+W) \Leftrightarrow Tv = Tv'$ b/c $v-v' = w \in W$ must be true in both cases. □

Coro:
If $U = F$, then $(V/W)^* \cong W^0$.

Thm:
There is a natural isomorphism $T: V \to V/W$ where
$T(v) = v+W$.

Here $W = Ker(T)$

The tensor product turns a bilinear function into a linear one

Thm:
or $K \in Ker(T)$

$\forall U$ vectorspace
$Hom(V \otimes W, U) = Hom(S/K, U) \cong \{T \in Hom(S, U) \mid T(K) = 0\}$
$\cong \{T: V \times W \to U \mid T \text{ bilinear}\}$

A linear map $T: S \to U \Leftrightarrow$ function $T: V \times W \to U$. That is
$T(v, w) \in U \quad \forall v \in V \ w \in W$

The condition $T(K) = 0$ means
$T(c(v,w) - (cv, w)) = 0 \Leftrightarrow T(c(v,w)) = T(cv, w)$
$T(c(v,w) - (v, cw)) = 0 \Leftrightarrow T(c(v,w)) = T(v, cw)$
$T((v+v', w) - (v,w) - (v',w)) = 0 \Leftrightarrow T(v+v', w) = T(v,w) + T(v',w)$
$T((v, w+w') - (v,w) - (v,w')) = 0 \Leftrightarrow T(v, w+w') = T(v,w) + T(v,w')$

That is T is bilinear.

This means $Hom(V \otimes W, U)$ is the space of bilinear maps
$T: V \times W \to U$.

Coro:

Any bilinear map $T: V \times W \to V$ induces a linear map $\bar{T}: V \otimes W \to U$ given by

$\bar{T}(v \otimes w) = T(v, w) \quad \forall v \in V \; w \in W.$

Remark:

$V \otimes W$ is spanned by all pure tensors $v \otimes w$.

We will use the above theorem for the tensor product b/c it is equivalent to the quotient space construction & more useful.

We can actually generalize the tensor product to n spaces.

Def:

We can actually generalize the tensor product to n spaces.

We use associativity

$V_1 \otimes V_2 \otimes V_3 \otimes \ldots \cong ((V_1 \otimes V_2) \otimes V_3) \otimes \ldots \cong V_1 \otimes (V_2 \otimes (V_3 \otimes \ldots))$

or switch to n-tuples.

Coro:

Then an n-linear map $V_1 \times \ldots \times V_n \to U \cong$ a linear map $V_1 \otimes \ldots \otimes V_n \to U$.

Example:

Let $V = F$. Find $a \otimes w$ where $a \in F = V$ & $w \in W$ where W is a vector space

$a \otimes w = (a \cdot 1) \otimes w = a(1 \otimes w)$

Thus $F \otimes W$ is spanned by vectors $1 \otimes w$ where $w \in W$.

$F \otimes W = \{1 \otimes w \mid w \in W\}$

Take $w, w' \in W$, then $\underbrace{1 \otimes w}_{\in F \otimes W} + \underbrace{1 \otimes w'}_{\in F \otimes W} = \underbrace{1 \otimes (w + w')}_{\in F \otimes W}$.

So $F \otimes W \cong W$.

Consider the bilinear map $T: F \times W \to W$, where $T(a, w) = aw$.

This enduces the linear map $\bar{T}: F \otimes W \to W$ & it's an isomorphism

Example:

Let V & W be vector spaces.

$\forall g \in V^*$ & $h \in W^*$ we have a bilinear map $T: V \times W \to F$ where

$T(v, w) = g(v)\, h(w)$

T is bilinear b/c g & h are linear & the product of two linear maps (given independent values) is bilinear.

This gives an induced linear map $\bar{T}: V \otimes W \to F$ where

$\bar{T}(v \otimes w) = g(v)\, h(w)$

Here $\bar{T} \in (V \otimes W)^*$

Def:

Let V, V', W, W' be vector spaces.
Let $T: V \to V'$ & $S: W \to W'$ be vector spaces.
This gives us a bilinear map $F: V \times W \to V' \otimes W'$

$$F(v, w) \mapsto (Tv) \otimes (Sw)$$ ← normally this is anonymous

This bilinear map gives us an induced linear map

$$T \otimes S = \overline{F}: V \otimes W \to V' \otimes W'$$

where

$$(T \otimes S)(v \otimes w) = \overline{F}(v \otimes w) = (Tv) \otimes (Sw)$$

This defines the <u>tensor product of linear maps.</u>

Remark:

This tensor product of linear maps naturally gives us the <u>Kronecker product</u> of matrices.

Thm:

we use index set rather than $1..n$ b/c this applies in infinite dimensional cases.

Let $\{v_i\}_{i \in I}$ be a basis for V.
Likewise $\{w_j\}_{j \in J}$ for W.

Then $\{v_i \otimes w_j\}_{i \in I, j \in J}$ is a basis for $V \otimes W$

& $\dim(V \otimes W) = \dim(V)\dim(W)$

Pf:

We show $\{v_i \otimes w_j\}_{i \in I, j \in J}$ is a basis for $V \otimes W$. The dimension property falls from that trivially.

First we show $\{v_i \otimes w_j\}_{i \in I, j \in J}$.
By definition $V \otimes W = \text{span}\{v \otimes w \mid v \in V, w \in W\}$
Since $\{v_i\}_{i \in I}$ is a basis for V, we can rewrite v as (& likewise W)

$$v = \sum_i a_i v_i \quad \& \quad w = \sum_j b_j w_j.$$

Thus

$$v \otimes w = \left(\sum_i a_i v_i\right) \otimes \left(\sum_j b_j w_j\right)$$

by the linearity & distributivity of the tensor product, we can rewrite this as

$$v \otimes w = \sum_i \sum_j a_i b_j (v_i \otimes b_j).$$

b/c we wrote $v \otimes w$ $\forall v \in V$ $w \in W$ as l.c. of $\{v_i \otimes w_j\}_{i \in I, j \in J}$

So $V \otimes W = \text{span}\{v_i \otimes w_j \mid i \in I, j \in J\}$.
Thus $\{v_i \otimes w_j\}_{i \in I, j \in J}$ spans $V \otimes W$.

Now we show $\{v_i \otimes w_j\}_{i\in I, j\in J}$ is linearly independent. To do this, we show $\sum_{i,j} a_{ij} (v_i \otimes w_j) = 0$ if & only if $\forall a_{ij} = 0$.

To do this we use the dual basis for V^* b/c given $v \in V$ where $v = \sum a_i v_i$
we can rewrite a_i (& thus v) as applications of dual basis on v. So
$a_i = f_i(v)$
$\Leftrightarrow v = \sum f_i(v) v_i$

Different V & V^*. Just generally true

Let $\{g_i\}_{i\in I}$ be the dual basis to $\{v_i\}_{i\in I}$ & $\{h_j\}_{j\in J}$ be the dual basis to $\{w_j\}_{j\in J}$. So
$g_k(v_i) = \delta_{k,i}$ & $h_l(w_j) = \delta_{l,j}$

Then $g_k \otimes h_l \in (V \otimes W)^*$
$$(g_k \otimes h_l)(v_i \otimes w_j) = g_k(v_i) h_l(w_j) = \delta_{k,i}\delta_{l,j} = \begin{cases} 1 & \text{if } (k,l) = (i,j) \\ 0 & \text{o/w} \end{cases}$$

By the definition of $(g_k \otimes h_l)$ & the nature of dual bases $\forall v, w$
$$v \otimes w = \sum_{i,j} \big((g_i \otimes h_j)(\underset{v\otimes w}{\underbrace{t}})\big)(v_i \otimes w_j) = \sum_{i,j} a_{ij}(v_i \otimes w_j)$$

Applying this to $\sum a_{ij}(v_i \otimes w_j) = 0$, we get
$(g_i \otimes h_j)(0) = 0 \ \forall i \in I\ j \in J$ (b/c $\{g_i \otimes h_j\}_{i\in I, j\in J}$ basis for $(V\otimes W)^*$)

Thus
$a_{ij} = 0 \ \forall i \in I\ j\in J$
& $\{v_i \otimes w_j\}_{i\in I\, j\in J}$ is linearly independent.

Thus $\{v_i \otimes w_j\}_{i\in I\, j\in J}$ is a basis for $V\otimes W$. □

<u>Thm</u>:
If $\dim(V) < \infty$ & $\dim(W) < \infty$, we have a canonical/natural isomorphism $\varphi : V^* \otimes W \to \mathrm{Hom}(V, W)$ where
$$(\varphi(\overset{\in V^*}{f} \otimes \overset{\in W}{w}))(\overset{\in V}{v}) = f(v)w$$
dimensions match so this has a shot.

<u>Pf</u>:
To have linear map $\varphi : V^* \otimes W \to \mathrm{Hom}(V, W)$, we need a bilinear map $T : V^* \times W \to \mathrm{Hom}(V, W)$.

Let $T(f, w) = T_{f,w}$ where $T_{f,w}(v) = f(v)w$. We show T is well defined & bilinear.

First, $T_{f,w}$ is indeed a linear map $V \to W$ b/c all components are linear.

We now check T is bilinear. This is linear (if you squint & believe) when you fix w & separately when you fix f. Thus T is bilinear & so is φ.

To show the map φ is an isomorphism, we show it sends a basis to a basis. (No need for bijective b/c this is sufficient.)

Pick basis $B = \{v_1, \ldots, v_n\}$ for V. Pick the dual basis $B^* = \{f_1, \ldots, f_n\}$ for V^*.
Pick basis $B' = \{w_1, \ldots, w_m\}$ for W.

For linear transformations $\mathrm{Hom}(V, W)$, we have a basis $\{e_{ij}\}_{1\le i\le m,\ 1\le j\le n}$
where $e_{ij}(v_k) = \delta_{jk} w_i$ $\quad 1 \le k \le n$.
keep same index ranges

A basis for $V^* \otimes W$ is $\{f_j \otimes w_i\}_{1\le i\le m,\ 1\le j\le n}$

Then $\varphi(f_j \otimes w_i) = T_{f_j, w_i}$, $T_{f_j, w_i}(v_k) = f_j(v_k) w_i = e_{ij}(w)$
Thus $T_{f_i, w_j} = e_{ij}$.

Thus $\varphi(f_j \otimes w_i) = e_{ij}$, meaning φ sends a basis to a basis, meaning φ is an isomorphism. $\square$

Extension of Scalars (& Other Field Stuff)

Let $F \subseteq K$ be a field extension. That is F is a closed subfield of K.

Let V be a vector space over F. Then $K \otimes_F V$ is a vector space over F. Note $\otimes_F = \otimes$ of vector spaces over F

We now define "scalar" multiplication $K \otimes_F V$

$a(b \otimes v) = (ab) \otimes v \quad \forall a, b \in K \quad v \in V$

With this definition $K \otimes_F V$ is a vector space over K & if $\{v_i\}_{i \in I}$ is a basis for V over F then $\{1 \otimes v_i\}_{i \in I}$ is a basis for $K \otimes_F V$ over K. In particular $\dim_K(K \otimes_F V) = \dim_F V$. ← we "don't gain" anything when scalars match

Examples:

Using this we get the following identities

i) $K \otimes_F F^n \cong K^n$ ii) $K \otimes_F F^{m \times n} \cong K^{m \times n}$ iii) $K \otimes_F F[x] \cong K[x]$

Rem:

$\dim_F(K \otimes_F V) = (\dim_F K)(\dim_F V)$ ← We get a larger space but as long as scalars match no more dimension

Chapter 6: Elementary Canonical Forms

How do we represent linear operators? In particular, how do we represent them well?

Def:
Let V be a vector space over field F where $\dim(V) = n < \infty$ w/ linear operator $T \in End(V) = Hom(V,V)$.

T is diagonalizable (or semi-simple) if $\exists B$ basis for V st the matrix of T wrt B $[T]_B$ is a diagonal matrix.

Let $B = \{v_1, \ldots, v_n\}$.
We say $[T]_B = diag(\lambda_1, \ldots, \lambda_n)$ iff

$$T = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

The j-th column of $[T]_B$ is
$[Tv_j]_B$ = coordinate vector of Tv_j wrt B.

Since $[T]_B = \begin{bmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, the j-th column is $\begin{bmatrix} 0 \\ \vdots \\ \lambda_j \\ \vdots \\ 0 \end{bmatrix} = \lambda_j e_j \ \forall j = 1 \ldots n$

Thus $Tv_j = \lambda_j v_j \ \forall j = 1, \ldots, n$.

When you get such a vector like v_j, you call it an eigenvector & λ_j the eigenvalue.

In other words, there exists a basis of only eigenvectors for T. (more later)

Def:
We say $v \in V$ an eigenvector (or characteristic vector) for $T \in End(V)$ w/ eigenvalue λ (& characteristic value λ) (a scalar) iff
$Tv = \lambda v$ & $v \neq 0$.

We exclude 0 b/c then $Tv = \lambda v$ for all scalars λ.

Likewise we say $\lambda \in F$ is an eigenvalue for $T \in End(V)$ w/ eigenvector $v \neq 0$ iff
$Tv = \lambda v$

Def:

The λ-eigenspace of $T \in End(T)$ is $Ker(T - \lambda I)$. More intuitively it is the set of all solutions v to $Tv = \lambda v$ (plus $v = 0$) or the set of all eigenvectors w/ eigenvalue λ.

Note: $det(T) = det([T]_B)$ for $T \in End(V)$ where B basis for V. This is well defined b/c $[T]_B, [T]_{B'}$ similar & determinant of similar matrices are equal.

Thm:

Let $T \in End(V)$ where $\dim(V) < \infty$. Then λ is an eigenvalue for T

$\iff T - \lambda I$ is singular (i.e. not invertible)

$\iff Ker(T - \lambda I) \neq \{0\}$

$\iff det(T - \lambda I) = 0$ ← The way we find all eigen values

& $det(\lambda I - T) = 0$ b/c $det(T - \lambda I) = (-1)^n det(\lambda I - T)$

Def:

i) The characteristic polynomial of T is
$f(\lambda) = det(\lambda I - T)$ ← some write $det(T - \lambda I)$. It may differ by sign

ii) This is indeed a polynomial $f(\lambda) \in F[\lambda]$ b/c it is a sum of linear factors. ← Hand wavy I know

iii) In particular it has degree n where $\dim(V) = n$ & is monic (i.e. its leading coefficient is one). ← Not true for $f(\lambda) = det(T - \lambda I)$

iv) We can explicitly write f as
$f(\lambda) = \lambda^n - tr(T)\lambda^{n-1} + \dots + det(T)$ ← where $\lambda = 0$ i.e. $f(0)$

↑ well defined for similar reasons to $det(T)$ being well defined

Thm:

The roots of the characteristic polynomial (i.e. λ st $det(\lambda I - T) = 0$) are precisely the eigenvalues for T.

This theorem is how we find eigenvalues in practice.

Rem:

Eigenvalue, eigenvector, & all work above applies to matrices $A \in F^{n \times n}$ where you define $T \in End(F^n)$ $T(v) = Av$.

Note that $[T]_B = A$ where B is the standard basis. (Hopefully this is intuitive)

Example:

Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Find the eigenvalues & eigenvectors.

To find the eigenvalues λ, we solve the characteristic polynomial
$f(\lambda) = det(\lambda I - A) = \begin{vmatrix} x & -1 \\ 0 & x \end{vmatrix} = x^2$

Thus 0 is an eigenvalue.

The 0-eigenspace is $Ker(A - 0 \cdot I) = Ker(A) = \{x \in F^2 \mid Ax = 0\}$.
$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_2 = 0$

Thus $Ker(A) = \{\begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \in F\} = span\{e_1\}$. This gives us a unique (up to scalar multiplication) eigenvector $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Note that A is not diagonalizable. This is b/c it has 1 eigenvector e_1, which cannot be a basis for F^2.

Thm:

Let V be a vector space over F w/ $\dim(V) = n < \infty$.
Let $T \in End(V)$ be an isomorphism.

T is diagonalizable (or semi-simple) if $\exists$ basis for V of eigenvectors of T. That is
$\exists$ basis $\{v_1, \dots, v_n\}$ st
$Tv_i = \lambda_i v_i \quad \forall i \in \{1, \dots, n\}$

Thm:

Let $A \in F^{n \times n}$ be an upper (or lower) triangular matrix.

$$A = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

The eigenvalues of A are the entries on the diagonal $\lambda_1, \dots, \lambda_n$. ← it is not clear there are enough (unique) eigenvectors for A to be diagonalizable.

This is true b/c the determinant of a triangular matrix is the product of the entries on the main diagonal.

$$f(x) = \begin{vmatrix} x - \lambda_1 & & * \\ & \ddots & \\ 0 & & x - \lambda_n \end{vmatrix} = (x - \lambda_1) \cdots (x - \lambda_n).$$

Thm:

Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Coro:

If $T \in End(V)$ where $\dim(V) = n < \infty$ & T has n distinct eigenvalues, then T is diagonalizable.

This holds b/c if you have n linearly independent vectors from an n-dimensional space, they form a basis.

Pf:

This is for the above theorem (not corollary). This is continued on the next page.

We use the following Lemma in this proof

Lemma: If $Tv = \lambda v$, then $g(T)v = g(\lambda)v$ for any $g(x) \in F[x]$.

Pf:

$Tv = \lambda v \Rightarrow T^2 v = T(Tv) = T(\lambda v) = \lambda(Tv) = \lambda \lambda v = \lambda^2 v$.

We can continue on like this inductively to give us $T^k v = \lambda^k v \quad \forall k \geq 0$.

This applies for $k=0$ b/c $T^0 = I$, $T^0 v = v = \lambda^0 v$.

By linearity $(\underbrace{\sum_{k=0}^{n} a_k T^k}_{\text{polynomial}}) v = (\underbrace{\sum_{k=0}^{n} a_k \lambda^k}_{\text{polynomial}}) v$. $\square$

Now the main proof.

Let V be a vector space w/ $\dim(V) = n < \infty$.
Let $T \in \mathrm{End}(V)$ be a linear operator on V w/ m distinct eigenvalues. That is $Tv_i = \lambda_i v_i \quad \forall i \in \{1, \dots, m\}$ where $\lambda_i \neq \lambda_j \quad \forall i \neq j$.

Suppose $c_1 v_1 + \dots + c_m v_m = 0$ for some $c_1, \dots, c_m$. We want to show these c's are trivial. That is $c_1 = \dots = c_m = 0$.

Here is one way. Consider repeated applications of T on $c_1 v_1 + \dots + c_m v_m = 0$.

$$
\begin{aligned}
c_1 v_1 + \dots + c_m v_m &= 0 \\
c_1 \lambda_1 v_1 + \dots + c_m \lambda_m v_m &= 0 \\
&\vdots \\
c_1 \lambda_1^{m-1} v_1 + \dots + c_m \lambda_m^{m-1} v_m &= 0
\end{aligned}
$$

This gives us the following square matrix w/ linearly independent rows (& columns) since all λ's are distinct

$$
\left.\begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_m \\ \vdots & & \vdots \\ \lambda_1^{m-1} & \cdots & \lambda_m^{m-1} \end{bmatrix}\right\} \text{Van der Monde matrix. } \det(\dots) = \prod_{i<j} (\lambda_i - \lambda_j)
$$

Since this matrix is linearly independent, it is invertible, meaning the system of linear equations has one solution which is the trivial $c_1 = \dots = c_m = 0$.

Here is another way. Recall the Lagrange interpolation polynomials
$$L_j(x) = \prod_{\substack{i=1 \\ i \neq j}}^{m} \frac{x - \lambda_i}{\lambda_j - \lambda_i} \Rightarrow L_j(v_i) = \delta_{ij}.$$

Now apply $L_j(T)$ to $c_1 v_1 + \dots + c_m v_m = 0$.

$$
\begin{aligned}
0 = L_j(T)(c_1 v_1 + \dots + c_m v_m) &= c_1 L_j(T) v_1 + \dots + c_m L_j(T) v_m \\
&= c_1 L_j(\lambda_1) v_1 + \dots + c_m L_j(\lambda_m) v_m \\
&= c_j v_j = 0
\end{aligned}
$$

Thus $c_j = 0 \quad \forall j$ (b/c eigenvectors $v_j \neq 0$ by definition).
Therefore eigenvectors corresponding to distinct eigenvalues are linearly independent.

Thm:

Let V be a vector space w/ $\dim(V) = n < \infty$.
Let $T \in End(V)$ w/ (distinct) eigenvalues $\lambda_1, \ldots, \lambda_m$ $(m \leq n)$ where $\lambda_i \neq \lambda_j$ $\forall i \neq j$. Note that these eigenvalues may have some multiplicity.

Consider the eigenspaces W_i

$W_i \doteq Ker(T - \lambda_i I) = \lambda_i$ eigenspace where
$d_i = \dim(W_i)$.

In this case we know the following are equivalent statements

T is diagonalizable

$\Leftrightarrow$ $f(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_m)^{d_m}$ ← characteristic polynomial factors

$\Leftrightarrow$ the multiplicity of eigenvalues matches dimension of eigenspaces (no linearly dependent eigenvectors w/ same eigenvalue)

$\Leftrightarrow$ $\sum_{i=1}^{m} d_i = \dim(V) = n$

Pf:

i) Suppose T is diagonalizable. Then there is basis of eigenvectors. Each eigenvector is in one of the eigenspaces W_i. Since all eigenvectors are linearly independent (as basis), the dimensions adds up to at least n $d_1 + \ldots + d_m = n$. Likewise the basis for V is given by a union of bases for $W_1, \ldots, W_m$.

ii) Suppose $d_1 + \ldots + d_m = n$. Then a union of bases $W_1, \ldots, W_m$ provides a basis for V (b/c each eigenspace has only $\{0\}$ in intersection & vectors w/ different eigenvalues are independent). Thus T is diagonal wrt this basis. Here ← applies to (i) to!

$$[T] = \begin{bmatrix} \overbrace{\lambda_1 \ldots \lambda_1}^{d_1} & & & & 0 \\ & \ddots & & & \\ & & \overbrace{\lambda_m \ldots \lambda_m}^{d_m} & & \\ 0 & & & & \end{bmatrix}$$

iii) Thus we can see the characteristic polynomial is
$f(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_m)^{d_m}$.

Rem:

In general, you can factor the characteristic polynomial as
$f(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_m)^{k_m}$.

However

$k_i \geq d_i$ where $d_i = \dim(Ker(T - \lambda_i I))$.

Rem:
If T is diagonalizable w/ (distinct) eigenvalues $\lambda_1, \dots, \lambda_m$ ($\lambda_i \neq \lambda_j \; \forall i \neq j$), then
$(T-\lambda_1 I)\cdots(T-\lambda_m I) = f(T) = 0.$ minimal polynomial that annihilates T

Thm: Cayley-Hamilton Theorem
$f(T)=0$ for any linear operator T where $f(x)$ is the characteristic polynomial. (degree is $\dim V$)

Def:
A polynomial $g(x) \in F[x]$ is an annihilating polynomial of T iff $g(T)=0$.

Pf:
First we show such non-zero polynomials $g(x)$ exist.

Recall $\dim(\mathrm{End}(V)) = n^2$. If you take n^2+1 linear operators, you get linearly dependent operators

Take $I, T, T^2, \dots, T^{n^2}$. These are linearly dependent so there is some linear combination that is zero, which gives us a non-zero annihilating polynomial.

Now we show the Cayley-Hamilton theorem.

A ring is a group w/ multiplication as well

Let $J = \{$annihilating polynomials of $T\} = \{g(x) \in F[x] \mid g(T)=0\}$. We just showed $J \neq \{0\}$. Note that J is closed under $+$ & $\cdot$
i) $\forall g_1, g_2 \in J \Rightarrow g_1 + g_2 \in J$
ii) $\forall g \in J$ & $\forall h \in F[x] \Rightarrow gh \in J$

This means J is an ideal of the ring $F[x]$, denoted $J \trianglelefteq F[x]$.

B/c polynomials are well studied as a ring we have a full description. Thus recall from abstract algebra the following theorem

Thm:
For every non-zero ideal in the ring of polynomials $J \trianglelefteq F[x]$ $J \neq 0$, there exists a unique monic polynomial $p \in J$ of min degree.
← coefficient is 1, so unique to scaling
Then $J = \{ph \mid h \in F[x]\}$ is just the set of multiples of p.
We call p the generator of the ideal J.
· Proved by division w/ remainder (?).

Monic = leading coefficient is one

Now w/ our $J = \{g(x) \in F[x] \mid g(T)=0\}$ the monic generator of this ideal is called the minimal polynomial of T.

Def:
The minimal polynomial of linear operator T is a unique non-zero monic polynomial of minimum degree st $p(T)=0$.

For $g(x) \in F[x]$, we have
$g(T)=0 \Leftrightarrow p(x) \mid g(x)$. ($p(x)$ divides $g(x)$)

Recall that the characteristic polynomial for linear transformation T is
$$f(x) = \det(xI - T).$$
It has roots precisely equal to the eigenvalues of T.

Thm:
Let $p(x)$ be the minimal polynomial for T & $f(x)$ the characteristic polynomial.

$p(x)$ & $f(x)$ have the same roots. (w/ possibly diff multiplicities)

Pf:
We want to show $p(\lambda) = 0$ $(\lambda \in F)$ iff λ is an eigenvalue of T.

Assume $p(\lambda) = 0$ for some $\lambda \in F$. Then $p(x) = (x - \lambda) q(x)$.
B/c p is the minimal polynomial, that is the smallest polynomial st $p(T) = 0$, & q has degree 1 less than p $(\deg(q) = \deg(p) - 1)$, we know $q(T) \neq 0$.
Thus $\exists v \in V$ st $q(T)v = w \neq 0$. Combining this w/ $p(T) = 0$, gives us
$$(T - \lambda I) w = (T - \lambda I) q(T) v = p(T) v = 0.$$
Since $w \neq 0$, w is an eigenvector w/ eigenvalue λ. □

Assume λ is an eigenvalue. Then $Tw = \lambda w$ for some $w \in W$ where $w \neq 0$.
Then $p(T) w = p(\lambda) w$. However, $p(T) = 0$ so $p(T) w = 0$.
Thus $p(\lambda) w = 0$ but since $w \neq 0$, $p(\lambda) = 0$.
Thus $p(\lambda) = 0$. □

Invariants & Subspaces

Def:
A subspace W of V is called <u>T-invariant</u> iff
$$Tw \in W \quad \forall w \in W.$$

This means the restriction of T on W $T|_W$ is still a linear transformation $T|_W : W \to W$, which can be useful for induction.

Example:
A 1-dimensional subspace $W = \text{span}\{w\} = Fw = \{cw \mid c \in F\}$ $(w \neq 0)$ is T-invariant. That means
$Tw = \lambda w$ for some $\lambda \in F$. ← choose λ to lead to eigenvalues, could be any letter

Thus we can see w is an eigenvector of T. & W is thus an eigenspace of T.

Example:

Every eigenspace $Ker(T-\lambda I)$ is T-invariant. Th...

You can see this by letting $w \in Ker(T-\lambda I)$. We know $Tw = 0$ b/c $w \in Ker(T-\lambda I)$.

We know the 0 vector is in all vector spaces, so we know $Tw = 0 \in Ker(T-\lambda I)$.

Example:

Let $S \in End(V)$ be a linear operator that commutes w/ $T \in End(V)$. That is $ST = TS$.

Then $Ker(S)$ & $Im(S)$ are T-invariant.
image (think: range)

First we show $Ker(S)$ is T-invariant. Let $v \in Ker(S)$. Then

$(TS)v = T(Sv) = T(0) = 0$.

Thus we know by $ST = TS$

$0 = (ST)v = S(Tv) = 0$

Thus $Tv \in Ker(S)$ when $v \in Ker(S)$. Therefore $Ker(S)$ is T-invariant

Now we show $Im(S)$ is T-invariant. Let $v \in Im(S)$. Then $\exists v' \in V$ st $v = Sv'$.

Then we can write Tv as

$Tv = T(Sv')$

$= S(Tv')$ b/c S commutes.

Then $Tv = S(Tv')$ (where $Tv' \in V$), so $Tv \in Im(S)$ given $v \in Im(S)$.

Therefore $Im(S)$ is T-invariant.

Example:

If we have such a commuting $S, T \in End(V)$, we know

$S = g(T)$ $\forall g(x) \in F[x]$.

What can we do with T-invariant subspaces?

Let B_w be a basis for W. Extend it to a basis B_v for V. Let $T \in End(V)$.

Let $m = dim(W) \leq n = dim(V)$.

We can represent T as a matrix. We know W is T-invariant iff A_3 is all zero

$$[T]_{B_v} = \begin{array}{cc} \left[\begin{array}{cc} A_1 & A_2 \\ A_3 & A_4 \end{array}\right] & \begin{array}{c} m \\ n-m \end{array} \\ \begin{array}{cc} m & n-m \end{array} & \end{array}$$

Further, we know $[T]_{B_w} = A_1$

In more detail, where $B_w = \{v_1, \ldots, v_m\}$ & $B_v = \{v_1, \ldots, v_m, \ldots, v_n\}$, The j-th column of $[T]_{B_v}$ is $[Tv_j]_{B_v}$.

$$Tv_j = \sum_{i=1}^{n} a_{ij} v_i \Rightarrow A = (a_{ij})$$

Then we know the entries of $[T]_{B_V}$

$$[T]_{B_V} = \left[\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1m} & a_{1m+1} & \cdots & a_{1n} \\ \vdots & & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} & a_{mm+1} & \cdots & a_{mn} \\ \hline a_{m+11} & \cdots & a_{m+1m} & a_{m+1m+1} & \cdots & a_{m+1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} & a_{nm+1} & \cdots & a_{nn} \end{array}\right] = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

Prop: Here are some properties of minimal/characteristic polynomials we know from the determinant.

i) The characteristic polynomial of $T|_W$ divides that of T.

ii) The minimal polynomial of $T|_W$ likewise divides that of T.

Pf:

i) From our earlier work above we know

$$A = [T]_{B_V} = \begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix} \quad \& \quad [T|_W]_{B_W} = A_1.$$

Let $f(x)$ be the characteristic polynomial of T & $g(x)$ be that of $T|_W$

$$f(x) = \det(xI - A) = \det\begin{bmatrix} xI - A_1 & -A_2 \\ 0 & xI - A_4 \end{bmatrix}$$
$$= \det(xI - A_1)\det(xI - A_4).$$
$$= g(x)\det(xI - A_4).$$

Here it is clear that the characteristic polynomial of $T|_W$ $g(x)$ divides that of T $f(x)$.

ii) Let $p(x)$ be the minimal polynomial of T & $q(x)$ be that of $T|_W$.

$$p(T) = 0 \Rightarrow p(T)v = 0 \quad \forall v \in V.$$

Thus since $W \subseteq V$

$$p(T)w = p(T|_W)w = 0 \quad \forall w \in W.$$

Thus $p(T|_W) = 0$.

We know from our theorems on minimal polynomials that for some polynomial p^* $p^*(T^*) = 0$ iff p^* divides T^*'s minimal polynomial.

Thus $q(x)$ divides $p(x)$. □

Rem:

When W is T-invariant, T induces a linear operator $\bar{T}$ on the quotient space V/W defined as
$$\bar{T}(v+W) = Tv + W \quad v \in V.$$

This is well defined b/c $w_1 + W = w_2 + W$ for all $w_1, w_2 \in W$. Since W is T-invariant we know $Tv \in W$ iff $v \in W$.

Rem:

A basis $B_{V/W}$ for V/W is
$$B_{V/W} = \{ v_{m+1} + W, \ldots, v_n + W \}.$$
We remove $v_1, \ldots, v_m$ b/c $v_1, \ldots, v_m \in W$ & $w + W$ is the zero vector for cosets & you don't include the zero vector in a basis.

In this case for $T \in End(V)$ where
$$[T]_{B_V} = \begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix} \quad \& \quad [T]_{B_W} = A_1$$
we know that the matrix for the induced linear operator $B_{V/W}$ is
$$[\bar{T}]_{B_{V/W}} = A_4$$

Thus where we can write the minimal polynomial p_V of V as the product of that for W p_W & of V/W $p_{V/W}$
$$p_V(x) = p_W(x) p_{V/W}(x).$$

Def:

For a T-invariant subspace $W \subseteq V$ & $v \in V$, define <u>T-conductor</u> of v into W as
$$\underline{S(v;W)} = \{ g(x) \in F[x] \mid g(T)v \in W \}.$$

This is an ideal in $F[x]$.

Thus since this is an ideal in $F[x]$ we know there exists a unique monic polynomial of minimum degree that generates it.

Thus $g \in S(v;W) \Leftrightarrow q \mid g$. ($q$ divides g).

Rem:

i) The minimal polynomial $p(x)$ of T is in $S(v;W)$
$$p \in S(v;W).$$
Further $q \mid p$.

1) For the zero subspace
$S(v; \{0\}) = \{ g(x) \in F[x] \mid g(T)v = 0 \}$.

We call this the T-annihilator of v.

Example:

Consider an λ-eigenspace W of V. That is $W = \ker(T - \lambda I)$.
By our properties of eigenspaces, we know W is T-invariant.

Notice that $T|_W = \lambda I_W$ where I_W is the identity operator of W.

The minimal polynomial of $T|_W$ is $x - \lambda \Rightarrow x - \lambda \mid p(x) \Rightarrow p(\lambda) = 0$. b/c $(\lambda - \lambda) \cdots = 0 \cdots = 0$

Thus the characteristic polynomial of $T|_W$ is $(x-\lambda)^d$ where $d = \dim(W)$.
Thus $(x-\lambda)^d \mid f(x)$.

Triangulable Operators

A triangulable operator is an operator that can be expressed as a triangular matrix relative to some basis.

Recall:

A linear operator $T \in End(V)$ is diagonalizable if $\exists$ basis $\mathcal{B}$ for V st $[T]_{\mathcal{B}}$ is diagonal.

Def:

A linear operator $T \in End(V)$ is <u>triangulable</u> if $\exists$ basis $\mathcal{B}$ for V st $[T]_{\mathcal{B}}$ is triangular.

In this triangular case where

$$[T]_{\mathcal{B}} = A = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

This is actually necessary & sufficient (theorem below)

we can easily find the characteristic polynomial $f(x)$ as
$f(x) = (x-\lambda_1)\cdots(x-\lambda_n)$.
Notice that $\lambda_1, \ldots, \lambda_n$ are our eigenvalues.

Thm:

A linear operator $T \in End(V)$ where $\dim(V) < \infty$ is triangulable iff the characteristic polynomial $f(x)$ is a product of linear factors. Likewise for the minimal polynomial $p(x)$.

Def:

A linear operator is called <u>split</u> if its minimal polynomial $p(x)$ can be written as / splits as a product of linear factors.

Coro:

Under a algebraically closed field (e.g. $\mathbb{C}$) which is one where all polynomials can be written as a product of linear factors, all linear operators are split & thus triangulable.

This is one reason why using complex numbers $\mathbb{C}$ is so popular in practice.

Pf: Triangulable & Char./Min. Polynomials pt. 1

Here we prove

(1) $T \in End(V)$ is triangulable

(2) $\Leftrightarrow$ char. polynomial $f(x)$ = product of linear factors

(3) $\Leftrightarrow$ min. polynomial $p(x)$ = product of linear factors

To show $(1) \Leftrightarrow (2) \Leftrightarrow (3)$, we prove $(1) \Rightarrow (2)$, $(2) \Rightarrow (3)$, $(3) \Rightarrow (1)$, thus giving us a circle which shows they are equivalent.

We have already shown $(1) \Rightarrow (2)$. The proof was if $T \in End(V)$ is triangulable then

$$[T] = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \Rightarrow f(x) = det(xI - [T]) = det\begin{bmatrix} x-\lambda_1 & & * \\ & \ddots & \\ 0 & & x-\lambda_n \end{bmatrix}$$

$$= (x-\lambda_1)\cdots(x-\lambda_n).$$

$\underline{(1) \Rightarrow (2)}$

We will show $(2) \Rightarrow (3)$ using Cayley-Hamilton's theorem. However, we haven't proven that yet. As such, we will prove $(3) \Rightarrow (1)$ next & the come back to $(2) \Rightarrow (3)$ after a diversion to prove Cayley-Hamilton's

Now we show $(3) \Rightarrow (1)$. Let $B = \{v_1, \ldots, v_n\}$ be basis for V. Suppose the minimal polynomials $p(x)$ of T is a product of linear factors.

Let $A = (a_{ij})$ be the matrix for T relative to B, so

$$Tv_j = \sum_{i=1}^{n} a_{ij} v_j$$

Note: upper triangular & lower triangular equivalent later

For the matrix A to be upper triangular iff

$$a_{ij} = 0, \quad \forall i > j$$

$$\Leftrightarrow Tv_j = \sum_{i=1}^{j} a_{ij} v_i \in span\{v_1, \ldots, v_j\} = V_j. \quad \forall j \in \{1, \ldots, n\}$$

Then

$$V_0 = \{0\} \subseteq V_1 = span\{v_1\} \subseteq V_2 = span\{v_1, v_2\} \subseteq \ldots \subseteq V_n = V$$

$$dim(V_0) = 0 \quad dim(V_1) = 1 \quad dim(V_2) = 2 \qquad dim(V_n) = n$$

This sequence of subspaces $V_0, \ldots, V_n$ is called a flag. (diversion)

Def:

A flag in vector space V where $dim(V) = n$ is a increasing sequence of subspaces

$$V_0 = \{0\} \subset V_0 \subset \ldots \subset V_{n-1} \subset V_n = V$$

such that

$$dim(V_j) = j \quad \forall j \in \{1, \ldots, n\}.$$

Note: If you have a flag you can always choose basis $B = \{v_1, \dots, v_n\}$ st $\{v_1, \dots, v_j\}$ is basis for V_j. You can prove this inductively by starting w/ the trivial basis $\emptyset$ for V_0 & repeatedly extending it to be a basis for V_{j+1}.

Thm:

A linear operator $T \in End(V)$ is triangulable iff it preserves some flag $V_0 \subset \dots \subset V_n$, that is each V_j is T-invariant.

Alright, we can almost restart our big proof. But first a lemma.

Lemma:

Let T be a split linear operator. Let $W \subset V$ be a T-invariant subspace. Then $\exists v \in V$ where $v \notin W$ w/ an eigenvalue λ st

$(T - \lambda I)v \in W$.

pf:

Pick some $u \in V$ where $u \notin W$ (possible b/c $W \subset V$). Consider the T-conductor of u into W

$S(u; W) = \{g(x) \in F[x] \mid g(T)u \in W\}$.

Let $q(x)$ be a generator of the ideal $S(u;W)$. Then we know $q(x)$ divides the minimal polynomial $p(x)$ $q(x) | p(x)$. This is true b/c $p(x) \in S(u;W)$. We know this b/c $p(T) = 0$ & $0 \in W$ is always true.

Since T is a split linear operator, we know $p(x)$ is a product of linear factors. Since $q(x)$ divides $p(x)$, we know

$q(x) =$ product of linear factors.

Thus $\exists \lambda \in F$ st $q(\lambda) = 0$, b/c $q(x)$ is product of linear factors. Therefore we can factor out this λ to give us

$q(x) = (x - \lambda) h(x)$ for some $h \in F[x]$.

Since $\deg(h) = \deg(q) - 1$, we know $h \notin S(u;W)$. We know this b/c since $q(x)$ is the generator of $S(u;W)$, it has minimal degree for polynomials in $S(u;W)$.

Since $h \notin S(u;W)$, we know $v = h(T)u \notin W$. We take this v & show $(T - \lambda I)v \in W$.

$(T - \lambda I)v$
$= (T - \lambda I)h(T)u$ (definition of v)
$= q(T)u \in W$ (definition of q)
$\Rightarrow q(T)u \in W$ ($q(x) \in S(u;W)$)

We have found a $v \in V$ where $v \notin W$ such that $(T - \lambda I)v \in W$.

Now we can finish $(3) \Rightarrow (1)$.

Pf: Triangulable & Char./Min. Polynomials pt. 2

Continuing our work on $(3) \Rightarrow (1)$, we construct a T-invariant flag by induction.

Start from $V_0 = \{0\}$. V_0 is trivially T-invariant. This is the base case.

Now inductively assume V_j-subspace is T-invariant & w/ basis $\{v_1, \dots, v_j\}$. Then using our lemma we find a $v_{j+1} \in V$ where $v_{j+1} \notin V_j$ such that $(T - \lambda_{j+1} I) v_{j+1} \in V_j$ for some $\lambda_{j+1} \in F$

Thus $Tv_{j+1} \in V_{j+1}$ where $V_{j+1} = V_j + F_{v_{j+1}} = \text{span}\{v_1, \dots, v_j, v_{j+1}\}$.
$\boxed{(3) \Rightarrow (1)}$

Rem:
For every split linear operator T $\exists$ basis $B = \{v_1, \dots, v_n\}$ st

$$[T]_B = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

If we take $V_j = \text{span}\{v_1, \dots, v_j\}$ $j \in \{1, \dots, n\}$, we get
$(T - \lambda_{j+1} I) v_{j+1} \in V_j \Rightarrow (T - \lambda_{j+1} I) v_i \in V_j \quad \forall i \in \{1, \dots, j+1\}$.

We will use this to prove Cayley-Hamilton's theorem.

Thm: Cayley-Hamilton's Theorem (Again)
If you have a linear operator $T \in End(V)$ w/ characteristic polynomial $f(x)$. We know that the following holds for the corresponding matrix polynomial (again called $f(X)$)
$f(T) = 0 \leftarrow 0$ matrix.

Pf:
We will prove this for only split linear operators.
What if the linear operator isn't split? If we extend to an algebraically closed field, then the linear operator ~~must~~ be split b/c we know we can factor the characteristic polynomial b/c we can completely factor all polynomials over an algebraically closed field (e.g. $\mathbb{C}$). We can always extend to an algebraically closed field b/c of topics in abstract algebra I haven't covered yet.

Let $T \in End(V)$ be a split linear operator. Then by our earlier part of the theorem, we know T is triangulable

$$[T]_B = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \quad \text{w/ } B = \{v_1, \dots, v_n\}$$

Thus the characteristic polynomial $f(x)$ is
$f(x) = (x - \lambda_1) \cdots (x - \lambda_n)$

Consider the following where $V_j = \text{span}\{v_1, \dots, v_j\}$.
$(T - \lambda_j I) v_i \in V_{j-1} \Leftrightarrow (T - \lambda_j I) V_i \subseteq V_{j-1} \quad \forall 1 \le i \le j \le n$
must be true, which you can convince yourself of by doing the subtraction.

(We use subspaces rather than vectors for simplicity below.)

Now we want to show $f(T)V = 0$ to show $f(T) = 0$, IR

$$f(T)V = (T-\lambda_1 I)\cdots\underbrace{(T-\lambda_n I)V_{n-1}}_{\subseteq V_{n-1}}$$

$$= (T-\lambda_1 I)\cdots\underbrace{(T-\lambda_{n-1})V_{n-1}}_{\subseteq V_{n-2}}$$

$$= \cdots$$

$$= \underbrace{(T-\lambda_1 I)V_1}_{\subseteq V_0}$$

$$= V_0 = \{0\} = 0$$

Thus $f(T)V = 0$. therefore $f(T) = 0$. □

Now with Cayley-Hamilton's Theorem proven we can finally tackle the last bit of our proof $(2) \Rightarrow (3)$.

<u>Pf</u>: Triangulable & Char./Min. Polynomials pt. 3

We now show $(2) \Rightarrow (3)$. That is we show if the characteristic polynomial $f(x)$ of some linear operator T is split (i.e. factorable), then the minimal polynomial $p(x)$ of T is split.

If $f(x)$ is split then $p(x)$ is split b/c $p(x)$ divides $f(x)$ by Cayley-Hamilton's theorem. $(2) \Rightarrow (3)$.

We therefore conclude that

$T \in End(V)$ is triangulable

$\Leftrightarrow$ char. polynomial $f(x)$ = product of linear factors / $f(x)$ split

$\Leftrightarrow$ min. polynomial $p(x)$ = product of linear factors / $p(x)$ split

Diagonalizable Operators

<u>Recall:</u>

Linear operator $T \in End(V)$ is diagonalizable iff $\exists$ basis $B = \{v_1, \ldots, v_n\}$ st

$$[T]_B = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

By the definition of linear operators being applied wrt a basis, we know

$T v_j = \lambda_j v_j \quad \forall j \in \{1, \ldots, n\}$

Again by doing simple matrix subtraction we get $(T - \lambda_j I) v_j = 0$.

With such a T we know the characteristic polynomial $f(x)$ is

$f(x) = (x-\lambda_1)\cdots(x-\lambda_n)$

& by Cayley-Hamilton's theorem we get

$f(T) = 0$

& also by simple application of the definition of the characteristic polynomial $f(x)$ we get

$f(T)v_j = 0 \quad \forall j \in \{1, \ldots, n\} \Rightarrow f(T) = 0$

Suppose $\lambda_1, \ldots, \lambda_k$ are distinct ($\lambda_i \neq \lambda_j$ for $1 \leq i < j \leq k$). Reorder the list such that all "2nd+ copies" are after k.

Take the minimal polynomial $p(x)$ to be

$$p(x) = (x-\lambda_1)\cdots(x-\lambda_k)$$

Since we already have all unique diagonal entries, we know

$$p(T) = 0 \quad b/c \quad p(T)v_j = 0 \quad \forall j \in \{1, \dots, n\}$$

We know this $p(x)$ is the minimal polynomial b/c all roots are distinct & if one root did not exist then $p(T) \neq 0$.

Thm:
A linear operator $T \in End(V)$ is diagonalizable if & only if its minimal polynomial $p(x)$ is a product of distinct linear factors.

Pf:
Our earlier proof proves

T diagonalizable $\Rightarrow$ $p(x)$ has distinct roots. $\square$

Now we prove

$p(x)$ has distinct roots $\Rightarrow$ T diagonalizable

Suppose $p(x) = (x-\lambda_1)\cdots(x-\lambda_k)$ where $\lambda_i \neq \lambda_j$ for $1 \leq i < j \leq k \leq n = \dim(V)$

Consider all eigenspace (W_j is λ_j-eigenspace)

$$W_j = Ker(T - \lambda_j I).$$

Take $W = W_1 + \cdots + W_k$ = span of eigenvectors.

We want to show $W = V$ b/c then we have V is spanned by distinct eigenvectors of T, meaning T has n distinct eigenvectors, meaning T is diagonalizable (take basis of eigenvectors).

Assume for contradiction that $W \neq V$. By our earlier lemma we know $\exists\, v \in V$ where $v \notin W$ & an eigenvalue λ st

$$(T - \lambda I)v \in W.$$

B/c $(T - \lambda I)v = w \in W$ & W is spanned by eigenvectors, we know we can write this as a sum (or linear combination) of eigenvectors

$$(T - \lambda I)v = w = w_1 + \cdots + w_k \quad \text{where } w_1, \dots, w_k \text{ are eigenvectors}$$

$$\Leftrightarrow w_j \in W_j \ \& \ Tw_j = \lambda_j w_j.$$

Likewise $\lambda = \lambda_i$ for some $i \in \{1, \dots, k\}$.

Now we can write the min. polynomial $p(x)$ as
$p(x) = (x-\lambda_i)h(x)$ where $h(x) = \prod_{i \neq j}(x-\lambda_j)$

Thus using $w_i = (T-\lambda_i I)v$, we get
$h(T)w = h(T)(T-\lambda_i I)v = p(T)v = 0$

Likewise since $w = w_1 + \cdots + w_k$ we get
$h(T)w = h(T)w_1 + \cdots + h(T)w_k$
$= h(\lambda_1)w_1 + \cdots + h(\lambda_k)w_k$ (b/c $f(T)v = f(\lambda)v$ where λ eigenvalue for v under T)
$= \underbrace{h(\lambda_i)}_{\neq 0} w_i$ (definition of h

Since $h(T)w = h(\lambda_i)w_i = 0$ & $h(\lambda_i) \neq 0$, we know $w_i = 0$.

Now consider $u = v - \sum_{j \neq i} \frac{w_j}{\lambda_j - \lambda_i}$

Now apply $(T-\lambda_i I)$ to u to get
$(T-\lambda_i I)u = (T-\lambda_i I)v - \sum_{j \neq i} \frac{(T-\lambda_i I)w_j}{(\lambda_j - \lambda_i)}$ replace T w/ eigenvalue
$= (T-\lambda_i I)v - \sum_{j \neq i} \frac{\cancel{(\lambda_j - \lambda_i)}w_j}{\cancel{(\lambda_j - \lambda_i)}}$
$= (T-\lambda_i I)v - w$
$= w - w$
$= 0$

B/c $(T-\lambda_i I)u = 0$, we know $u \in W_i$. However, that would mean $v \in W_i$ b/c of closure under vector addition.

This is a contradiction so $W = V$, V is spanned by $k=n$ distinct eigenvectors, & thus T is diagonalizable.

Coro:

If T is diagonalizable & W is a T-invariant subspace, then $T|_W$ is also diagonalizable.

Pf:

The minimal polynomial of $T|_W$ divides the minimal polynomial of T.

Thus if T's min. polynomial is a product of linear factors then so is $T|_W$'s.

Simultaneous Diagonalization

Def

A set of linear operators $\mathcal{F} \subset End(V)$ can be simultaneously diagonalized if $\exists$ basis B for V st $[T]_B$ is diagonal for all $T \in \mathcal{F}$.

Recall that all diagonal matrices commute w/ each other.

Thm:

A set $\mathcal{F} \subset End(V)$ of linear operators is simultaneously diagonalizable iff

i) Every $T \in \mathcal{F}$ is diagonalizable

&

ii) $TS = ST \quad \forall T, S \in \mathcal{F}$.

Pf: $\Rightarrow$

Suppose $B = \{v_1, \ldots, v_n\}$ such that it diagonalizes all linear operators $T \in \mathcal{F}$ (i.e. $[T]_B$ diagonal $\forall T \in \mathcal{F}$).

Then for every $T \in \mathcal{F}$, B is a basis of eigenvectors, that is $Tv_i = \lambda_i v_i \quad \forall i \in \{1, \ldots, n\}$. (Note that λ_i depends on T.)

Take some other $S \in \mathcal{F}$. Similarly,

$Sv_i = \mu_i v_i \quad \forall i \in \{1, \ldots, n\}$.

We now show T & S commute by considering

$(TS)v_i = T(Sv_i) = T(\mu_i v_i) = \mu_i T(v_i) = \mu_i \lambda_i v_i$ &

$(ST)v_i = S(Tv_i) = S(\lambda_i v_i) = \lambda_i S(v_i) = \lambda_i \mu_i v_i$.

Since $(TS)v_i = (ST)v_i$, we know TS & ST agree on a basis. Thus $TS = ST \quad \forall T, S \in \mathcal{F}$ (ii)

Trivially, we know every $T \in \mathcal{F}$ is diagonalizable b/c it is diagonalized by B. (i)

We've thus proven the forward direction.

Pf: $\Leftarrow$

Suppose every $T \in \mathcal{F}$ is diagonalizable & $TS = ST \quad \forall T, S \in \mathcal{F}$.

To prove that $\mathcal{F}$ is simultaneously diagonalizable we use induction.

Pick some $T \in \mathcal{F}$. Since it is diagonalizable it has an eigenvalue $\lambda \in F$ & a corresponding λ-eigenspace $W = Ker(T - \lambda I)$.

In the case where $W = V$, then $T = \lambda I$, which is trivially diagonalizable in every basis.

In the case where $W \subset V$ is a proper subspace, then W is S-invariant $\forall S \in \mathcal{F}$. We know this b/c the operators commute.

Let us briefly show this. Let $w \in W = \ker(T - \lambda I)$

$Tw = 0$
$\Rightarrow STw = 0$
$\Rightarrow TSw = 0$
$\Rightarrow Sw \in \ker(T - \lambda I) = W$
$\Rightarrow W$ is S-invariant

Consider the restrictions $S|_W$. Then $S|_W$ is diagonalizable b/c W is S-invariant & S is diagonalizable. (earlier proof)

Moreover, since all $S, R \in \mathcal{F}$ commute & are invariant on W, $S|_W$ & $R|_W$ commute $\forall S, R \in \mathcal{F}$.

By induction (or really recursion), there is a basis B_W for W st $[S]_{B_W}$ is diagonal $\forall S \in \mathcal{F}$.
Essentially, we keep going until we hit the trivial space. At each level we get an eigenvector $w \in W$.

Do this for every eigenspace of T. Since T is diagonalizable, when we union all these bases, we get a basis for V in which all $S \in \mathcal{F}$ is diagonal.

□

Rem:

There is a similar result for simultaneous triangulation.
That is every set of commuting, triangulable/split linear operators can be made simultaneously triangular wrt to some basis.

Note: Commuting is sufficient but not necessary.

Direct Sums & Projections

Def:

Let $W_1, \ldots, W_k$ be subspaces of a vector space V. The sum of these subspaces is defined as

could be ∞-dimensional

$$\sum_i^k W_i = W_1 + \ldots + W_k = \{w_1 + \ldots + w_k \mid w_i \in W_i\} = \operatorname{span}(W_1 \cup \ldots \cup W_k).$$

We know $\sum_i W_i$ is a subspace of V (by closure)

Def:

Let $W = W_1 + \ldots + W_k$ for some subspaces $W_1, \ldots, W_k$ of vector space V.
We say W is a direct sum denoted $W = W_1 \oplus \ldots \oplus W_k$ iff there is a unique way to write

$w = w_1 + \ldots + w_k$ where $w_i \in W_i$)

for all $w \in W$.

Equivalently iff $W_1, \ldots, W_k$ are independent ~~in that $W_i \cap W_j = \{0\}$ $i \neq j$~~ or ~~likewise~~

$w_1 + \ldots + w_k = 0$ where $w_i \in W_i$ iff $w_1 = \ldots = w_k = 0$.

or likewise

$\{w_1, \ldots, w_k\}$ is linearly independent $\forall w_i \in W_i$ where $w_i \neq 0$.

Example:

Consider $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$. Let $W_i = \text{span}\{v_i\} = \mathbb{F}v_i$
Let $W_i = \text{span}\{v_i\} = \mathbb{F}v_i$. Then $\mathcal{B}$ is a basis if & only if $V = W_1 \oplus \ldots \oplus W_n$

Example:

Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues $T \in \text{End}(V)$.
Consider the eigenspaces $W_i = \ker(T - \lambda_i I)$.

Recall that if you have eigenvectors $v_1, \ldots, v_k$ w/ eigenvalues $\lambda_1, \ldots, \lambda_k$ where all the eigenvalues are distinct, then the vectors are linearly independent.

Using this we know $W = W_1 \oplus \ldots \oplus W_k$.

Note that $W = V$ if & only if T is diagonalizable (i.e. $k = \dim(V)$).

Example:

Consider $V = \mathbb{F}^{n \times n}$.
Let $W_1 = \{A \in V \mid A \text{ is symmetric} \Leftrightarrow A^t = A\}$ } kinda like
Let $W_2 = \{A \in V \mid A \text{ is skew-symmetric} \Leftrightarrow A^t = -A\}$ } even & odd

We show $V = W_1 \oplus W_2$.

Take some matrix $A \in V$. We want to write $A = B + C$ where $B \in W_1 \Leftrightarrow B^t = B$ & $C \in W_2 \Leftrightarrow C^t = -C$. We use these identities along w/ the properties of transpose to give us

$$A = B + C$$
$$A^t = B - C \longleftarrow (A^t = B^t + C^t)$$

We can then use elimination to give us

$$B = \tfrac{1}{2}(A + A^t)$$
$$C = \tfrac{1}{2}(A - A^t)$$

This assumes we can divide by 2, that is the characteristic of the field isn't two $\text{char}(\mathbb{F}) \neq 2$.

We thus have shown how to write any matrix $A \in V$ as a sum of $B \in W_1$ & $C \in W_2$. By the definition of W_1 & W_2 we know $W_1 \cap W_2 = \{0\}$, so W_1 & W_2 are independent. (Work below)

Thus we conclude $V = W_1 \oplus W_2$.

We again show that W_1 & W_2 are independent.

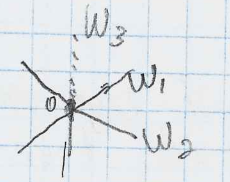
Suppose $A = B + C = B' + C'$ for some $B, B' \in W_1$ & $C, C' \in W_2$. We show $B = B'$ & $C = C'$.
Then $B - B' = C - C'$. Using the transpose we get

$$B - B' = C - C' \Rightarrow (B - B')^T = (C - C')^T \Rightarrow B - B' = C + C'$$

Therefore $B - B' = C - C' = 0$, so $B = B'$ & $C = C'$.
Thus W_1 & W_2 are independent.

<u>Example:</u>

Consider $V = \mathbb{R}^3$ w/ 3 lines W_1, W_2, W_3. We know $\dim(W_i) = 1$ & in particular $W_i = F w_i$ where $w_i \in V$ & $w_i \neq 0$.



Graph of 3 independent lines W_1, W_2, W_3.

$W_1 \cap W_2 = \{0\}$
$\Leftrightarrow \{w_1, w_2\}$ are linearly independent
$\Leftrightarrow W_1 + W_2 = W_1 \oplus W_2$ is direct
$\Leftrightarrow W_1$ & W_2 span a plane

Note that knowing every pair of lines is independent is not sufficient to show $\{W_1, W_2, W_3\}$ are independent b/c they may all fall w/in the same plane. That is

$$W_1 \oplus W_2 \text{ \& } W_1 \oplus W_3 \text{ \& } W_2 \oplus W_3 \not\Rightarrow W_1 \oplus W_2 \oplus W_3.$$

However, the converse is true

$$W_1 \oplus W_2 \oplus W_3 \Rightarrow W_1 \oplus W_2 \text{ \& } W_1 \oplus W_3 \text{ \& } W_2 \oplus W_3.$$

Note further than

$$W_1 \oplus W_2 \oplus W_3 \Leftrightarrow \{w_1, w_2, w_3\} \text{ basis for } V$$

<u>Rem:</u>

$\oplus$ is commutative & associative like you'd expect.

Thus $W_1 \oplus W_2 \oplus W_3 = (W_1 \oplus W_2) \oplus W_3$

<u>Thm:</u>

A sum $W_1 + \dots + W_k$ is direct if & only if

$$W_j \cap (W_1 + \dots + W_{j-1}) = \{0\} \quad \forall j = 2 \dots k \Leftrightarrow W_j \cap \left(\sum_{i \neq j} W_i\right) = \{0\} \quad \forall j = 2, \dots, k$$

<u>Pf:</u>

This is an application of $W_1 \oplus W_2 \oplus W_3 = (W_1 \oplus W_2) \oplus W_3$.

This condition is equivalent to

$$W_1 \oplus W_2, \quad (W_1 \oplus W_2) \oplus W_3, \quad ((W_1 \oplus W_2) \oplus W_3) \oplus W_4, \text{ etc.}$$

If you use induction the proof falls as such.

<u>Notation:</u>

$$W = W_1 \oplus \dots \oplus W_k = \bigoplus_{i=1}^{k} W_i$$

I also like $a_1 + \dots + a_k = \overset{k}{\underset{i=1}{+}} a_i$

Thm:

Let B_i be a basis for W_i $\forall i \in \{1, \dots, k\}$ & $W = W_1 + \dots + W_k$. Then

Then $W = W_1 \oplus \dots \oplus W_k$ if & only if
$B_i \cap B_j = \emptyset$ $\forall i \neq j$ & $B_1 \cup \dots B_k$ basis for W
which is equivalent to
the disjoint union $\bigsqcup_{i=1}^{k} B_i$ is a basis for W.

Coro:

Let V & W be independent spaces.
$\dim(V \oplus W) = \dim(V) + \dim(W)$ ⟵ from disjoint union of basis is basis

Likewise $W = \bigoplus_{i=1}^{k} W_i$
$\dim(\bigoplus_{i=1}^{k} W_i) = \sum_{i=1}^{k} \dim(W_i)$

Pf: $\Rightarrow$

Suppose $W = W_1 \oplus \dots \oplus W_k$. We show the B_i's form a disjoint basis for W.
We show the bases are disjoint.
Suppose $v \in B_i \cap B_j$ for some $i \neq j$. Then $v \in W_i \cap W_j$. However W_i & W_j is independent so $W_i \cap W_j = \{0\}$. Therefore $v = 0$. However no basis contains 0. This is a contradiction so the basis is disjoint.

We show the bases are a (disjoint) union for W.
Take $w \in W$. Then b/c W is a direct sum $\exists!$
$\exists!$ $w = w_1 + \dots + w_k$ $(w_i \in W_i)$
Since $w_i \in W_i$, we know $w_i \in \text{span}(B_i)$. Thus b/c each $w_i \in W_i$, we know the sum w is in span(B): $w \in \text{span}(B)$.

Now we show linear independence. For clarity index B_i by J_i. Then B is indexed by $J = \bigsqcup_{i=1}^{k} J_i$

Take $\sum_{j \in J} c_j v_j = 0$ where $B_i = \{v_j\}_{j \in J_i}$. We show $c_j = 0$ $\forall j \in J$.

We split this sum to basis for the corresponding subspaces.

$$\sum_{j \in J} c_j v_j = (\underbrace{\sum_{j \in J_1} c_j v_j}_{\in W_1}) + \dots + (\underbrace{\sum_{j \in J_k} c_j v_j}_{\in W_k})$$

b/c W is a direct sum we know this sum of $\sum_{j\in J_i} c_j v_j \in W_i$ must be unique. Therefore
$$\sum_{j\in J_i} c_j v_j = 0 \quad \forall i \in \{1,\dots,k\}.$$

Thus $c_j = 0$ $\forall j \in J$ & B is linearly independent.

Thus the disjoint union B forms a basis for W. □

Pf: $\Leftarrow$

Suppose the disjoint union $B = \bigsqcup_{i=1}^{k} B_i$ is a basis for $W = \sum_{i=1}^{k} W_i$ where B_i is a basis for W_i for all $i = 1,\dots,k$. We claim
$$W = \mathrm{span}(B) = \bigoplus_{i=1}^{k} W_i.$$

We again index the B's the same.

Take some $w \in W$. We know there is a unique $w = \sum_{j\in J} c_j v_j$ b/c B is a basis for W.

Since B is a disjoint union of the B_i's we know there is a unique way to write
$$w = \sum_{j\in J} c_j v_j = \underbrace{\Big(\sum_{j\in J_1} c_j v_j\Big)}_{\in W_1} + \dots + \underbrace{\Big(\sum_{j\in J_k} c_j v_j\Big)}_{\in W_k}$$

B/c these summands are unique, we conclude $W = W_1 \oplus \dots \oplus W_k$.

Projections

Let's start w/ a familar example.

Example:

Consider $V = \mathbb{R}^2 = W_1 \oplus W_2$ where $\overset{x\text{-axis}}{W_1 = \mathbb{R}e_1}$ & $\overset{y\text{-axis}}{W_2 = \mathbb{R}e_2}$ where $e_1 = (1,0)$ & $e_2 = (0,1)$.

$\forall v = (x,y) \in \mathbb{R}^2$ $v = \underbrace{xe_1}_{\in W_1} + \underbrace{ye_2}_{\in W_2}$. Further this sum is unique.

We get our classic projections
$$v = (x,y) = \underbrace{(x,0)}_{\substack{\text{Projection Onto} \\ x\text{-axis} \,/\, W_1}} + \underbrace{(0,y)}_{\substack{\text{Projection Onto} \\ y\text{-axis} \,/\, W_2}}$$

We can generalize this to any lines. Further we can do any vector space which is a direct sum of subspaces.

Def:

Let $V = W_1 \oplus \dots \oplus W_k$. Then $\forall v \in V$ $\exists! v = w_1 + \dots + w_k$ where $w_i \in W_i$.

Define the <u>projection of v onto W_i</u> as
$$E_i(v) = w_i.$$

Thm:

$E_1, \dots, E_k$ are linear operators on V w/ the properties

i) $E_i^2 = E_i$ or E_i is idempotent $\forall i \in \{1,\dots,k\}$

ii) $E_i \circ E_j = 0$ $\forall i \neq j$ or likewise $E_i E_j = [0]$ $\forall i \neq j$ (as matrix)

iii) $E_1 + \dots + E_k = I$ when E_i is defined as matrix wrt any basis
iv) $Im(E_i) = W_i$ (by definition)

Pf:

E_i is a linear operator on V. Let $v, v' \in V$ & $c \in F$.
B/c $v, v' \in V$ we know that uniquely
$v = w_1 + \dots + w_k$ & $v' = w_1' + \dots + w_k'$

$E_i(v+v') = w_i + w_i' = E(v) + E(v')$
&
$E_i(cv) = cw_i = cE(v)$. □

Thm:
Conversely if we have linear operators $E_1, \dots, E_k$ that satisfy the previous theorem's i, ii, & iii then
$V = W_1 \oplus \dots \oplus W_k$ where $W_i = Im(E_i)$.

Pf:
Suppose we have some such linear operators $E_1, \dots, E_k$ satisfying (i), (ii), & (iii).

Given property (iii), we write every $v \in V$ as
$$v = Iv = (E_1 + \dots + E_k)v = \underbrace{E_1 v}_{\in W_1} + \dots + \underbrace{E_k v}_{\in W_k}.$$
Thus we have a sum $V = W_1 + \dots + W_k$. We show this sum is direct by using (i) & (ii).

Consider $v = w_1 + \dots + w_k$. We show this is unique. B/c $Im(E_i) = W_i$, we know $\forall w_i \in W_i$ $w_i = E_i u_i$. We write
$v = w_1 + \dots + w_k = E_1 u_1 + \dots + E_k u_k$

We apply E_i & use that
$E_i w_j = E_i E_j u_j = 0 \quad \forall i \neq j$ by (ii) &
$E_i w_i = E_i E_i u_i = E_i u_i = w_i$ by (i)
Thus $w_i = E_i v$ is uniquely determined from v.

Thus the sum is unique & $V = W_1 \oplus \dots \oplus W_k$.

Example:

Consider $\mathbb{R}^2 = \mathbb{R}e_1 \oplus \mathbb{R}e_2$ where $e_1=(1,0)$ & $e_2=(0,1)$.
Then $\forall v \in \mathbb{R}^2$

$$v=(x,y)=(x,0)+(y,0)=xe_1+ye_2$$

$$\Rightarrow \begin{cases} E_1 v = E_1(x,y)=(x,0)=xe_1 \\ E_2 v = E_2(x,y)=(0,y)=ye_2 \end{cases}$$

Example:

Consider $\mathbb{R}^2 = \mathbb{R}e_1 \oplus \mathbb{R}(e_1+e_2)$.
Then $\forall v \in \mathbb{R}^2$

$$v=(x,y)=(x-y,0)+(y,y)$$

$$\Rightarrow \begin{cases} E_1 v = E_1(x,y)=(x-y,0) \\ E_2 v = E_2(x,y)=(y,y) \end{cases}$$

Notice, that the spaces effect the other projections, even if the projecting space doesn't change.

Def:

A $\underset{E^2=E,}{\underline{\text{projection}}}$ on a vector space V is a linear operator E such that that is an idempotent linear operator.

Lemma:

Let E be a projection operator on V. Let $\overbrace{R=Im(E)}^{\text{range}}$ & $\overbrace{N=Ker(E)}^{\text{null space}}$.
R & N are subspaces of V. then we know

i) $V = R \oplus N$ (like a rank nullity theorem),

ii) $v = Ev + (I-E)v \quad \forall v \in V$. That is $(I-E)v \in N$,

iii) $E|_R = I$ & $E|_N = 0$.
this is <u>why</u> it's idempotent

iv) $I-E$ is a projection onto N. That is $Im(I-E)=N$ & $Ker(I-E)=R$
This comes from solving the above equations.

Example:

Consider $V=F^{n\times n}$ & $T \in End(V)$ where $T(A)=A^t$.

Here you can easily see $T^2=I$ b/c $(A^t)^t = A$. From this we get

$$(T-I)(T+I)=0$$

which means the minimal polynomial of T divides x^2-1, so the eigenvalues are 1 & -1.

Going further w/ these eigenvalues we see the eigenspaces are the symmetric matrices $(\lambda=1)$ & skew symmetric matrices $(\lambda=-1)$.

$$W_1 = \{A \in V \mid A^t = A\} = Ker(T-I),$$

$$W_2 = \{A \in V \mid A^t = -A\} = Ker(T+I).$$

We earlier showed you can write every matrix as the sum of a symmetric matrix & skew symmetric matrix by doing

$$A = \tfrac{1}{2}(A+A^t) + \tfrac{1}{2}(A-A^t).$$

Thus we know $V=W_1 \oplus W_2$. Using this fact we can consider the projections onto symmetric & skew symmetric matrices by

$$E_1 = \tfrac{1}{2}(I+T) \quad \& \quad E_2 = \tfrac{1}{2}(I-T).$$

We can generalize this!

Prop:

Let V be a vector space over $\mathbb{C}$ (including ∞ dimensional). Let T be a linear operator of finite order, that is $T \in End(V)$ where $T^N = I$ for some $N \geq 1$.

Then the eigenvalues of T are the N-th roots of 1 (unity) & $V = W_0 \oplus \ldots \oplus W_{N-1}$ where $W_j = Ker(T - e^{2\pi i j/N} I)$ $0 \leq j \leq N-1$.
The corresponding projections are

$$E_j = 1/N \sum_{k=0}^{N-1} e^{-2\pi i j k/N} T^k.$$ ← kinda like averaging

In particular T is diagonalizable b/c it has N eigenvalues.

Invariant Direct Sums

These are direct sums of subspaces that are all invariant under some linear Operator. It is important for induction.

Thm:

A direct sum $V = W_1 \oplus \ldots \oplus W_k$ is T-invariant if & only if each subspace W_i is T-invariant, that is

$Tw_i \in W_i \quad \forall w_i \in W_i$.

This gives us the restriction operators

$T_i = T|_{W_i} \in End(W_i)$.

Using this we can write T for V. Take some $v \in V$.
B/C $V = W_1 \oplus \ldots \oplus W_k$, $\exists!\ v = w_1 + \ldots + w_k$. Thus

$\underline{Tv = Tw_1 + \ldots + Tw_k = T_1 w_1 + \ldots + T_k w_k}$

Using this definition we write $T = T_1 \oplus \ldots \oplus T_k = (T_1, \ldots, T_k)$. This gives us a (meta) linear map $T \to T_1 \oplus \ldots \oplus T_k$ which is in $End(W_1) \oplus \ldots \oplus End(W_k)$.

Conversely, any k-tuple $(T_1, \ldots, T_k)$ where $T_i \in End(W_i)$ defines a $T \in End(V)$ such that $V = W_1 \oplus \ldots \oplus W_k$ is T-invariant.

Lemma:

Let B_i be a basis for W_i $\forall i = 1, \ldots, k$. Then $B = \coprod_{i=1}^{k} B_i$ is a basis for V.

If a direct sum is T-invariant then

$$[T]_B = \begin{bmatrix} A_1 & \ldots & 0 \\ 0 & & A_k \end{bmatrix}$$ where $A_i = [T_i]_{B_i}$

Note: $[T]_B$ is block diagonal

Def:
The direct sum of matrices $A_1, \ldots, A_k$ is defined as the block-diagonal matrix
$$A_1 \oplus \cdots \oplus A_k = \begin{bmatrix} A_1 & \cdots & 0 \\ 0 & & A_k \end{bmatrix}$$

Recall that given $V = W_1 \oplus \cdots \oplus W_k$ we get projection operators $E_1, \ldots, E_k \in End(V)$ by
$$v = w_1 + \cdots + w_k = E_1 v + \cdots + E_k v, \text{ so } \quad E_i v = w_i \ \forall i = 1..k.$$

These projections satisfy the necessary properties
i) $E_i^2 = E_i \quad \forall i$
ii) $E_i E_j = 0 \quad \forall i \neq j$
iii) $E_1 + \cdots + E_k = I$
iv) $W_i = Im(E_i) \quad \forall i$

Recall further the direct sum decomposition $V = W_1 \oplus \cdots \oplus W_k$ is T-invariant iff
$\left.\begin{array}{l} \Leftrightarrow T w_i \in W_i \quad \forall w_i \in W_i \\ \Leftrightarrow w_i = E_i v \quad T E_i v \in W_i = Im(E_i) \\ \Leftrightarrow T E_i v = E_i u \text{ for some } u \in V \quad \forall v \in V \end{array}\right\}$ This doesn't go anywhere

More clearly
$$v = w_1 + \cdots + w_k \quad \text{where } w_i = E_i v \in W_i$$
$$\Rightarrow Tv = \underbrace{Tw_1}_{\in W_1} + \cdots + \underbrace{Tw_k}_{\in W_k} \Big\} \text{ b/c subspaces } T\text{-invariant}$$

Thus b/c W_i is T-invariant $Tw_i = TE_i v = E_i T v$. This is actually sufficient to show W_i is T-invariant

Thm:
A direct sum is T-invariant iff T commutes w/ the corresponding projection operators $E_1, \ldots, E_k$.

If you recall, all diagonal matrices commute. This gives some nice properties.

Diagonalizable Operators

Suppose T is diagonalizable w/ distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. This gives us eigenspaces
$$W_i = Ker(T - \lambda_i I).$$

B/c T is diagonalizable we know $V = W_1 \oplus \cdots \oplus W_k$. Further
$$T|_{W_i} = \lambda_i I.$$

In terms of matrices, the matrix of the operator is block diagonal
$$[T] = \begin{bmatrix} A_1 & \cdots & 0 \\ 0 & & A_k \end{bmatrix}$$
where
$$A_i = [T_i] = \begin{bmatrix} \lambda_i & \cdots & 0 \\ 0 & & \lambda_i \end{bmatrix}.$$

Thus the whole matrix is diagonal (duh).

This gives us minimal polynomial $p(x)$
$$p(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

B/c $V = W_1 \oplus \dots \oplus W_k$, we know $\exists!\ v = w_1 + \dots + w_k$. B/c the W_i's are eigenspaces, we can simplify the calculator of Tv to
$Tv = Tw_1 + \dots + Tw_k = \lambda_1 w_1 + \dots + \lambda_k w_k$.

We can generalize this to polynomials of the linear operator so for all $g(x) \in F[x]$

$$g(T)v = g(T)w_1 + \dots + g(T)w_k$$
$$= g(\lambda_1)w_1 + \dots + g(\lambda_k)w_k.$$

We want to find the polynomials g_i such that
$g_i(\lambda_i) = 1\ \forall i$ & $g_i(\lambda_j) = 0\ \forall i \neq j$.

Recall:

To find these g_i's we want the Lagrange interpolation polynomials

$$L_j(x) = \prod_{\substack{1 \le i \le k \\ i \neq j}} \frac{x - \lambda_i}{\lambda_j - \lambda_i}$$

By this definition of L_j,
$$L_j(\lambda_i) = \delta_{ij}$$

Applying the Lagrange polynomials to our problem

$$L_j(T)v = L_j(T)w_1 + \dots + L_j(T)w_k$$
$$= L_j(\lambda_1)w_1 + \dots + L_j(\lambda_k)w_k$$
$$= w_j = E_j v$$

Thus we have explicitly found the projections of $V = \underbrace{W_1 \oplus \dots \oplus W_k}_{\text{eigenspaces of } T}$, so
$$E_j = L_j(T).$$

Thm:

If the minimal polynomial of $T \in \mathrm{End}(V)$ is
$$p(x) = (x - \lambda_1)\cdots(x - \lambda_k) \quad w/\ \lambda_i \neq \lambda_j\ \forall i \neq j$$
then T is diagonalizable w/ eigenspaces
$$W_i = \ker(T - \lambda_i I).$$
So $V = W_1 \oplus \dots \oplus W_k$ w/ projections
$$E_i = L_i(T).$$

Pf:

To prove $V = W_1 \oplus \dots \oplus W_k$, we check that $E_i = L_i(T)$ satisfy the earlier properties

i) $E_i^2 = E_i$

ii) $E_i E_j = 0 \quad \forall i \neq j$

iii) $E_1 + \dots + E_k = 0$

we do this only using that $p(T)=0$.

First we show (i), that is $E_i^2 = L_i(T)^2 = E_i = L_i(T)$. Equivalently we show $L_i(T)^2 - L_i(T) = L_i(T)(L_i(T)-1) = 0$.

The polynomial $L_i(x)(L_i(x)-1)=0$ for $x=\lambda_j$ $\forall j=1,\dots,k$. This is b/c
$L_i(\lambda_j)=0$ $\forall i\neq j$ & $L_i(\lambda_i)=1$
B/c $L_i(x)[L_i(x)-1)$ has roots $\lambda_1,\dots,\lambda_k$, then we know the minimal polynomial $p(x)$ divides $L_i(x)(L_i(x)-1)$.

B/c $p(x)$ divides $L_i(x)(L_i(x)-1)$ & $p(T)=0$, we know $L_i(T)(L_i(T)-1)=0 \Leftrightarrow L_i(T)^2 = L_i(T)$.
We do this for all i.

Now we show (ii). We show $E_iE_j=0$ $\forall i\neq j$, or equivalently $L_i(T)L_j(T)=0$.
We know
$L_i(x)L_j(x)=0$ for $x=\lambda_1,\dots,\lambda_k$
b/c at least one will be 0. B/c of this we know $p(x)$ divides $L_i(x)L_j(x)$.
Since $p(T)=0$, we know $L_i(T)L_j(T)=0$ $\forall i\neq j$.

Now we show (iii). We show $E_1+\dots+E_k=I$. Consider $L_1(x)+\dots+L_k(x)-1$.
We know
$L_1(x)+\dots+L_k(x)-1=0$ for $x=\lambda_1,\dots,\lambda_k$.
Thus $L_1(x)+\dots+L_k(x)-1$ has roots $\lambda_1,\dots,\lambda_k$. Thus $p(x)$ divides it. Since $p(T)=0$, we know
$L_1(T)+\dots+L_k(T)=I$.

From (i), (ii), & (iii), it follows that $V=W_1\oplus\dots\oplus W_k$ where
$$W_i = Im(L_i(T)) = Im\Big(\prod_{\substack{1\le j\le k\\ i\neq j}} T-\lambda_j I\Big)$$

To finish this off we just need to show $W_i = Ker(T-\lambda_i I)$. We do this by double set inclusion.
If $w=\Big(\prod_{j\neq i} T-\lambda_j I\Big)v\in W_i$ then $(T-\lambda_i I)w = p(T)v=0$. Thus $w\in Ker(T-\lambda_i I)$.
Conversely, take any $u\in Ker(T-\lambda_i I)$. We show $u\in W_i$. B/c $V=W_1\oplus\dots\oplus W_k$ we know $u=w_1+\dots+w_k$, thus
$Tu = Tw_1+\dots+Tw_k = \lambda_1 w_1+\dots+\lambda_k w_k$
But this should be
$\lambda_i u = \lambda_i w_1 + \dots + \lambda_i w_k$
Thus we get
$\lambda_1 w_1 = \lambda_i w_1,\ \lambda_2 w_2 = \lambda_i w_2,\ \dots,\ \lambda_k w_k = \lambda_i w_k$
But these should be distinct so we know $w_j=0$ $\forall j\neq i$. Thus we conclude $u\in W_i$.

We have shown double set inclusion so $W_i = Img(E_i) = Ker(T-\lambda_i I)$ as claimed. □

Remark:
For all $g(x)\in F[x]$, we have
$g(T) = g(\lambda_1)E_1 + \dots + g(\lambda_k)E_k$ where $E_j = L_j(T)$.

Note: This is just the lagrange interpolation formula again.

In particular
$T = \lambda_1 E_1 + \dots + \lambda_k E_k$ ← Spectral decomposition of T

Primary Decomposition

Example:

Consider $T = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ on $\mathbb{R}^3$ (wrt standard basis).

$Te_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 2e_1 + e_2$

$Te_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = 2e_2$

$Te_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = -e_3$

You can see e_2 & e_3 are eigenvectors. If you find the characteristic polynomial $f(x)$ you get

$$f(x) = \det(xI - T) = (x-2)^2(x+1).$$

Thus we find eigenvalues 2 & -1. We find eigenspaces

$$\mathrm{Ker}(T-2I) = \mathbb{R}e_2$$
$$\mathrm{Ker}(T+I) = \mathbb{R}e_3.$$

Since the eigenspaces don't span $\mathbb{R}^3$, T is not diagonalizable.

Consider $T-2I$. $(T-2I)e_2 = 0$ but $(T-2I)e_1 = e_2$. However notice $(T-2I)^2 e_1 = 0$. So

$$(T-2I)^2 : e_1 \mapsto 0 \ \& \ e_2 \mapsto 0.$$

Thus $\mathrm{Ker}(T-2I)^2 = \mathrm{span}\{e_1, e_2\}$. Now we can write $\mathbb{R}^3$ as

$$\begin{aligned} \mathbb{R}^3 &= \mathrm{span}\{e_1, e_2\} \oplus \mathrm{span}\{e_3\} \\ &= \mathrm{Ker}(T-2I)^2 \oplus \mathrm{Ker}(T+I) \end{aligned}$$

We call $\mathrm{Ker}(T-2I)^2$ a generalized eigenspace b/c it's kinda like an eigenspace.

Def:

The generalized λ-eigenspace of T is the set of all $v \in V$ st

$$(T-\lambda I)^m v = 0 \text{ for some } m \geq 1.$$

We'll see later you can bound m

What we're trying to do is break up the characteristic/minimal polynomials. We do this by describing how to factor them.

Lemma:

Let $p(x) \in F[x]$ st $p(x) = q_1(x) q_2(x)$ where $q_1(x)$ & $q_2(x)$ are relatively prime, i.e. $\gcd(q_1, q_2) = 1$.

Then $\exists h_1(x), h_2(x) \in F[x]$ st
$q_1 h_1 + q_2 h_2 = 1$

This is very similar to relative primality w/ the integers. It's just polynomials now.

Pf:

Consider $J = \{ q_1 a_1 + q_2 a_2 \mid a_1, a_2 \in F[x] \}$. J is an ideal of $F[x]$. That is it is closed under addition + of any two elements & multiplication by any $b \in F[x]$.

Here $J = \langle q_1, q_2 \rangle$ is the ideal generated by q_1 & q_2.

We know $J = \langle g \rangle$ can be generated by some $g \in F[x]$, b/c J is an ideal that is
$J = \{ g\, a \mid a \in F[x] \}$.

But $q_1 \in J$ & $q_2 \in J$. Therefore g divides q_1 & q_2.
But since $\gcd(q_1, q_2) = 1$, we know g is constant. Thus $1 \in J$.

Thus $\exists h_1, h_2 \in F[x]$ st $q_1 h_1 + q_2 h_2 = 1$.

B/c 1 is in the ideal we actually know J is the whole ring
$J = F[x]$.

Let V be a vector space over F. Consider $T \in End(V)$ w/ minimal polynomial $p(x) = q_1(x) q_2(x)$ where $\gcd(q_1, q_2) = 1$.

We again need another lemma.

Lemma:

Consider $T \in End(V)$ w/ minimal polynomial $p(x) = q_1(x) q_2(x)$. Then
1) $V = W_1 \oplus W_2$ where $W_i = Ker(q_i(T))$
2) the corresponding projections E_1 & E_2 are also polynomials of T.
3) the minimal polynomial of $T|_{W_i}$ is $q_i(x)$.

Pf:

We want to find projections E_1 & E_2 st
i) $E_i^2 = E_i$ $\iff$ $E_1^2 = E_1$ & $E_2^2 = E_2$
ii) $E_i E_j = 0$ $\forall i \neq j$ $\iff$ $E_1 E_2 = E_2 E_1 = 0$ ← here they're commutative
iii) $E_1 + E_2 = I$
iv) $W_i = Im(E_i)$ $\iff$ $W_1 = Im(E_1)$ & $W_2 = Im(E_2)$

Note in this case that $E_1 E_2 = E_2 E_1 = 0$ & $E_1 + E_2 = I$ then
$E_1 = E_1 I = E_1(E_1 + E_2) = E_1^2 + E_1 E_2 = E_1^2$ & $E_2 = E_2^2$ by same logic

So we just show (ii) & (iii).

We want explicitly $E_1 = g_1(T)$ & $E_2 = g_2(T)$, so
$g_1(T) g_2(T) = 0$
$g_1(T) + g_2(T) = I$

We want $W_1 = Im(E_1) = Ker(q_1(T))$, that is $E_1 v = g_1(T) v \in Ker(q_1(T))$.
Using the kernel requirement
$q_1(T) g_1(T) v = 0 \ \forall v \in V \iff q_1(T) g_1(T) = 0$.
Similarly we want $q_2(T) g_2(T) = 0$.

$g_1(T) g_2(T) = 0$ (T)+

Equivalently, we want p to divide $q_1 g_1$, $q_2 g_2$, $g_1 g_2$, & $g_1 + g_2 - 1$.

$g_1(T) + g_2(T) = I$

Recall that $p = q_1 q_2$ & $q_1 h_1 + q_2 h_2 = 1$ for some $h_1, h_2 \in F[x]$.

We let $g_1 = q_2 h_2$ & $g_2 = q_1 h_1$. Here we know
$g_1 + g_2 - 1 = 0$ b/c $q_1 h_1 + q_2 h_2 = 1$.
Likewise we can easily verify the products as $p = q_1 q_2$.

Thus we have verified that p divides all the given polynomial's b/c $q_1 g_1 = q_1 q_2 h_2$, $q_2 g_2 = q_2 q_1 h_1$, $g_1 g_2 = q_1 q_2 h_1 h_2$, $g_1 + g_2 - 1 = 0$.

Let $E_1 = g_1(T) = q_2(T) h_2(T)$ & $E_2 = g_2(T) = q_1(T) h_1(T)$.
We have already verified above that E_1 & E_2 satisfy the necessary properties to be projections. (2)

Thus $V = W_1 \oplus W_2$ where $W_i = Im(E_i)$. (1)

We now only need to show $W_i = Ker(q_i(T))$ (1) & $q_i(x)$ is the minimal polynomial of $T|_{W_i}$ (3).

We already have shown $W_i \subseteq Ker(q_i(T))$. To reiterate the some $v \in V$. Then
$E_1 v = q_2(T) h_2(T) v \in Im(E_1)$
So we know
$q_1(T) E_1 v = \underbrace{q_1(T) q_2(T)}_{p(T)=0} h_2(T) v = 0 \Rightarrow E_1 v = w_1 \in Ker(q_1(T))$

Thus $W_1 \in Ker(q_1(T))$. Likewise for W_2.

Now we show $Ker(q_i(T)) \subseteq W_i$. Consider $v \in Ker(q_1(T))$.
Then $q_1(T) v = 0$. Then
$E_2 v = q_1(T) h_1(T) v = h_1(T) q_1(T) v = 0$
Since $E_2 v = 0$, we know $E_1 v = v$, so $v \in W_1$. Likewise for W_2.

Thus $W_i = Ker(q_i(T))$. (1)

We now show (3). We show only for W_1 but W_2 is shown similarly.

Take $w \in W_1 = \ker(q_1(T))$. Then $q_1(T)w = 0$. Since $w \in W_1$, we can rewrite this as where $T_1 = T|_{W_1}$.

$$q_1(T_1)w = 0 \implies q_1(T_1) = 0.$$

Since $q_1(T_1) = 0$, we know the minimal polynomial $p_1(x)$ of $T_1 = T|_{W_1}$ divides $q_1(x)$.

Consider some $h(T_1) = 0$. Then $h(T_1)w = 0$ $\forall w \in W_1$. Since $w \in W_1$ we can write w as

$$w = E_1 v$$

Thus

$$h(T_1)w = h(T)w = h(T)E_1 v = h(T)q_2(T)h_2(T)v = 0 \quad \forall v \in V$$

Thus $h(T)q_2(T)h_2(T) = 0$. Thus p divides hq_2h_2.

B/c $q_1h_1 + q_2h_2 = 1$, it follows q_1 & q_2h_2 are coprime, that is $\gcd(q_1, q_2h_2) = 1$.

Since $p = q_1q_2$ divides hq_2h_2, we conclude q_1 divides hh_2. Since q_1 & h_2 are coprime, q_1 divides h.

Since q_1 divides all polynomials $h(x) \in F[x]$ where $h(T_1) = 0$, we conclude q_1 is the minimal polynomial of T_1. (3)

Thus our lemma is proven. □

We can recursively/inductively apply this lemma which let's use factor our polynomials in two until we reach powers of irreducible polynomials.

This gives us our theorem.

Thm: Primary Decomposition Theorem
Let $T \in End(V)$ be a linear operator w/ minimal polynomial $p(x) = q_1(x) \cdots q_k(x)$ where $\gcd(q_i, q_j) = 1$ $\forall i \neq j$. Then

1) $V = W_1 \oplus \cdots \oplus W_k$ where $W_i = \ker(q_i(T))$
2) the corresponding projections $E_1, \ldots, E_k$ are polynomials of T
3) the minimal polynomial of $T|_{W_i}$ is $q_i(x)$.

Pf:
Induction on k using the previous lemma.

We have thus shown we can write $p(x) = p_1(x)^{r_1} \cdots p_k(x)^{r_k}$ w/ irreducible $p_i(x)$. Then $q_i(x) = p_i(x)^{r_i}$ where $\gcd(q_i, q_j) = 1$ $\forall i \neq j$.

Example:
1) In $\mathbb{C}[x]$ the irreducible monic polynomials are $(x - \lambda)$ where $\lambda \in \mathbb{C}$.

2) In $\mathbb{R}[x]$ the irreducible monic polynomials are $(x - \lambda)$ where $\lambda \in \mathbb{R}$ or $(x - \lambda)(x - \bar{\lambda})$ where $\lambda \in \mathbb{C}$ & $\lambda \notin \mathbb{R}$.

Split Linear Operators (aka Triangulable)

For a split linear operator T the minimal polynomial $p(x)$ is a product of linear operators
$p(x) = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k}$.
B/C of this split we know
$V = W_1 \oplus \cdots \oplus W_k$ where $W_i = \ker(T-\lambda_i I)^{r_i}$ = generalized λ_i-eigenspace

Example:
Here we'll be solving ODEs w/ constant coefficients.

Consider U = space of smooth functions of t. Let $T = \frac{d}{dt} \in \mathrm{End}(U)$.
Take some $p(x) = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k} \in \mathbb{C}[x]$.

By our earlier theorem,
$V = \ker(p(T))$ = space of solutions to ODE $p(\frac{d}{dt})f(t) = 0$.
Then $p(T) = 0$ on V, by definition so $V = W_1 \oplus \cdots \oplus W_k$ where $W_i = \ker(T-\lambda_i I)^{r_i}$
by our primary decomposition theorem.

Thus we've split up solving our potentially very complex polynomial to solving linear factors which is very easy.

Solve the ODE
$$\left(\frac{d}{dt} - \lambda\right)^r f(t) = 0$$
Write $f(t) = e^{\lambda t} g(t)$. Then $\left(\frac{d}{dt}\right)^r g(t) = 0$ for some $g(t) \in \mathbb{C}[t]_{\deg \le r-1}$.

Therefore, V (our space of solutions) has a basis of
$$\{ t^j e^{\lambda_i t} \mid 1 \le i \le k \ \& \ 0 \le j \le r_i - 1 \}$$
As expected $\dim(V) = r_1 + \cdots + r_k = \deg(p)$.

Now consider a split operator T w/ minimal polynomial
$p(x) = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k}$
So $V = W_1 \oplus \cdots \oplus W_k$ where $W_i = \ker(T-\lambda_i I)^{r_i}$ = generalized λ_i-eigenspace.

Def:
The <u>semi-simple part</u> (or diagonalized part of T) of T is the linear operator D defined by
$D|_{W_i} = \lambda_i I_{W_i} \ \forall i = 1 \ldots k$. Thus
$D = \lambda_1 E_1 + \cdots + \lambda_k E_k$

Since $V = W_1 \oplus \cdots \oplus W_k$ (recall $W_i = \ker(T-\lambda_i I)^{r_i}$), for all $v \in V$ $\exists!$ $v = w_1 + \cdots + w_k$

Thus $Dv = Dw_1 + \dots + Dw_k = \lambda_1 w_1 + \dots + \lambda_k w_k$.

- Essentially we devise a linear operator w/ eigenspaces of T's generalized eigenspaces.

By construction D is diagonalizable w/ $W_i = \ker(D - \lambda_i I)$.

Consider $\underline{N = T - D}$. (or $T = D + N$) Then on W_i we have $N^{r_i} = (T - \lambda_i I)^{r_i} = 0$. Thus $N^{r_i} w_i = 0 \ \forall w_i \in W_i$

If we take the largest r_i's $r = \max(r_1, \dots, r_k)$, then $N^r w_i = 0 \ \forall w_i \in W_i$.

Then $N^r = 0$ is <u>nilpotent</u>

<u>Def:</u>
The N defined above is the <u>nilpotent part</u> of T. It is unique.

Jordan-Chevalley Decomposition

The Jordan-Chevalley is our decomposition of linear operator T into
$T = D + N$
where D is the semi-simple part of T & N the nilpotent one.

<u>Thm:</u>
Let T be a linear operator on a finite dimensional vector space V.
T can be written <u>uniquely</u> as
$T = D + N$
where
D is diagonalizable/ semi-simple,
N is nilpotent, &
D & N commute $DN = ND$.

Furthermore, there exist polynomials $g(x), h(x) \in F[x]$ st
$D = g(T)$, $N = h(T)$, & $g(0) = h(0) = 0$.

<u>Pf:</u>
First we show such a D & N <u>exist</u>.

Assume $V = W_1 \oplus \dots \oplus W_k$ where $W_i = \ker(T - \lambda_i I)^{r_i}$ is generalized eigenspace for T.
By definition of D from earlier we know
$D|_{W_i} = \lambda_i I_{W_i}$ b/c $W_i = \ker(D - \lambda_i I)$ is an eigenspace for D.
So D is diagonalizable

Let $N = T - D$, then $N^{r_i}|_{W_i} = 0$, so $N^r = 0$ where $r = \max(r_1, \dots, r_k)$. Thus N is nilpotent.

By the definition of D & N we know $T = D + N$.

We know $\exists g(x) \in F[x]$ st $g(T) = D$ b/c
$D = \lambda_1 E_1 + \dots + \lambda_k E_k$
& E_i is a polynomial of T, Thus $\exists g(x) \in F[x]$ st $g(T) = D$.

Consider $h(x) = x - g(x)$. Then $h(T) = T - g(T) = T - D = N$. Thus $\exists h(x) \in F[x]$ st $N = h(T)$.

B/C D & N are polynomials of T, we know they commute $DN = ND$.

Now we just show that we can choose $g(x), h(x) \in F[x]$ st $g(0) = h(0) = 0$.
It is sufficient to show $g(0)=0$ b/c $h(0) = 0 - g(0) = 0$ if & only if $g(0)=0$.

Note that $\lambda_1, \dots, \lambda_k$ are eigenvalues of T. Take some eigenvector $v \in V$ where $v \neq 0$ w/ $Tv = \lambda_i v$. Thus $v \in W_i$.
Since $v \in W_i$ we know

$$Dv = \lambda_i v = g(T)v = g(\lambda_i)v$$

We have thus show $g(\lambda_i) = \lambda_i$.

If some $\lambda_i = 0$, we can easily choose $g(0)=0$.

Suppose however all eigenvalues are non-zero $\forall \lambda_i \neq 0$. Then the minimal polynomial of T is

$$p(x) = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k}.$$

Since all $\lambda_i \neq 0$, we know $p(0) \neq 0$.
Then we can replace $g(x)$ w/

$$\tilde{g}(x) = g(x) - \frac{g(0)}{p(0)} p(x).$$

Here we know

$$\tilde{g}(T) = g(T) - \cancel{\frac{g(0)}{p(0)} p(T)} = g(T) = D$$

$$\tilde{g}(0) = g(0) - \frac{g(0)}{\cancel{p(0)}} \cancel{p(0)} = 0$$

Thus such a D & N w/ polynomials $g(x)$ & $h(x)$ exist.
<u>(existence)</u>

Now we show such a D & N are <u>unique</u>.

Suppose $T = D + N = D' + N'$ w/
D' diagonalizable,
N' nilpotent, &
$D'N' = N'D'$.

We show $D = D'$ & $N = N'$.

To do this we show $D'T = TD'$ & $N'T = TN'$.

$$D'T = D'(D'+N') = D'D' + D'N'$$
$$TD' = (D'+N')D' = D'D' + N'D'$$

Here we can see $D'T - TD' = 0$, so $D'T = TD'$. $N'T = TN'$ follows likewise.

By similar logic, we know D' & N' commute w/ any polynomial of T. Since $D = g(T)$ & $N = h(T)$, we know
$D'D = DD'$ & $N'N = NN'$.

Since both N & N' are nilpotent, there exists $m \in \mathbb{N}$ st
$N^m = (N')^m = 0$

B/c N & N' commute, we know
$(N' - N)^{2m} = 0$
as you can expand it w/ the usual binomial theorem & you'll always have at least one of N or N' have a power of at least m, so they are all 0.

Consider $D - D' = N' - N$. Since D & D' commute, they can be diagonalized simultaneously.

Recall: If you have some diagonalizable operators $T_1, \ldots, T_k$ that all commute. Then $T_1, \ldots, T_k$ can be diagonalized simultaneously.

Since D & D' can be diagonalized simultaneously, $\exists$ basis of eigenvectors for D & D'. This same basis is a basis of eigenvectors for $D - D'$. Thus $D - D'$ is diagonalizable.

If for $v \neq 0$, $(D - D')v = \mu v = (N - N')v$, we know
$(D - D')^{2m} v = \mu^{2m} v = (N - N')^{2m} v = 0$
$\Rightarrow \mu^{2m} = 0$
$\Rightarrow \mu = 0$,

Thus $D - D' = N - N' = 0$ & $D = D'$ & $N = N'$. Thus the D & N are unique (unique)

Cyclic Subspaces

Def:

Let V be a finite dimensional vector space w/ $T \in End(V)$. Pick some $v \in V$ & consider $v, Tv, T^2 v, T^3 v, \ldots$

The T-cyclic subspace generated by v is $Z(v; T) =$

$$Z(v; T) = span\{v, Tv, T^2 v, \ldots\} = \{g(T)v \mid g(x) \in F[x]\}$$

This is the minimum subspace of V that contains v & is T-invariant.

Sum of powers of T is polynomial

It is clearly a subspace b/c it's a span & it is clearly T-invariant. It is intuitive that it is minimum but not clear.

By the pigeonhole principle (valid b/c V is finite dimensional), $\exists$ minimum $k \in \mathbb{N}$ st $T^k v \in span\{v, Tv, \ldots, T^{k-1} v\}$.

We say k is the max number such $\{v, Tv, \ldots, T^{k-1} v\}$ is linearly independent.

Thus we know $\{v, Tv, \ldots, T^{k-1} v\}$ is a basis for $Z(v; T)$.

B/c $\{v, Tv, \ldots, T^{k-1}v\}$ is a basis for $Z(v; T)$, we know $\forall v \in Z(v; T)$

$$v = c_0 v + c_1 Tv + \ldots + c_{k-1} T^{k-1} v = g(T)v \quad \text{where } g(T) \in F[x]_{\deg \leq k-1}$$

In particular

$$T^m v \in \operatorname{span}\{v, Tv, \ldots, T^{k-1}v\} = Z(v; T) \text{ for all } m \geq k.$$

We thus know $T^k v = g(T)v$. This is equivalent to $h(T)v = 0$ where

$$h(x) = x^k - g(x) = x^k - c_{k-1}x^{k-1} - \ldots - c_1 x - c_0.$$

We can see $h(x)$ is a monic polynomial of degree k. In particular, $h(x)$ is the minimal polynomial st $h(T)v = 0$.

<u>Recall:</u>

The T-annihilator of v is

$$S(v; \{0\}) = \{ g(x) \in F[x] \mid g(T)v = 0 \}$$

This is an ideal in $F[x]$ generated by the monic polynomial $p_v(x)$ of minimum degree called the T-annihilator (polynomial) of v.

Since $p_v(x)$ is the generator of $S(v; \{0\})$ we know that $g(T)v = 0$ iff $p_v(x)$ divides $g(x) \in F[x]$

Since the minimum polynomial $p(x)$ of T has $p(T) = 0$ by definition, we know $p_v(x)$ divides $p(x)$.

<u>Thm:</u>

The cyclic subspace $Z(v; T)$ generated by $v \in V$ where $v \neq 0$ has a basis $\{v, Tv, \ldots, T^{k-1}v\}$ where $k = \deg(p_v(x))$.

Hence $\dim(Z(v; T)) = \deg(p_v(x))$.

Moreover $T|_{Z(v;T)}$ has minimal polynomial $p_v(x)$.

Let's try to find the matrix A of $T|_{Z(v;T)}$ relative to the ordered basis $\{v, Tv, \ldots, T^{k-1}v\}$.

$$A = \underbrace{\begin{bmatrix} 0 & 0 & & 0 & c_0 \\ 1 & 0 & \ddots & 0 & c_1 \\ 0 & 1 & \ddots & 0 & c_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 1 & c_{k-1} \end{bmatrix}}_{\text{Companion matrix of } p_v(x)}$$

This makes sense as we are finding $T(v + Tv + \ldots + T^{k-1}v)$. So

$$T(v + Tv + \ldots + T^{k-1}v) = Tv + \ldots + T^{k-1}v + T^k v = Tv + \ldots + T^{k-1}v + (c_0 v + c_1 Tv + \ldots + c_{k-1}T^{k-1}$$

Where do we get these cs? from the minimum polynomial / annihilator
$p_v(x) = x^k - c_{k-1}x^{k-1} - \dots - c_1 x - c_0$.

We can construct such a matrix A from any monic polynomial $h(x)$ of deg $= k$.
Here A is called the <u>companion matrix</u> of A.

<u>Def:</u>
If the cyclic subspace $Z(v;T)$ is the whole space V, that is
$Z(v;T) = V$
Then we say v is a <u>cyclic vector</u>.

<u>Coro:</u>
If v is a cyclic vector then

1) $p_v(x) = p(x) = f(x)$ where $p_v(x)$ is the T-annihilator of v,
$p(x)$ is the minimum polynomial, &
$f(x)$ is the characteristic polynomial.

2) If A is the companion matrix of a monic polynomial $h(x)$ then
$h(x) =$ minimal polynomial $=$ characteristic polynomial.

<u>Thm:</u> Cyclic Decomposition Theorem
For all finite dimensional subspaces & $T \in End(V)$ $\exists v_1, \dots, v_r \in V^V$ st $p_i(x) = p_{v_i}(x)$
satisfy
& $p_1 | p_2, p_2 | p_3, \dots, p_{r-1} | p_r$ (p_1 divides p_2, ...)
$V = Z(v_1;T) \oplus \dots \oplus Z(v_r;T)$.

The number r & polynomials $p_1, \dots, p_r$ are unique.

<u>Def:</u>
$p_1(x), \dots, p_r(x)$ are called the <u>invariant factors</u> of T.

<u>Properties:</u>
$p_1, p_2, \dots, p_{r-1}$ all divide p_r, so p_r is the "largest." We know

1) $p_r(x) = p(x) =$ min. polynomial of T

2) $p_1(x) \cdots p_r(x) = f(x) =$ char. polynomial of T

<u>Coro:</u>
There is a cyclic vector of T if & only if $p(x) = f(x)$.

<u>Background:</u> Rational Form
Let $T \in End(V)$ be a T-invariant direct sum $V = W_1 \oplus \dots \oplus W_r$ if
$Tw \in W_i$ $\forall w \in W_i$. Pick basis B_i for W_i.

Then $B = \bigcup_{i=1}^{r} B_i$ is a basis for V.

Then $[T]_B$ is a block diagonal matrix where
$[T]_B = \begin{bmatrix} A_1 & \dots & 0 \\ 0 & & A_r \end{bmatrix}$ where $A_i = [T|_{W_i}]_{B_i}$.

Then from the cyclic decomposition theorem
$$V = Z(v_1; T) \oplus \cdots \oplus Z(v_r; T).$$

__Thm:__ Rational form
Let V be a vector space. Then $\exists$ a basis B for V st
$$[T]_B = \begin{bmatrix} A_1 & \cdots & 0 \\ 0 & \cdots & A_r \end{bmatrix}$$
where A_i is the companion matrix of $p_i(x)$ & $p_1(x), \ldots, p_r(x)$ are the invariant factors of T.

This matrix is a unique canonical form called the __rational form__

__Def:__
If $V = W \oplus W'$ we say subspace W' is __complementary__ to W.

Let's prove the cyclic decomposition for a special case, nilpotent operators.

__Pf:__ Cyclic Decomposition for Nilpotent Operators
Let V be a vector space over field F w/ $\dim(V) = n$.

Let $N \in End(V)$ be a nilpotent operator, that is $N^k = 0$ for some $k \geq 1$.
Then the minimal polynomial $p(x)$ divides x^k (annihilating polynomial of N). If you take the minimum k then $p(x) = x^k$ b/c then $p(T) = 0$ & $p(x) \in F[x]$ is the minimum degree polynomial to do so.

Suppose that the invariant factors $p_1(x), \ldots, p_r(x)$ such that $p_i(x) = x^{k_i}$ where $k_1 \leq k_2 \leq \cdots \leq k_r = k$.

Since $\dim(V) = n$ & N is nilpotent, the char. polynomial is $f(x) = x^n$.

The product of $p_1(x), \ldots, p_r(x)$ is $x^{k_1 + k_2 + \cdots + k_r}$. Since $p_1(x) \cdot \ldots \cdot p_r(x) = f(x)$ we know $k_1 + \cdots + k_r = n$.

The companion matrix A_i of $p_i(x) = x^{k_i}$ is
$$A_i = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ 1 & \ddots & \vdots & 0 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} = [N|_{Z(v_i; N)}]_{B_i}$$
where $B_i = \{v_i, Nv_i, \ldots, N^{k_i - 1} v_i\}$ & $N^{k_i} = 0$ b/c $(p_{v_i}(x) = x^{k_i})$ (annihilator of v_i).

We expect $B = \bigsqcup_{i=1} B_i$ to be a basis for V.

We visualize the action of N on $v_1, \ldots, v_r$ as

$$v_1 \xrightarrow{N} Nv_1 \longrightarrow \cdots \longrightarrow N^{k_1} v = 0 \quad k_1 \text{ vectors} \neq 0$$

$$v_2 \xrightarrow{N} Nv_2 \longrightarrow \cdots \longrightarrow N^{k_2} v = 0 \quad k_2 \text{ vectors} \neq 0. \text{ Note } k_2 > k_1$$

$$\vdots$$

$$v_r \xrightarrow{N} Nv_r \longrightarrow \cdots \longrightarrow N^{k_r} v = 0 \quad k_r = k \text{ vectors} \neq 0$$

Here it should be obvious if the chains are a basis then for all $v \in V$
$N^k v = 0$
This is b/c in the largest case k_r, $k \geq k_r$. Thus the vector has "progressed along the chain" far enough for it to become 0. This is the

To prove the cyclic decomposition theorem for N. We have to show V has a basis B of such chains
$B = \{v_1, Nv_1, \ldots, N^{k_1 - 1} v_1, \ldots, v_r, Nv_r, \ldots, N^{k_r - 1} v_r\}$.

We prove this by induction on $\dim(V)$.

If $\dim(V)$, then $V = Fv_1$ for any $v_1 \neq 0$. By the definition of N, we know $Nv_1 = 0$
Thus B does have a basis $V = \{v_1\}$ b/c any non-zero vector forms a basis for a 1-dim space.

Now assume the statement is true for all spaces w/ $\dim < \dim(V)$.
We want to find an N-invariant proper subspace of N.

Consider $Im(N)$ as our proper subspace. It is indeed proper by the rank-nullity theorem
$\dim(Im(N)) + \dim(Ker(N)) = \dim(V)$
Thus to show $\dim(Im(N)) < \dim(V)$, we show $\dim(Ker(N)) = 0$.

$\dim(Ker(N)) \neq 0$ b/c $N^k v = N(N^{k-1} v) = 0$ where $N^{k-1} v \neq 0$. Since there is a non-zero vector N^{k-1} in the kernel, the kernel is non-trivial & $\dim(Ker(N)) > 0$.

Thus $Im(N) \subset V$ & $\dim(Im(N)) < \dim(V)$. Since $Im(N)$ is an N-invariant subspace (b/c clearly $Nv \in Im(N)$ $\forall v \in Im(N) \subset V$) & has smaller dimension than V, we use the inductive assumption.
That is $W = Im(N)$ has a basis of non-zero vectors from the chains

$$w_1 \xrightarrow{N} \cdot \longrightarrow \cdots \longrightarrow 0$$

$$\vdots$$

$$w_s \xrightarrow{N} \cdot \longrightarrow \cdots \longrightarrow 0 \quad \text{where } s < r$$

Every such vector $w_i \in W = Im(N)$. Thus $w_i = Nv_i$ for some $v_i \in V$.
Thus we can extend the chains (next page).

$v_1 \xrightarrow{N} w_1 \longrightarrow \cdots \longrightarrow 0$

$\cdots$

$v_s \xrightarrow{N} w_s \longrightarrow \longrightarrow \cdots \longrightarrow 0$

Now we need to show the vectors we added don't mess up the basis properties.

<u>Claim</u>: These vectors are (still) linearly independent

<u>Proof</u>:

Suppose

$$c_{10} v_1 + c_{11} N v_1 + \cdots + c_{1\ell_1} N^{\ell_1} v_1 + \cdots + c_{s0} v_s + c_{s1} N v_s + \cdots + c_{s\ell_s} N^{\ell_s} v_s = 0$$

(Probably should be using Σ notation)

Apply N to both sides. We're essentially moving everything along, giving us (use $w_i = N v_i$)

$$c_{10} w_1 + c_{11} N w_1 + \cdots + c_{1\ell_1 - 1} N^{\ell_1 - 1} w_1 + \cdots + c_{s0} v_s + c_{s1} N w_1 + \cdots + c_{s\ell_s - 1} N^{\ell_s - 1} w_s = 0$$

Since $\{w_1, Nw_1, \ldots, w_s, Nw_s, \ldots\}$ forms a basis, we know

$$c_{10} = \cdots = c_{1\ell_1 - 1} = \cdots = c_{s0} = \cdots = c_{s\ell_1 - 1} = 0$$

Thus all coefficients but the last are zero. That is we now must show

$$c_{1\ell_1} N^{\ell_1} v_1 + \cdots + c_{s\ell_s} N^{\ell_s} v_s = c_{1\ell_1} N^{\ell_1 - 1} w_1 + \cdots + c_{s\ell_s} N^{\ell_s - 1} w_s = 0$$

Since $N^{\ell_1 - 1} w_1, \ldots, N^{\ell_s - 1} w_s$ is part of a basis for W, we know

$$c_{1\ell_1} = \cdots = c_{s\ell_s} = 0$$

Since all c's are zero, the vectors are linearly independent

Since the vectors $B' = \{v_1, Nv_1, \ldots, v_s, Nv_s, \ldots\}$ are linearly independent, B' is a candidate basis. (Note: It might not span V yet.)

We know we can extent B' to a basis for V. That is $\exists B'' \subset V$ st $B' \sqcup B''$ is a basis for V, Since $B' \sqcup B''$ is a disjoint union, we know $V = W' \oplus W''$ where $W' = \text{span}(B')$ & $W'' = \text{span}(B'')$.

Note that W' is N-invariate. However, the complementary subspace might not be N-invariant. We need to show it is.

Consider some $v \in B''$. Then $Nv \in Im(N) = W$ obviously. However, we have just shown that $\forall v \in W$ $\exists v' \in V'$ st $Nv = Nv'$.

Since $Nv = Nv'$, we know $Nv - Nv' = N(v - v') = 0$, so $v - v' \in Ker(N)$. If we replace v w/ $v - v'$ then $v - v'$ is still part of linearly independent set B'' st $B' \sqcup B''$ is a basis for V.

If we do this for all $v \in B''$, we get $B'' \subset Ker(N)$. Trivially then, W'' is N-invariate b/c all vectors go to 0.

Thus the basis $B = B' \sqcup B''$ is a union of chains from B' & B'' which consists of "single link chains" like

$$\left.\begin{array}{l} v'' \xrightarrow{N} 0 \\ \bullet \longrightarrow 0 \\ \circ\circ\circ \\ \bullet \longrightarrow 0 \end{array}\right\} v'' \in B''$$

<u>Remark:</u>
$r = \dim(Ker(N)) = Nullity(N)$ & $k_1, \ldots, k_r$ are uniquely determined from N.
In the above proof r is the number of chains.

Jordan (Canonical) Form

Consider some $T \in End(V)$ where $\dim(V) = n < \infty$ & T is a split linear operator.

This always happens over the field $F = \mathbb{C}$

B/C T is split we know that the min. polynomial $p(x)$ & char. polynomial $f(x)$ can be written

$$p(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_\ell)^{k_\ell}$$
$$f(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_\ell)^{d_\ell}$$

where $\lambda_i \neq \lambda_j$ $\forall i \neq j$ & $1 \leq k_i \leq d_i$ $\forall i$.

By the primary decomposition theorem, we have
$V = W_1 \oplus \cdots \oplus W_\ell$ where $W_i = Ker(T - \lambda_i I)^{k_i}$ are the generalized eigenspaces.
Note that $d_i = \dim(W_i)$. Clearly W_i is T-invariant.

Define N_i as $N_i = (T - \lambda_i I)|_{W_i} \in End(W_i)$. By construction N_i is nilpotent b/c $N_i^{k_i} = 0$ on W_i.
Thus, its min. polynomial is x^{k_i}.

We then use the rational form of N_i to get the Jordan form of the matrix.

Let B_i be a basis for W_i st $[N_i]_{B_i}$ is in rational form. Let $R_i = [N_i]_{B_i}$.

Notice $T|_{W_i} = N_i + \lambda_i I$ by the definition of N_i. Let $A_i = [T|_{W_i}]_{B_i} = \lambda_i I + R_i$.

The disjoint union $B = \bigsqcup_{i=1}^{\ell} B_i$ is a basis for V &
$$A = [T]_B = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_\ell \end{bmatrix}$$
is block diagonal. This is the <u>Jordan form</u> of T.

The rational form R_i of the nilpotent operator N_i is itself block diagonal.
$$R_i = \begin{bmatrix} R_{i,1} & & 0 \\ & \ddots & \\ 0 & & R_{i,r_i} \end{bmatrix} \text{ where each } R_{i,j} \text{ is } R_{i,j} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}.$$

The min. polynomial of N_i is x^{k_i} where k_i is the multiplicity of the eigenvalue λ_i as a root of the min. polynomial $p(x)$ of T. Thus the invariant factors of N_i are powers of x: $x^{k_{i,1}}, x^{k_{i,2}}, \dots, x^{k_{i,r_i}}$ where $1 \le k_{i,1} \le \dots \le k_{i,r_i} = k_i$ & $k_{i,1} + \dots + k_{i,r_i} = d_i = \dim(W_i)$.

The number r_i is the nullity of N_i, that is $r_i = \dim(\mathrm{Ker}(N_i))$, which is $\dim(\mathrm{Ker}(T-\lambda_i I))$ or the dimension of the λ_i eigenspace. So
$r_i = \dim(\mathrm{Ker}(N_i)) = \dim(\mathrm{Ker}(T-\lambda_i I))$.
The size of R_{ij} is $k_{ij} \times k_{ij}$.

By our earlier definition of A_i we get
$A_i = \lambda_i I + R_i = \begin{bmatrix} J_{i,1} \dots & 0 \\ 0 & J_{i,r_i} \end{bmatrix}$ where $J_{i,j} = \begin{bmatrix} \lambda_i & \cdots & \cdots & 0 \\ 1 & \lambda_i & & \vdots \\ \vdots & 1 & \ddots & \vdots \\ 0 & \cdots & 1 & \lambda_i \end{bmatrix}$
We call $J_{i,j}$ a Jordan block & it is $k_{ij} \times k_{ij}$.

Since $A = \begin{bmatrix} A_1 \dots & 0 \\ 0 & A_l \end{bmatrix}$ & $A_i = \begin{bmatrix} J_{i,1} & 0 \\ 0 & J_{i,r_i} \end{bmatrix}$, we have the diagonal of $A = [T]_B$ is made up of Jordan blocks.

<u>Thm:</u> Jordan Form
For every split linear operator $T \in \mathrm{End}(V)$, there is a basis B for V st the matrix $[T]_B$ is in Jordan form, that it is block diagonal w/ Jordan blocks on the diagonal. That is

$$A = [T]_B = \begin{bmatrix} J_{1,1} \dots & & & 0 \\ & J_{1,r_1} \dots & & \\ & & J_{l,r_1} \dots & \\ 0 & & & J_{l,r_l} \end{bmatrix}$$

We know that this matrix $A = [T]_B$ is unique if we order the blocks.

<u>Note:</u> You may ones above or below the diagonal in Jordan Form. It is equivalent. (Reversing the basis switches b/w the two.)

<u>Properties:</u> Jordan Form
1) A linear transformation has only its eigenvalues $\lambda_1, \dots, \lambda_l$ along the main diagonal. Each λ_i appears d_i times which is the multiplicity of λ_i as a root of the char. polynomial $f(x)$.

See next page

2) The number r_i of Jordan blocks w/ eigenvalue λ_i is $r_i = \dim(W_i)$.

3) The largest Jordan block w/ eigenvalue λ_i has size k_i where k_i is the multiplicity of λ_i as root of min. polynomial $p(x)$.

Example:

Let $T = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{End}(\mathbb{R}^3) = \mathbb{R}^{3\times 3}$. Normally we would do $\mathbb{C}^3$, but we do $\mathbb{R}^3$ b/c the char. polynomial factors over $\mathbb{R}$

We find the char. polynomial $f(x)$ of T as $f(x) = (x-1)^3$.

We thus find k st $N = T - I$ is nilpotent.

$N = T - I = \begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, likewise $N^2 \neq 0$ but $N^3 = 0$.

Thus the min polynomial $p(x)$ of T is $p(x) = (x-1)^3$.

B/c this has one eigenvalue we know V is a cyclic space

Since $\ker(N) = \ker(T - I)$ has dim 1 (i.e. the 1-eigenspace has dim 1), we have 1 Jordan block $[T]_B$

$[N]_B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow [T]_B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

Notice $[N]_B$ is also in rational form.

How do we find the rational form of N? You need to pick some $v \in V$ st $N^2 v \neq 0$ b/c then we have chain

$$v \to Nv \to N^2 v \to 0$$

Then $B = \{v, Nv, N^2v\}$ is a basis of $V = \mathbb{R}^3$ st N is in rational form & thus T is in Jordan form.

Let's find the rational form of T. It is the companion matrix of the min. polynomial $p(x)$

$p(x) = (x-1)^3 = x^3 - 3x^2 + 3x - 1 = x^3 - (3x^2 - 3x + 1)$

so the rational form of T is

$[T]_{B'} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{bmatrix}$

The basis B' that gives this rational form is $B' = \{v, Tv, T^2v\}$.

We can use the same vector $v \in V = \mathbb{R}^3$ as earlier, since $p_v(x) = (x-1)^3$. So $(T-I)^2 v = N^2 v \neq 0$. We do just need any $v \in V = \mathbb{R}^3$ st $T^2 v \neq 0$ tho.

To find $v \in V = \mathbb{R}^3$ just test it out when you find N^2.

Example:

Let $T = \begin{bmatrix} -1 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \in \mathrm{End}(\mathbb{R}^3) = \mathbb{R}^{3\times 3}$. Again normally we'd take $\mathbb{C}^3$, but $\mathbb{R}^3$ is enough to split T.

The char. polynomial $f(x)$ of T is thus $f(x) = (x+1)^3$.

We now find the mini. polynomial $p(x)$ by finding a k st $N^k = 0$ where $N = T+I$ is the nilpotent operator.

$$N = T+I = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow N^2 = 0 \Rightarrow p(x) = (x+1)^2$$

Since $k = 2 < 3$, we know $V = \mathbb{R}^3$ isn't a cyclic space under T & instead there will be 2 subspaces.

Notice by inspection of N that

$Ker(N) = Ker(T+I) = span\{e_1, e_3\}$ ← Basis

so the nullity of N is 2 & thus the dimension of the -1 eigenspace is 2. Thus again we have 2 Jordan blocks.

Since there are only 2 Jordan blocks, the Jordan form of T must be

$$[T]_B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad \text{or} \quad [T]_B = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$ ← depending on order of Jordan blocks.

We could if we allowed it, have the ones above the diagonal.

By this same B we get the rational form of N

$$[N]_B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

How do we find this basis $B = \{v_1, v_2, v_3\}$? By inspection of B, we need

$$\begin{matrix} Nv_1 = 0 \\ Nv_2 = v_3 \\ Nv_3 = 0 \end{matrix} \iff \begin{matrix} v_1 \xrightarrow{N} 0 \\ v_2 \xrightarrow{N} Nv_2 = v_3 \xrightarrow{N} 0 \end{matrix}$$

Note that the kernel of N has a basis $\{v_1, v_3\}$.

We have 2 ways to find $B = \{v_1, v_2, v_3\}$.

1) Start from our earlier basis for $Ker(N)$, namely $\{e_1, e_3\}$. Then we search for $v_3 \in Ker(N)$ & find v_2 st $Nv_2 = v_3$.

Note that $Im(N) = span\{v_3\}$. Since $Im(N)$ is its column space & $N = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, we know $v_3 = 2e_1$ will work.

If you look at the action of N on the basis you see
$Ne_1 = 0, \quad Ne_2 = 2e_1, \quad Ne_3 = 0$

Thus we can see $v_1 = e_3$, $v_2 = e_2$, & $v_3 = 2e_1$ will work for B.

2) Find v_2 st $N v_2 \neq 0$. Then let $v_3 = N v_2$.

Now let's find the rational form of T. We need to find the invariant factors. We know the number of invariant factors of T is 2 b/c that is the dimension of the eigenspace.

The invariant factors for T are polynomials $p_1(x)$, $p_2(x)$ st p_1 divides p_2, $p_2 = p$ & $p_1 p_2 = f$.

B/c $p_2 = p$, we know $p_2 = (x+1)^2$. This leaves $p_1(x) = x+1$ as the only option. By rewriting $p_1(x)$ & $p_2(x)$ we get the rational form, as

$p_1(x) = x - (-1)$
$p_2(x) = x^2 - (-2x - 1)$

$$\Rightarrow [T]_{B'} = \left[\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & 0 & -1 \\ 0 & 1 & -2 \end{array}\right]$$

To find $B' = \{u_1, u_2, u_3\}$ we need $Tu_1 = -u_1$, $(T+I)u_1 = 0$,
$Tu_1 = -u_1$, $(T+I) = u_1$
$Tu_2 = u_3$, $Tu_3 = T^2 u_2 = -u_2 - 2Tu_2$ ← Come from $(T+I)^2 u_2 = 0$.

We want u_2 such that it's annihilator is $p_{u_2}(x) = (x+1)^2$, but $(T+I)u_2 \neq 0$.

We can take $u_2 = e_2$ b/c $Ne_2 = (T+I)u_2 = 2e_1 \neq 0$. Then $Tu_2 = Te_2 = 2e_1 - e_2$
Then take $u_3 = e_3$.

The cyclic decomposition of T is
$V = Z(u_1; T) \oplus Z(u_2; T)$

$Z(v_1; N) = \text{span}\{e_3\}$, $v_1 = e_3$ $\qquad$ $Z(v_2; N) = \text{span}\{e_1, e_2\}$, $v_2 = e_2$

Bilinear Forms

<u>Def:</u>
Consider a vector space V. A <u>bilinear form</u> on V is a bilinear map $f: V \times V \to F$ which is bilinear (i.e. linear in each argument when the other is fixed). More explicity for all $v, v', w, w' \in V$ & $c \in F$

$f(v+v', w) = f(v, w) + f(v', w)$
$f(cv, w) = cf(v, w) = f(v, cw)$
$f(v, w+w') = f(v, w) + f(v, w')$

Recall our work on tensor products. Every bilinear map $V \times V \to F$ has a corresponding linear map on the tensor product $V \otimes V \to F$.

We can view this corresponding tensor form as a linear functional on $V \otimes V$.

The set of all bilinear forms is a vector space, as are all linear transformations.

$$(f+g)(v, w) = f(v, w) + g(v, w)$$
$$(cf)(v, w) = c f(v, w)$$

This vector space is the dual of $V \otimes V$, that is $(V \otimes V)^*$.

Example:

Consider $V = \mathbb{R}^n$ w/ the dot product, which is an inner product

$(X, Y) = X^t Y =$ dot product ← dot product is symmetric

By definition all inner products are bilinear forms. That is expected b/c bilinear forms are a generalization of inner products.

Example:

Consider $V = \mathbb{R}^2$. Define bilinear form $d(X, Y)$ on $X = \begin{bmatrix} a \\ c \end{bmatrix}$ & $Y = \begin{bmatrix} b \\ d \end{bmatrix}$ as

$d(X, Y) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ ← d is skew-symmetric

Def:

We call a bilinear form f <u>symmetric</u> iff

$f(v, w) = f(w, v) \quad \forall v, w \in W$

We call it <u>skew-symmetric</u> iff

$f(v, w) = -f(w, v)$

Example:

Consider some arbitrary vector space V w/ $g, h \in V^*$.

we define a bilinear form

$f(v, w) = g(v) h(w)$

Essentially $f = g \otimes h \in V^* \otimes V^* \subseteq (V \otimes V)^*$.

Note that $V^* \otimes V^* = (V \otimes V)^*$ if V is finite dimensional.

Example:

Fix some matrix $A \in F^{n \times n}$ & let bilinear form $f(X, Y)$ be defined

$f(X, Y) = X^t A Y = (X, AY)$ ← dot product $(X, Y) = X^t Y$

where X & Y are column matrices $X, Y \in F^n$

We can actually use the above example to get a canonical form of every bilinear form.

Example: cont.

Continuing the previous example f is symmetric/skew-symmetric iff A is likewise symmetric/skew-symmetric.

Recall A is symmetric if $A = A^t$ & skew symmetric iff $A = -A^t$

This likewise holds for degeneracy

Def:

The matrix in the above example is called the Gram matrix. In particular where you identify your bilinear form w/ the matrix A.

We can construct A by
$A = (a_{ij})$ where $a_{ij} = f(e_i, e_j)$ where $\{e_1, \dots, e_n\}$ standard basis for F^n.

We can generalize/apply this Gram matrix to any vector space V.

Let $B = \{v_1, \dots, v_n\}$ be a basis for vector space V. For any $v, w \in V$ we know
$v = \sum_{i=1} x_i v_i \iff [v]_B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = X$ & $w = \sum_{j=1} y_j v_j \iff [w]_B = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = Y$

Using this we can thus define $f(v, w)$ as
$f(v, w) = f(\sum_i x_i v_i, \sum_j y_j v_j) = \sum_i \sum_j x_i f(v_i, v_j) y_j = X^t A Y$ where $A = (a_{ij})$.

Def:

Using the above work for some bilinear form f on vector space V w/ basis B, we can write the matrix of f wrt B as
$[f]_B = A = (a_{ij})$ where $a_{ij} = f(v_i, v_j)$.

Thm:

For any bilinear f on vector space V, we write for any $v, w \in V$
$f(v, w) = [v]_B^t [f]_B [w]_B$ where $[f]_B = (a_{ij})$ & $a_{ij} = f(v_i, v_j)$.

Further, the map $f \mapsto [f]_B$ is a vector space isomorphism from the space of all linear forms $(V \otimes V)^*$ to $F^{n \times n}$, where $n = \dim(V)$.

Remark:

The isomorphism $(V \otimes V)^* \cong F^{n \times n}$ depends on the choice of basis B while $V^* \otimes V \cong \mathrm{End}(V)$ doesn't

Let B & B' be basises for vector space V. Let P be the transition matrix from B' to B, that is
$[v]_B = P[v]_{B'}$ $\forall v \in V$.

We claim then that for a bilinear form
$[f]_{B'} = P^t [f]_B P$.

Thm:

Given basis B & B' for V w/ transition matrix P from B' to B,
For any bilinear form f
$[f]_{B'} = P^t [f]_B P$

Pf:

Consider $v, w \in V$. We know
$$f(v,w) = [v]_B^t [f]_B [w]_B = [v]_{B'}^t [f]_{B'} [w]_{B'}$$

Since P is the transition matrix from B' to B, we know
$$[v]_B = P[v]_{B'} \quad \& \quad [w]_B = P[w]_{B'}.$$
Thus
$$f(v,w) = (P[v]_{B'})^t [f]_B (P[w]_{B'}) = [v]_{B'}^t P^t [f]_B P [w]_{B'}$$

Using this & the earlier equation we get
$$[f]_{B'} = P^t [f]_B P. \square$$

Rem:

For any bilinear form $f \in (V \otimes V)^*$, f is symmetric/skew-symmetric if & only if $[f]_B$ is symmetric/skew-symmetric for some basis B in V.

Indeed if $[f]_B$ is such for some basis B of V, then it is such for all basis.

Note:
$$(P^t A P)^t = P^t A^t P$$

Def:

The <u>rank of a bilinear form</u> f is the rank of any matrix representing it. That is
$$\text{rank}(f) = \text{rank}([f]_B).$$

This indeed is well-defined b/c it is independent of choice of basis. We know this b/c
$$[f]_{B'} = P^t [f]_B P \quad \text{where } P \text{ is the transition matrix } B' \to B.$$
Since P is the transition matrix, P & P^t are non-singular (i.e. invertible). Thus they do not affect the rank.

Rem:

A linear form is non-degenerate/non-singular if
$$\text{rank}(f) = \dim(V).$$
In other words $[f]_B$ is invertible.

Def:

If you fix an input vector on the bilinear form $f \in (V \otimes V)^*$, you get a linear functional, that is an element of V^*.

In particular you get two, a left one & a right one.

$L_f(v): V \to F$ $\quad [L_f(v)](w) = f(v,w)$
$R_f(w): V \to F$ $\quad (R_f(w))(v) = f(v,w)$ $\Big\}$ $(L_f(v))(w) = f(v,w) = (R_f(w))(v)$.

Notice that $L_f: V \to V^*$ & $R_f: V \to V^*$

Recall:
For any linear transformation T, $\operatorname{rank}(T) = \dim(Im(T)) = \operatorname{rank}([T]_B)$ where B is any basis.

Thm:
Pick a basis $B = \{v_1, \ldots, v_n\}$ for V & dual basis $B^* = \{g_1, \ldots, g_n\}$ for V^*.
Then
$$[f]_B = [L_f]_B^{B^*} = [R_f]_B^{B^*}$$

This can be proved by the construction of $[f]_B$ & $[L_f]_B^{B^*}$ & $[R_f]_B^{B^*}$.

Coro:
$\operatorname{rank}(f) = \operatorname{rank}(L_f) = \operatorname{rank}(R_f)$.

Thm:
A bilinear form $f: V \times V \to F$ (or $f \in (V \otimes V)^*$) is non-generate/non-singular iff
iff $L_f: V \to V^*$ is an isomorphism
iff $R_f: V \to V^*$ is an isomorphism

Equivalently $\forall v \in V\ v \neq 0\ \exists w \in V$ st $f(v,w) \neq 0$
& $\forall w \in V\ w \neq 0\ \exists v \in V$ st $f(v,w) \neq 0$. $\Big\}$ From inner products, there is no vector orthogonal to all other vectors

Def:
The kernel/radical of a bilinear form f is defined by
$$\ker(f) = \underbrace{\{v \in V \mid f(v,w) = 0\ \forall w \in V\}}_{L_f} = \underbrace{\{w \in V \mid f(v,w) = 0\ \forall v \in V\}}_{R_f}$$

Rem:
If f is symmetric, then $R_f = L_f$.

If f is skew-symmetric, then $R_f = -L_f$.

Symmetric Bilinear Forms

Recall:
A bilinear form $f \in (V \otimes V)^*$ is symmetric iff $f(v,w) = f(w,v)$.

Thm:
Let V be a vector space where $n = \dim(V) < \infty$ & $\operatorname{char}(F) = 0$ (i.e. the characteristic of the field is 0, i.e. you can divide by any non-zero elements)

Then for all symmetric bilinear forms f there exists a basis B for V st $[f]_B$ is diagonal.

Pf:
This proof is done by induction on the dimension of V.

Trivially, if $\dim(V)=1$, then $[f]_B$ is diagonal b/c any non-zero 1×1 matrix is diagonal.

Suppose now that this is true for some subspace W' w/ dimension k. We show it is true for a superspace W w/ dimension $k+1$. That is $W' \subset W$ & W' is a hyperspace/hyperplane of W. (We don't do exactly this.)
We pick some non-zero vector $w \in W$ that is not orthogonal to itself, that is $f(w,w) \neq 0$. ~~$\forall$ of W.~~

Suppose that $f(v,v)=0$ $\forall v \in V$. Then

$$f(v+w, v+w) = f(v,v) + f(w,w) + 2f(v,w) \leftarrow \text{polarization identity}$$
$$= 2f(v,w)$$

But $f(v+w, v+w)=0$. Thus $f(v,w)=0$ for any $v, w \in V$.
Thus $[f]_B = 0$.

Now pick some $v \in V$ st $f(v,v) \neq 0$. Consider $W = Fv = \text{span}\{v\}$ & $W^\perp = \{w \in V \mid f(v,w)=0\}$. then $V = W \oplus W^\perp$ since $\forall u \in V$

$$u = \underbrace{\frac{f(u,v)}{f(v,v)} v}_{W} + \underbrace{w}_{W^\perp}.$$

Find $f(u,v)$.

$$f(u,v) = \frac{f(u,v)\cancel{f(v,v)}}{\cancel{f(v,v)}} + f(v,w)$$

So $f(v,w)=0 \Rightarrow w \in W^\perp$ & $W \cap W^\perp = \{0\}$. ← v not orthogonal to self.

Then by induction $W^\perp$ has a basis $\{v_1, \dots, v_{n-1}\}$ st
$f(v_i, v_j)=0$ $\forall i \neq j$.
And by definition we know when we add $v_n = v$, we again have
$f(v_i, v_j) = 0$. $\forall i \neq j$.
Thus we have basis $B = \{v_1, \dots, v_n\}$. □

Coro:
If $\text{char}(F)=0$, then for all $A \in F^{n\times n}$ where $A = A^t$ (i.e. for all symmetric matrices), there exists an invertible $P \in F^{n \times n}$ st P^tAP is diagonalizable.

Coro:
1) For the complex numbers $F = \mathbb{C}$, there is always a basis B st $[f]_B = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{bmatrix}$ where $\text{rank}(f) = r = \text{number of } 1\text{s}$

d) For the real numbers $F=\mathbb{R}$, there is always a basis B st

$$[f]_B = \begin{bmatrix} 1_{\ddots r_+} & & & 0 \\ & 1 & & \\ & & -1_{\ddots r_-} & \\ & & & -1 \\ 0 & & & 0_{\ddots 0} \end{bmatrix}$$

where $\operatorname{rank}(f) = r_+ + r_-$

& $r_+ - r_-$ is called the signature of f.

If you take a vector space $\mathbb{R}^n$ w/ an inner product f w/ such a Gram matrix $[f]_B$, you get a space called $\mathbb{R}^{r_+, r_-}$.

In particular the Minkowski space is $\mathbb{R}^{3,1}$! The Minkowski space is useful in relativity.

Review

Example: Jordan Form

When asked about Jordan form, you find the characteristic polynomial to find the eigenspaces & their basis.

Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 8 & -9 \\ 2 & 4 & -4 \end{bmatrix}$. We find char. polyn. $f(x)$

$$f(x) = \det(xI - A) = \begin{vmatrix} x-2 & 0 & 0 \\ -3 & x-8 & 9 \\ -2 & -4 & x+4 \end{vmatrix} = (x-2)((x-8)(x+4)+36)$$

$= (x-2)(x^2-4x+4)$

$= (x-2)^3$

This gives use one eigenvalue 2. We find eigenvectors / a basis for this eigenspace

eigenvectors/basis vectors

To find such x, we solve $(A-2I)x=0$. Let $N = A-2I$.

We do a brief diversion to find the minimal polynomial b/c

We know $N^3 = (A-2I)^3 = 0$ b/c $f(A) = (A-2I)^3 = 0$ (Note: $f(x) = (x-2)^3$). We know $f(A)=0$ by the Cayley-Hamilton theorem.

We know the min. polyn. $p(x)$ divides $f(x) = (x-2)^3$. Thus to find the min polynomial, we just check if $N=0$ or $N^2=0$.

$N = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 6 & -9 \\ 2 & 4 & -6 \end{bmatrix} \neq 0$

$N^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0$

Thus we know the minimal polynomial is $p(x) = (x-2)^2$.

We know A is thus not diagonalizable b/c it's minimal polynomial is not a product of unique factors.

Now we actually solve $Nx = (A-2I)x = 0$

$$\begin{bmatrix} 0 & 0 & 0 \\ 3 & 6 & -9 \\ 2 & 4 & -6 \end{bmatrix} \begin{matrix} \\ \tfrac{1}{3} \\ R_3 - \tfrac{2}{3}R_2 \end{matrix} \quad \begin{array}{c} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \quad\ \ s \quad t \end{array} \Rightarrow x = \begin{bmatrix} -2s+3t \\ s \\ t \end{bmatrix} \Rightarrow x \in \operatorname{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus the 2-eigenspace $\ker(A-2I) = \ker(N)$ has a basis is $B' = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Now we find the Jordan form J of A. B/C $p(x) = (x-2)^2$ we know we only have 2s along the diagonal & we have 2 Jordan blocks. Thus

$$J = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

To find a basis $\mathcal{B} = \{v_1, v_2, v_3\}$ such that $J = [A]_{\mathcal{B}}$, we need vectors v_1, v_2, v_3 st

$$\begin{aligned} Av_1 &= 2v_1 + v_2 & & Nv_1 = v_2 \\ Av_2 &= 2v_2 & \iff & Nv_2 = 0 \\ Av_3 &= 2v_3 & & Nv_3 = 0 \end{aligned}$$

Under such a basis note that $[N]_{\mathcal{B}}$ is actually in rational form. Let's use the forms w/ N b/c they're simpler.

Here easily v_2, & v_3 are eigenvectors. Finding v_1 is more difficult tho. Let's start by picking in the image of N

Let's find $\mathrm{Im}(N)$. In general the image of a matrix is the span of its column / columnspace.

$$\mathrm{Im}(N) = \mathrm{span}\left\{ \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ -9 \\ -6 \end{bmatrix} \right\} = \mathrm{span}\left\{ \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} \right\}.$$

Note that $\mathrm{nullity}(N) = 2$ (from earlier) & $\mathrm{rank}(N) = 1$. This checks out by the rank-nullity theorem b/c $\mathrm{rank}(N) + \mathrm{nullity}(N) = 3$.

Take $v_2 = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} = Ne_1$. Thus $v_1 = e_1$. Now we just pick

v_3 st $\{v_2, v_3\}$ is a basis for $\ker(N)$, that is the eigenspace. Take $v_3 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$. (Could have picked either one.)

Let's find the rational form of A. Recall the cyclic decomposition theorem. By this theorem the cyclic decomposition of N is $Z(v_1; N) \oplus Z(v_3; N) = \mathrm{span}\{v_1, v_2\} \oplus \mathrm{span}\{v_3\}$.

$\ldots = A - 2I$, we know $Z(v_1; A) = Z(v_1; N) = \text{span}\{v_1, v_2\}$.
However, to put A in rational form we need basis $\{v_1, Av_1\}$ for $Z(v_1; A)$.

Thus to put A in rational form we need basis $\{u_1, u_2, u_3\}$ such that
$u_1 = v_1$
$u_2 = Au_1 = Av_1$
$u_3 = v_3$

For the upper block of the rational form (for $Z(v_1; A)$), we need to write A^2u_1 in terms of u_1 & u_2. We can do this w/ the minimum polynomial $p(x) = (x-2)^2 = x^2 - 4x + 4 = x^2 - (4x - 4)$. This gives us $A^2 = 4A - 4I$. We know this b/c $p(A) = 0$ by the Cayley Hamilton theorem.

Thus $A^2u_1 = 4Au_1 - 4Iu_1 = 4u_2 - 4u_1$, which gives us the following block for A

$\begin{bmatrix} 0 & -4 \\ 1 & 4 \end{bmatrix} \leftarrow$ companion matrix of $(x-2)^2 = x^2 - (4x - 4)$

Recall $Au_3 = Av_3 = 2v_3 = 2u_3$ b/c u_3 is an eigenvector w/ eigenvalue 2.
This gives us block
$[2] \leftarrow$ companion matrix of $x - 2 = x - (2)$

Thus our rational form for A as

$$\left[\begin{array}{cc|c} 0 & -4 & 0 \\ 1 & 4 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$$

The companion matrix blocks make sense b/c A has invariant factors $(x-2)$, $(x-2)^2$.

Example:
Consider $A = \begin{bmatrix} 2 & 4 & -6 \\ 3 & 24 & -33 \\ 2 & 16 & -22 \end{bmatrix}$. Find the Jordan form of A. (Note: A not diagonalizable)

$$f(x) = \begin{vmatrix} x-2 & -4 & 6 \\ -3 & x-24 & 33 \\ -2 & -16 & x+22 \end{vmatrix} = (x-2)(x^2 - 2x) + 4(-3x - \cancel{66} + \cancel{66}) + 6(\cancel{32} + 2x - \cancel{48})$$
$$= x(x-2)^2 \cancel{-12x} \cancel{+12x}$$
$$= x(x-2)^2$$

Since the min. poly. divides $f(x)$, we know $p(x) = x(x-2)$ or $p(x) = x(x-2)^2$.
If $A(A-2I) = 0$, then A is diagonalizable b/c the min. polyn. would be a product of unique factors. We are given A isn't diagonalizable thus $p(x) = x(x-2)^2$.

Since $p(x) = x(x-2)^2$, we know the Jordan form J of A is

$$J = \left[\begin{array}{cc|c} 2 & 0 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]$$

To find a basis for this we need a basis for $\ker(A-2I)^2$ & $\ker(A)$, that is the generalized eigenspaces. By the primary decomposition theorem we know $V = \ker(A-2I)^2 \oplus \ker(A)$.

Let $B = \{v_1, v_2, v_3\}$ be the basis st $J = [A]_B$. Then we know
$Av_1 = 2v_1 + v_2 \Leftrightarrow N_1v_1 = (A-2I)v_1 = v_2$, $N_1v_2 = 0$ & $Av_3 = N_2v_3 = 0$.

We're running low on time & the rest of the example is purely computational, so let's move on.

Example: Bilinear Forms

Consider vector space V w/ $L_1, L_2 \in V^* = V \to F$.
We know by our theorem in class that
$f(v, w) = L_1(v) L_2(w)$
is a bilinear form on V.

Prove that such a bilinear form has rank 1, that is $\text{rank}(f) = 1$, & conversely prove that every bilinear form of rank 1 can be written as such.

Recall from our isomorphism in class that
$f \leftrightarrow R_f : V \to V^*$ where $(R_f(w))(v) = f(v, w)$. ← Remember $(L_f(v))(w) = f(v, w)$.
Further recall our theorem that
$\text{rank}(f) = \text{rank}(R_f) = \text{rank}(L_f) = \dim(\text{Im}(f))$.

If $f(v, w) = L_1(v) L_2(w) = (R_f(w))(v)$. By definition of R_f
$R_f(w) = L_2(w) L_1 \in V^*$ $(L_2(w) \in F)$.
Thus $\text{Im}(R_f) = \text{span}\{L_1\}$. Thus $\text{rank}(R_f) = 1$.
But since $\text{rank}(f) = \text{rank}(R_f)$, we know $\text{rank}(f) = 1$. (Forwards)

Conversely suppose $\text{rank}(f) = 1$. Then $\text{rank}(R_f) = 1$. Thus $\dim(\text{Im}(R_f)) = 1$ & $\text{Im}(R_f) = \text{span}\{L_1\}$ for some $L_1 \neq 0$ in V^*.

Thus $R_f(w) = \alpha L_1$ where α is a scalar that depends on w.
We write $\alpha = L_2(w)$ so $L_2 : V \to F$. Thus
$R_f(w) = L_2(w) L_1$.
Since $R_f(w)$ is linear $L_2(w)$ must be linear. Thus $L_2 \in V^*$.

Thus we know $(R_f(w)) = L_2(w) L_1(v)$. But $(R_f(w)) = f(v, w)$ by definition. Thus $f(v, w)$ can be written as necessary. (Backward)

We have thus shown both directions.