PU MAT 217 Notes

Aathreya Kadambi

January 2022

Contents

1 Lecture 1: Vector Spaces

1.1 Groups

Definition (Group). A **group** $(G, \cdot)$ is a set G equipped with $\cdot : G \times G \to G$

$$(a, b) \mapsto a \cdot b \in G$$

Such that the following holds

(1) $a(bc) = (ab)c$ for all $a, b, c \in G$.

(2) $\exists$ an element $e \in G$ satisfying

$$e \cdot a = a \cdot e = a$$

for all $a \in G$.

(3) $\forall a \in G$, $\exists b \in G$ such that $ab = ba = e$ and b is called the *multiplicative inverse* of a.

Example. $(\mathbb{Z}, +)$ is a group. $(\mathbb{Q}, +)$ is a group. $(\mathbb{R}, +)$ is a group. $(\mathbb{R}\backslash\mathbb{Q}, +)$ is not a group (0 is not in the set). $(\mathbb{R}, \times)$ is not a group (0 does not have a multiplicative inverse). $(\mathbb{R}\backslash\{0\}, \times)$ is a group.

Definition (Abelian Group). A group G is called an **abelian group** if $ab = ba$ for all $a, b \in G$. All the groups in the above example are abelian groups.

Example. $\{e\}$ (e is the identity) is called the trivial group. It has one element, which is the inverse of itself.

Example. $\mathbb{Z}_2 = \{0, 1\}$. Our operation is defined as follows:

$$0 + 0 = 0$$
$$0 + 1 = 1$$
$$1 + 0 = 1$$
$$1 + 1 = 0$$

Definition (Field). A **field** $(\mathbb{F}, +, \cdot)$ is a set $\mathbb{F}$ equipped with $+ : G \times G \to G$ and $\cdot : G \times G \to G$ with $a + b \in G$ and $a \cdot b \in G$ for all $a, b \in G$ such that the following holds:

(1) $\mathbb{F}$ is an abelian group under "+" (we denote by 0 the additive identity).

(2) $\mathbb{F}\backslash\{0\}$ is an abelian group under "$\cdot$" (we denote by 1 the multiplicative identity).

(3) For any three elements $a, b, c \in \mathbb{F}$,

$$a \cdot (b + c) = ab + ac$$

Example. $(\mathbb{R}, +, \times)$ is a field. $(\mathbb{Q}, +, \times)$ is a field.

Example. $\mathbb{Z}_2 = \{0, 1\}$. We can now write the following tables:

+	0	1
0	0	1
1	1	0

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

Since both tables are symmetric over the diagonal, the groups are both Abelian. Combined with the distributive property for $\mathbb{Z}$, this is a field.

1.2 Complex Numbers

Definition (Complex Number). A complex number is a pair $(a, b) \in \mathbb{R} \times \mathbb{R}$ of real numbers. We can also denote $a + bi = (a, b)$.

Definition (Complex Addition). Consider $a + bi$ and $c + di$.

$$(a, b) + (c, d) = a + bi + c + di = (a + c) + (b + d)i = (a + c, b + d)$$

Definition (Complex Multiplication). Consider $a + bi$ and $c + di$.

$$(a, b) \cdot (c, d) = (a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i = (ac - bd, ad + bc)$$

In particular, $i \cdot i = \sqrt{-1}$.

Example. $(\mathbb{C}, +)$ is an abelian group with the identity $0 = (0, 0)$. $(\mathbb{C}\backslash\{0\}, \cdot)$ is an abelian group with the identity $1 = (1, 0)$ and the inverse exists since it is simply the result of solving the system of linear equations resulting when you solve $(a + bi)(c + di) = 1$. By distributivity for $\mathbb{R}$, $(\mathbb{C}, +, \cdot)$ is a field.

1.3 Vector Spaces

Definition (Vector Space). A triple $(V, +, \cdot)$ and a field $\mathbb{F}$ are called a **vector space** where $+$ is called addition and $\cdot$ is called scalar multiplication if

(A) $(V, +)$ is an Abelian group.

(M1) $1 \cdot u = u$ for all $u \in V$.

(M2) $a(bu) = (ab)u$ for all a, b in $\mathbb{F}$ and u in V.

(M3) $(a + b) \cdot u = a \cdot u + b \cdot u$ for all $a, b \in \mathbb{F}$ and $u \in V$.

Remark. Scalar multiplication is $\cdot : \mathbb{F} \times V \to V$. 1 is the multiplicative identity in $\mathbb{F}$. Also, notice that theres addition in $\mathbb{F}$ and addition in $(V, +, \cdot)$ which are different and multiplication in the field and scalar multiplication which are both different but we are often lazy and just notate them with the same symbol.

Example. $\mathbb{F}^n = \{(a_1, \ldots, a_n) \mid a_j \in \mathbb{F}\ 1 \leq j \leq n\}$

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

$$\lambda(a_1, \ldots, a_n) = (\lambda a_1, \ldots, \lambda a_n)$$

for all $\lambda \in \mathbb{F}$ and $(a_1, \ldots, a_n) \in \mathbb{F}^n$. $\mathbb{F}^n$ is a vector space over $\mathbb{F}$.

Example. $\mathbb{C}$ is a vector space over $\mathbb{R}$.

2 Lecture 2: Subspaces

Lemma. Let G be a group. Then

(1) e is unique

(2) any element a has a unique inverse in G.

Proof.

(1)

Suppose e_1 and e_2 are two distinct elements. Then

$$e_2 = e_1 \cdot e_2 = e_2 \cdot e_1 = e_1$$

(2)

Suppose a has two inverses v and w.

$$w = we = w(av) = (wa)v = ev = v$$

■

Corollary. Let V be a vector space over $\mathbb{F}$.

(1) $0 \cdot U = \varnothing$ where $0 \in \mathbb{F}$ and $\varnothing \in V$ for all $u \in V$.

(2) $\lambda\varnothing = \varnothing$ for all $\lambda \in \mathbb{F}$.

Proof.

(1)

$$O \cdot u = (0 + 0) \cdot u = 0 \cdot u + 0 \cdot u$$
$$(-0u) + (0u) = (-0u) + (0u) + (0u)$$
$$\varnothing = 0 \cdot u$$

■

(2)

Same idea

Corollary. $(-1) \cdot u = -u$ for all $u \in V$.

Example. $\mathbb{F}^n = \{(x_1, \ldots, x_n) \mid x_j \in \mathbb{F}\ 1 \leq j \leq n\}$

$$(x_1, \ldots, x_n) + (y_1, \ldots y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$$

with $\lambda \in \mathbb{F}$.

$$\mathbb{F}^\infty = \{(x_1, x_2, \ldots) \mid x_j \in \mathbb{F},\ j \in \mathbb{Z}^+\}$$

is a vector space.

Example. Given a set S, a field $\mathbb{F}$, define set $\mathbb{F}^S$, all functions $f : S \to \mathbb{F}$,

$$(f+g)(x) = f(x) + g(x)$$

$$(\lambda f)(x) = \lambda \cdot f(X)$$

Example. $(\mathbb{F}_2)^S$ where $S = \{0, 1, \dots, n-1\}$ has 2^n elements. $(\mathbb{F}_3)^S$ where $S = \{0, 1, \dots, n-1\}$ has 3^n elements.

Remark. $\mathbb{F}_2$ is just a field with two elements.

Definition (Subspace). Let V be a vector space over $\mathbb{F}$. A subset $W \subseteq V$ is a **subspace** if W is a vector space over F.

Remark. The above might sound confusing, sometimes when we refer to the vector space we just refer to its set and not its operations for ease.

Example. $W = \{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{F}\}$, W is a subspace of $\mathbb{F}^3$ because

$$(x_1, x_2, 0) + (y_1, y_2, 0) = (x_1 + y_1, x_2 + y_2, 0) \in W$$

$$(x_1, x_2, 0) = (\lambda x_1, \lambda x_2, 0) \in W$$

Proposition. Let V be a vector space over $\mathbb{F}$. A subset $W \neq \emptyset \subseteq V$ is a subspace if and only if

(1) $\varnothing \in W$

(2) $u + v \in W$ for all u and v in W.

(3) $\lambda u \in W$ for all λ in $\mathbb{F}$ and u in W.

Example. Define $\mathbb{R}^\infty = \{(x_1, \dots, x_n, \dots) \mid x_j \in \mathbb{R}, j \in \mathbb{Z}_{>0}\}$

$$l_2(\mathbb{R}) = \{(x_1, \dots, x_n, \dots) \mid \sum_{j=1}^{\infty} |x_j|^2 < \infty\}$$

We can check that $l_2(\mathbb{R})$ is a subspace:

- $\varnothing = (0, \dots, 0, \dots) \in l_2(\mathbb{R})$
- $\lambda(x_1, \dots, x_n, \dots) = (\lambda x_1, \dots, \lambda x_n, \dots)$ since

$$\sum_{j=1}^{\infty} (\lambda x_j)^2 = \lambda^2 \sum_{j=1}^{\infty} (x_j)^2 < \infty$$

- Given $(x_1, \dots, x_n, \dots) \in l_2(\mathbb{R})$ and $y_1, \dots, y_n, \dots) \in l_2(\mathbb{R})$,

$$\sum_{j=1}^{\infty} (x_j + y_j)^2 \leq \sum_{j=1}^{\infty} (2x_j^2 + 2y_j^2) = 2\sum_{j=1}^{\infty} x_j^2 + 2\sum_{j=1}^{\infty} y_j^2 < \infty$$

Example. $\mathbb{R}^3$, N all of the solutions (x_1, x_2, x_3) of

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0 \end{cases}$$

N is a subspace of $\mathbb{R}^3$.

Proposition. V is a vector space, and suppose U and W are subspaces of V.

- $U \cup W \neq \emptyset$ is a subspace of V
- If U and W do not contain each other, $U \cup W$ is not a subspace.

Remark. The union is not a linear operation because it doesn't preserve the linear structure of the space.

Question: Given subspaces U and W of V, construct the smallest subspace that contains $U \cup W$.

Remark. All linear combinations.

Definition. Let U and W be subsets of V. Then $U + W = \{u + w \mid u \in U, w \in W\}$.

Proposition. U and W are subspaces of V. Then $U + W$ is the smallest subspace that contains $U \cup W$.

Proof.
Notice that $U \subseteq U + W$ since $0 \in W$, and $W \subseteq U + W$ since $0 \in U$. Thus, $U \cup W \subseteq U + W$.
For all $a + b \in U + W$ and $c + d$ in $U + W$,

$$(a + b) + (c + d) = (a + c) + (b + d) \in U + W$$

$$(a + b \in U + W \Rightarrow \lambda(a + b) = \lambda a + \lambda B \in U + W$$

Let Z be any arbitrary subspace of V containing $U \cap W$. Then $U + W \subseteq Z$.
For all $u \in U$ and $w \in W$, $u, w \in U \cup W \subseteq Z$, so $u + w \in Z$. Thus, $U + W \subseteq Z$. ∎

Example.

$$U = \{(x, 2x, y, 2y) \in \mathbb{F}^4 \mid x, y \in \mathbb{F}\}$$

$$V = \{x, 2x, y, y) \in \mathbb{F}^4 \mid x, y \in \mathbb{F}\}$$

$$U + V = \{x, 2x, y, z) \in \mathbb{F}^4 \mid x, y, z \in \mathbb{F}\}$$

3 Precept 1: Historical Motivation to Linear Algebra

3.1 Introduction to Linear Systems

Goal of Linear Algebra: Develop systematic methods to solving systems of linear equations.

Example. 5 cows and 2 sheep cost \$10. 2 cows and 5 sheep \$8. How much does 1 cow or 1 sheep cost?

Solution. Let x be the cost of a cow and y be the cost of a sheep.

$$\begin{aligned} 5x + 2y &= 10 \\ 2x + 5y &= 8 \end{aligned}$$

Dividing the first row by 5, subtracting two of the first row from the second row to eliminate x, and multiplying the second equation by $\frac{5}{21}$ we get

$$\begin{aligned} x + \frac{2}{5}y &= 2 \\ y &= \frac{20}{21} \end{aligned}$$

which then lets us get $x = \dfrac{34}{21}$.

Today, we will attempt to create an algorithm to solving systems of linear equations. We have two operations with equations:

- Multiply a row by a nonzero scalar
- Add rows
- Swap rows

Geometrically, $5x + 2y = 10$ and $2x + 5y = 8$ are lines (or similar figures), which is why its called *linear* algebra. Here, we are working with the real numbers, so $\mathbb{F} = \mathbb{R}$.

Harder Example.

$$\begin{aligned} x + y - z &= -2 \\ 3x - 5y + 13z &= 18 \\ x - 2y + 5z &= k \end{aligned}$$

Elimination on this matrix yields

$$\begin{aligned} x + z &= 1 \\ y - 2z &= -3 \\ 0 &= k - 7 \end{aligned}$$

so k must be 7 and since we can plug in any real number for z, we have infinitely many solutions. There are three possible number of solutions that a system of linear equations can have:

- 0
- 1
- infinitely many

Remark. It's important to remember that we cannot divide by zero. In generic fields that are not the real numbers, we have to be more careful about this.

Example. When does $\{(x, y, z) \mid ax + by + cz = k\} \subseteq \mathbb{F}^3$ form a subspace? When $k = 0$ because then it will pass through the origin.

3.2 Matrices, Vectors, and Gauss

Jordan Elimination

Idea: turn this process into an algorithm.

Example.

$$\begin{aligned} 4x_1 + 3x_2 + 2x_3 - x_4 &= 4 \\ 5x_1 + 4x_2 + 3x_3 - x_4 &= 4 \\ -2x_1 - 2x_2 - x_3 + 2x_4 &= -3 \\ 11x_1 + 6x_2 + 4x_3 + x_4 &= 11 \end{aligned}$$

we encode this into a data structure.

Definition (Matrix). An $n \times m$ matrix over a field $\mathbb{F}$ is a rectangular array

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \dots & a_{nm} \end{bmatrix}$$

with $a_{ij} \in \mathbb{F}$ for all i and j. Matrices A and B are equal when their size and entries are all the same. We also have special types of matrices:

- If $n = m$, A is a square matrix
- If A is square and $a_{ij} = 0$ whenever $i \neq j$, A is a diagonal matrix
- If $a_{ij} = 0$ whenever $i > j$, A is upper triangular
- If $a_{ij} = 0$ whenever $i < j$ A is lower triangular
- If $a_{ij} = 0$ for all i and j, A is a zero matrix

Example.

$$\begin{aligned} x + y - 2z &= 5 \\ 2x + 3y + 4z &= 2 \end{aligned}$$

We turn this into a matrix of coefficients: the coefficient matrix.

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

and we can make this into an *augmented* matrix with the other sides of the equation:

$$\left[\begin{array}{ccc|c} 1 & 1 & -2 & 5 \\ 2 & 3 & 4 & 2 \end{array}\right]$$

We can now do the same operations that we did on regular matrices to augmented matrices to solve systems. We get:

$$\left[\begin{array}{ccc|c} 1 & 0 & -10 & 13 \\ 0 & 1 & 8 & -8 \end{array}\right]$$

The reasons that this matrix is so nice are that one, the leftmost nonzero entries are 1, the leftmost nonzero entries are alone in their columns, and the leftmost nonzero entries form a staircase.
The solutions to this matrix are:

$$\left\{ \begin{bmatrix} 13+10t \\ -8-8t \\ t \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; t \in \mathbb{R} \right\}$$

An algorithm for solving linear equations is called Gauss-Jordan elimination. The idea is that we work equation by equation top to bottom. Suppose you've dont all previous equations and you get to the ith equation:

$$cx_j + \ldots = b$$

where c is nonzero. We divide by c, to make the row

$$x_j + \ldots = \frac{b}{c}$$

Finally, we eliminate x_j from all other equations by subtracting multiples of this row. Finally, go to the next equation. The algorithm stops if either you get zero = nonzero, a contradiction, or you get a consistent system and rearrange the equations to get the staircase shape.

On the matrix side, we call these steps "elementary row operations". They are:

1. Divide a row by a nonzero scalar
2. Subtract a multiple of one row from another one
3. Rearrange rows

If M is the augmented matrix you start with, the output of the algorithm is called $\operatorname{rref}(M)$ ("reduced row echelon form").

Definition (Reduced Row Echelon Form). A matrix is in **reduced row echelon form** if it satisfies the leftmost nonzero entries are 1, the leftmost nonzero entries are alone in their columns, and the leftmost nonzero entries form a staircase.

Example.

$$\begin{aligned} x+y &= 1 \\ 2x-y &= 5 \\ 3x+4y &= 2 \end{aligned}$$

The associated matrix is

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & -1 & 5 \\ 3 & 4 & 2 \end{array}\right]$$

Eliminating, we go through the following steps: (setting $\mathbb{F} = \mathbb{R}$)

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

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

4 Lecture 3:

Definition. Let $U_1, \ldots, U_m \subseteq V$. We define

$$U_1 + \ldots + U_m = \{w_1 + \ldots w_m \mid w_j \in U_j,\ 1 \leq j \leq m\}$$

Lemma. Let $U_1, \ldots U_m$ subspaces of V. Then $U_1 + \ldots + U_m$ is the smallest subspace that contains $\cup_{j=1}^{m} U_j$.

Definition. Let $U_1, \ldots, U_m \in V$

(1) $\lambda_1 U_1 + \lambda_2 U_2 + \ldots + \lambda_m U_m$ is called a linear combination ($\lambda_i \in \mathbb{F},\ 1 \leq j \leq m$).

(2) $\text{span}(U_1, \ldots, U_m) = \{\lambda_1 U_1 + \lambda_2 U_2 + \ldots + \lambda_m U_m \mid \lambda_i \in \mathbb{F}, 1 \leq j \leq m\}$

(3) $\text{span}() = \{\emptyset\}$

Lemma. Let $U_1, \ldots, U_m \in V$. Then $\text{span}(U_1, \ldots, U_m)$ is the smallest subspace that contains $U_1, \ldots, U_m$.

Remark. $\text{span}(U_1, \ldots, U_m) = \sum_{i=1}^{m} \text{span}(U_i)$

Definition.

(1) A vector space V is called **finite dimensional** if there exists a subset $S = \{V_1, \ldots, V_m\}$ such that $V = \text{span}(V_1, \ldots, V_m)$.

(2) A vector space V is called **infinite dimensional** if it is not finite dimensional.

Example.

(1) Given a field $\mathbb{F}$, a function $P : \mathbb{F} \to \mathbb{F}$, $P(Z) = a_m Z^m + a_{m-1} z^{m-1} + \ldots + a_1 z + a_0$, $0 \neq a_m \in \mathbb{F}$, $a_j \in \mathbb{F}$ for all $0 \leq j \leq m$ is called a polynomial. There is a one to one correspondance from $P(Z)$ to $(a_0, a_1, \ldots, a_m)$.

$P(\mathbb{F})$ is the space of all polynomials over $\mathbb{F}$.
This is infinitely dimensional because suppose it is finitely dimensional, then the polynomial with a degree larger than all of the degrees in the finite set that spans it is not in the span, a contradiction.

$P(\mathbb{F})$ is isomorphic to the space of all tuples with finitely many elements (only the first value must be nonzero).

(2) $C[a, b]$ continuous functions on $[a, b]$ $f : [a, b] \to \mathbb{R}$ is infinitely dimensional.

(3) $R[a, b]$ Riemann integrable functions on $[a, b]$ is infinintely dimensional since it is a superset of $C[a, b]$.

Example.

$$U \equiv \{(x, 2x, y, 2y) \in \mathbb{F}^4 \mid x, y \in \mathbb{F}\}$$
$$V \equiv \{(x, 2x, y, y) \in \mathbb{F}^4 \mid x, y \in \mathbb{F}\}$$
$$U + V = \{(x, 2x, y, z) \in \mathbb{F}^4 \mid x, y, z \in \mathbb{F}\}$$

Definition (Direct Sum). $U_1 + \ldots + U_m$ is a **direct sum** if for all $w \in U_1 + \ldots + U_m$, there exists a unique representation $w = w_1 + \ldots w_m$ where $w_j \in U_j$, for $1 \leq j \leq m$.

Proposition. Let $U_1, \ldots, U_m \subseteq V$ subspaces. Then the following are equivalent:

(1) $U_1 + \ldots + U_m$ is a direct sum.

(2) If $w_1 + \ldots w_m = 0$, $w_j \in U_j$ with $1 \leq j \leq m$ then $w_j = 0$ for all $1 \leq j \leq m$.

Proof of one direction.

$$w = w_1 + \ldots + w_m$$

$w_j \in U_j$

$$w = \hat{w_1} + \ldots + \hat{w_m}$$

$\hat{w_j} \in U_j$

$$0 = (w_1 - \hat{w_1}) + \ldots + (w_m - \hat{w_m})$$

with $(w_j - \hat{w_j}) \in U_j$ for all j. Applying 2,

$$w_j = \hat{w_j}$$

for all j, so the sum is a direct sum. ∎

Lemma. Let $U, V \in W$ be subspaces. Then (1) is equivalent to (2):

(1) $U + V$ is a direct sum

(2) $U \cap V = \{0\}$

Proof.
First we show that (1) implies (2). Taking $z \in U \cap V$,

$$z = u \in U \subseteq U + V$$

$$z = v \in V \subseteq U + V$$

$$0 = u - v \in U + V$$

$$u = v = 0$$

so $z = 0$. We now show that (2) implies (1).

$$u + v = 0 \Rightarrow u = v = 0$$

$$u = -v \Rightarrow u, v \in U \cap V$$

By (2), $u = v = 0$. ∎

We can use this idea to make a stronger proposition than before:

Proposition. Let $U_1, \ldots, U_m \subseteq V$ subspaces. Then the following are equivalent:

(1) $U_1 + \ldots + U_m$ is a direct sum.

(2) If $w_1 + \ldots w_m = 0$, $w_j \in U_j$ with $1 \leq j \leq m$ then $w_j = 0$ for all $1 \leq j \leq m$.

(3) $U_i \cap \sum_{j \neq i} U_i = \{0\}$, $1 \leq i \leq m$.

Example.

$$U_1 = \{(x, x+y, 0) \in \mathbb{F}^3 \mid x, y \in \mathbb{F}\}$$
$$U_2 = \{(0,0,z) \in \mathbb{F}^3 \mid z \in \mathbb{F}\}$$
$$U_3 = \{0, y, y) \in \mathbb{F}^3 \mid y \in \mathbb{F}\}$$

$U_1 + U_2 + U_3$ is not a direct sum since

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

$U_2 + U_3 = \{(0, y, z) \in \mathbb{F}^3 \mid y, z \in \mathbb{F}\}$. Notice that

$$U_1 \cap (U_2 + U_3) \neq \{(0,0,0)\}$$

$$U_i \cap U_j = \{(0,0,0)\}$$

for all $i \neq j$.

V is a finite dimensional vector space over $\mathbb{F}$.

Definition. Given $V_1, \dots, V_m \in V$,

(1) They are linearly independent if $\lambda_1 V_1 + \dots + \lambda_m V_m = 0$, $\lambda_i \in \mathbb{F}$ implies that $\lambda_i = 0$ for all i.

(2) They are linearly dependent if there exist scalars $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ not all zero such that $\lambda_1 V_1 + \dots + \lambda_m V_m = 0$.

By convention, a collection of zero vectors is linearly independent.

Lemma (Linear Dependence Lemma). Suppose $v_1, \dots, v_m$ are linearly dependent. Then there exists $1 \leq j \leq m$ such that

(1) $U_j \in \text{span}(U_1, \dots, U_{j-1})$.

(2) $\text{span}(U_1, \dots, U_m) = \text{span}(U_1, \dots, U_{j-1}, U_{j+1}, \dots, U_m)$

Proof.

There exist scalars $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ not all zero such that $\lambda_1 U_1 + \dots + \lambda_m U_m = 0$. Let j be the maximal one such that $\lambda_j \neq 0$. $U_j = -\dfrac{\lambda_1 U_1 + \dots + \lambda_{j-1} U_{j-1}}{\lambda_j} \in \text{span}(U_1, \dots, U_{j-1})$.

Proposition. $V = \text{span}(w_1, \dots, w_m)$ Let $u_1, \dots, u_k$ be linearly independent. Then $k \leq m$.

Proof.

$$V = \text{span}(U_1, w_1, w_2, \dots, w_m)$$

By the Linear Dependence Lemma,

$$\text{span}(U_1, w_1, \dots, w_m)$$
$$\text{span}(U_1, U_2, w_1, \dots, w_{j-1}, w_{j+1}, \dots w_m)$$
$$= \text{span}(U_1, U_2, \{w_1, \dots, w_m\} \backslash \{w_{j_1}, w_{j_2}\})$$
$$\vdots$$
$$\Rightarrow m \geq k$$

Corollary. V is finite dimensional. If $U \subseteq V$ is a subspace, then U is finite dimensional.

Proof.

If U is the zero space, we are done. Suppose U is not. Suppose U is $\text{span}(v_1)$ with $v_1 \in U \cap V$, then we are done. Otherwise,

5 Lecture 4: Bases & Dimension

Bases is the plural form of basis. For this course we will work with finite dimensional vector spaces, which have finitely many vectors that can be used to generate (span) them.

Definition (Basis). If $v_1, v_2, \ldots, v_m$ are linearly independeent, they are called a **basis** of $\text{span}(v_1, v_2, \ldots, v_m)$.

Example.

(1) $P_m(\mathbb{F})$: all polynomials with degree less than or equal to m. A basis for this space is $\{1, x, x^2, \ldots, x^m\}$. Since for the following statement to hold for all x,

$$\lambda_0 + \lambda_1 x + \ldots + \lambda_m x^m = 0 \Rightarrow \lambda_i = 0 \forall i$$

they are linearly independent.

(2) $\mathbb{R}^3 : (1, 0, 0), (0, 1, 0), (0, 0, 1)$ is a basis.

(3) $N = \{(x, y, z \in \mathbb{R}^3 \mid 2x + y + z = 0\}$. A basis is $\{(-1, 1, 1), (0, -1, 1)\}$. Any vector (x, y, z) can be generated with $-x(-1, 1, 1) + (x + z)(0, -1, 1)$

Lemma. Any generating list of a finite dimensional vector space can be reduced to a basis.

Proof.
$V = \text{span}(v_1, \ldots, v_m)$. If $V = \{0\}$, done. Assume $V \neq \{0\}$. If $v_1, \ldots, v_m$ is linearly indepenent, done. Assume $v_1, \ldots, v_m$ is linearly dependent. By the Linear Dependence Lemma, there exists $1 \leq j \leq m$ such that $V = \text{span}(v_1, \ldots, v_m) = \text{span}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$. Repeating this step until the generating set is linearly dependent. ∎

Theorem (Existence of a basis). Let V be a finite dimensional vector space. Then, V has a basis.

Proof.
By definition, there exists a generating list that is finite, so there must be a basis. ∎

Corollary. Consider a finite dimensional vector space V. Let $A = \{w_1, \ldots, w_k\}$ be a set of linearly independent vectors in V. Then A can be extended to a basis of V.

Proof.
By the Existience of a basis theorem, V has a basis $\{u_1, \ldots, u_m\}$. Then

$$V = \text{span}(w_1, \ldots, w_k, u_1, \ldots, u_m)$$

Recall the following theorem:

Lemma. $\{w_1, \ldots, w_m\}$ linearly dependent. Then there exists $1 \leq j \leq m$ such that

(1) $w_j \in \text{span}(w_1, \ldots, w_{j-1})$

(2) $\text{span}(w_1, \ldots, w_m) = \text{span}(w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_m)$

Remark. I believe there's a much simpler proof. Start with the set $\{w_1, \ldots, w_k\}$, and suppose $\text{span}(w_1, \ldots, w_k) \neq V$. Then $\exists x \in V$ with $x \notin \text{span}(w_1, \ldots, w_k)$. Then, add x to the set and start again with the new set $\{w_1, \ldots, w_k, x\}$. Suppose for contradiction that this process does not end. Then, eventually the set will be the entire vector space. However, this one of its basis (since it is finite dimensional) is a subset of the above set, a contradition.

Lemma. $w_1, \ldots, w_{k-1}$ linearly independent. $w_k \notin \text{span}(w_1, \ldots, w_{k-1})$ if and ony if $w_1, \ldots, w_k$ is linearly independent.

Lemma. Let $\{w_1, \ldots, w_k\}$ be a basis of V. Then for any vector $v \in V$, there exists a unique representation

$$v = \lambda_1 w_1 + \ldots + \lambda_k w_k$$

Proof.
Suppose this is not true. Then there existis two distinct represenations

$$v = \lambda_1 w_1 + \ldots + \lambda_k w_k$$

$$v = \hat{\lambda_1} w_1 + \ldots + \hat{\lambda_k} w_k$$

$$0 = (\lambda - \hat{\lambda_1}) w_1 + \ldots + (\lambda_k - \hat{\lambda_k}) w_k$$

which means $\lambda_j = \hat{\lambda_j}$ for all j by linear independence. ∎

Theorem. Let U be a subspace of V. Then there exists a subspace W such that $V = U \oplus W$. Then we define W to be a **complement** of U.

Proof.
Let $B = \{u_1, \ldots, u_m\}$ be a basis of U. Then by the last corollary, B can be extended to a basis $\overline{B} = \{u_1, \ldots, u_m, w_1, \ldots, w_k\}$ of V. Then, define $W \stackrel{\text{def}}{=} \text{span}(w_1, \ldots, w_k)$. Then we will prove that $U \cap W = \{0\}$. $0 \neq v = \lambda_1 u_1 + \ldots + \lambda_m u_m = \mu_1 w_1 + \ldots + \mu_k w_k$. Since $\overline{B}$ is a basis of V,

$$0 = \lambda_1 u_1 + \ldots + \lambda_m u_m - \mu_1 w_1 - \ldots - \mu_k w_k$$

$$\Rightarrow \lambda_1 = \ldots = \lambda_m = \mu_1 = \ldots = \mu_k = 0$$

a contradiction. ∎

Lemma. Any two bases of a finite dimensional space has the same length.

Proof.
Consider two bases $B_1 = \{u_1, \ldots, u_m\}$ and $B_2 = \{w_1, \ldots, w_k\}$. Then using the reduction lemma, $m \geq k$ and $k \geq m$, so $m = k$. ∎

Definition (Dimension). Let V be a finite dimensional vector space. We define the dimension $\dim(V) =$ the length of nay basis of V.

Example. $N = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 2x_1 + x_2 + x_3 = 0\}$ has $\dim(N) = 2$.

Consider $y(t)$ such that $y'''(t) = 0$. The collection of all solutions has dimensionality three.

Proposition. If $U \subseteq V$ is a subspace, then $\dim U \leq \dim V$. Moreover, "$=$" if and only if $U = V$.

Proposition.

(1) Any linearly independent list with length $\dim(V)$ is a basis of V.

(2) Any spanning list with length $\dim(V)$ is also a basis of V.

Theorem. $U, V \subseteq W$ subspaces. Then

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

In particular, $U + V$ is a direct sum if and only if $\dim(U+V) = \dim U + \dim V$.

Proof.
Consider a basis $B = \{w_1, \ldots, w_k\}$ of $U \cap V$. Then we have the following extensions.

$$B_U = \{w_1, \ldots, w_k, u_1, \ldots, u_m\}$$

$$B_V = \{w_1, \ldots, w_k, v_1, \ldots, v_n\}$$

We want to show that $\dim(U+V) = m + n + k$. In other words, we would like to show that

$$D = \{w_1, \ldots, w_k, u_1, \ldots, u_m, v_1, \ldots, v_n\}$$

is a basis of $U+V$. Clearly, D is a spanning list, so we simply must show that they are linerly independent. In other words, we would like to show that

$$a_1 w_1 + \ldots + a_k w_k + b_1 u_1 + \ldots + b_m u_m = -c_1 v_1 + \ldots + (-c_n v_n)$$

implies that all coefficients are 0.

6 Precept 2:

Example. Is $T = \{(x, y, z) \mid x^2 + y^2 + z^2 = 0\} \subseteq \mathbb{F}^3$ a subspace? If $\mathbb{F} = \mathbb{R}$, $T = \{\vec{0}\}$, so T is a subspace! If $\mathbb{F} = \mathbb{C}$, $(\sqrt{w}, w, w^{3/2})$ is a solution. Also, since $(1, i, 0) \in T$ and $(0, -i, 1) \in T$ but the sum isnt in T, it is not a subspace. If $\mathbb{F} = \mathbb{F}_2$, $(x + y + z)^2 = x^2 + y^2 + z^2 = 0$, so $x + y + z = 0$. We have a subspace!

6.1 Matrix Algebra

We can look at matrix multiplication as follows:

$$A\vec{x} = \begin{bmatrix} \vec{v_1} & \dots & \vec{v_m} \end{bmatrix} = x_1\vec{v_1} + \ldots + x_m\vec{v_m}$$

We have some rules for $A\vec{x}$:

(a) $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$ for $\vec{x}, \vec{y} \in \mathbb{F}^m$.

(b) $A(k\vec{x}) = k \cdot A\vec{x}$ where k is a scalar.

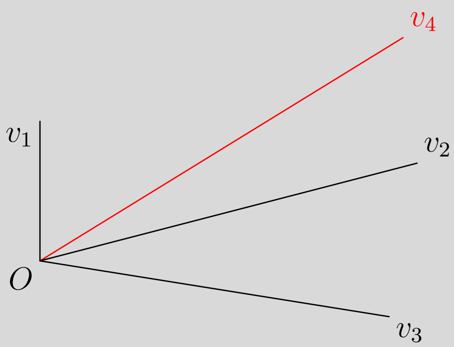
Remark. We can write a linear system with augmented matrix

$$\begin{bmatrix} A \mid \vec{b} \end{bmatrix}$$

in "matrix form" as

$$A\vec{x} = \vec{b}$$

Example.



How many ways can we represent $\vec{v_4}$ in terms of a linear combination of $\vec{v_1}$, $\vec{v_2}$, and $\vec{v_3}$? our answer is infinitely many. Proof: Note that we are working in two dimensions, so any pair of vectors in $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ is a basis for $\mathbb{R}^2$. Thus, there are at least 3 solutions, which means there must be infinitely many solutions (since any system of linear equations over $\mathbb{R}$ must have either 0, 1, or infinitely many solutions). Another way to look at this is that after row reducing our matrix for the system, we get a matrix such as

$$\left[\begin{array}{cccc|c} 1 & & & & * \\ & 1 & & & * \\ & & 1 & 0 & * \\ & & & 1 & * \\ & & & 0 & 0 \\ & & & 0 & 0 \end{array}\right]$$

and we can plug in anything we want for the variables corresponding to the zero rows.

Example. Consider the vectors

$$v_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$v_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, v_5 = \begin{bmatrix} 3 \\ 4 \\ 0 \\ 0 \end{bmatrix}, v_6 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Notice that

$$\text{span}(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}) = \text{span}(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix})$$

so the set of those three vectors is a basis for the span of all six. However, it is not a basis for $\mathbb{F}_4$.

Example. Do $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 5 \\ 4 \end{bmatrix}$ form a basis of $\mathbb{F}^3$? No since $3\vec{v_1} + \vec{v_2} = \vec{v_3}$. If we wanted to determine this computationally, we could consider the following equation:

$$a_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ 4 \end{bmatrix}$$

Trnaslating this into an augmented matrix, we have

$$\left[\begin{array}{cc|c} 1 & 3 & 6 \\ 1 & 2 & 5 \\ 1 & 1 & 4 \end{array}\right]$$

The reduced form of this matrix is

$$\text{rref}(\left[\begin{array}{cc|c} 1 & 3 & 6 \\ 1 & 2 & 5 \\ 1 & 1 & 4 \end{array}\right]) = \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

so $a_1 = 3$ and $a_2 = 1$!

Example. Consider the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$. Is this a basis for $\mathbb{F}_3$? Then we must determine whether they are linearly independent, or

$$a_1\vec{v_1} + a_2\vec{v_2} + a_3\vec{v_3} = \vec{0}$$

Considering the matrix representation and finding the reduced row-echelon form, we get:

$$\text{rref}(\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 3 & 6 & 0 \end{array}\right]) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

so the only solution is $a_1 = a_2 = a_3 = 0$. Therefore, these vectors are independent as desired.

Remark. The columns of A corresponding to columns with leading 1s in rref(a) are linearly independent.

7 Lecture 5: Dimensions and Maps

7.1 Dimensions

Proposition. $\dim V = n$, $\vec{v_1}, \ldots, \vec{v_n} \in V$ then the following are equivalent:

(1) $\vec{v_1}, \ldots, \vec{v_n}$ form a basis.

(2) $\vec{v_1}, \ldots, \vec{v_n}$ are linearly independent.

(3) $\text{span}(\vec{v_1}, \ldots, \vec{v_n}) = V$

Proof. *(Later)*
Review the following proposition:

Proposition. V is a finite dimensional vector space, and $U_1, U_2 \subseteq V$ are subspaces, then

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Proof done last class.

7.2 Maps

Definition (Map/Function/Homomorphism). Consider two vector spaces V or W over $\mathbb{F}$. We define a linear map/function/homomorphism) a function $T : V \to W$ if $\forall u, v \in V$, $\lambda \in \mathbb{F}$,

(1) $T(u + v) = Tu + Tv$

(2) $T(\lambda u) = \lambda T(u)$

For ease we sometimes staate these two conditions together with the condition $T(\lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 T(u_1) + \lambda_2 T(u_2)$.

Definition. We define $\text{Hom}(V, W)$ (V and W vector spaces over $\mathbb{F}$) as the set of all maps from V to W that are linear.

Properties of Linear Maps.

(0) $T\vec{0} = \vec{0}$ (Property 1 implies $T(u) = T(u) + T(0)$)

(1) We have a special map (zero map) which maps all vectors to zero.

(2) $I \in$ Hom(V,V), $V \to V$, $v \mapsto v$ is the identity map.

(3) $T \in \text{Hom}(F^n, F^m)$, $T \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} \in F^m$

Example.

(1) $\cdot(x + 1) : P(F) \to P(F)$ $f \mapsto (x^2 + 1)f$ is a linear map.

Remark. When thinking about vector spaces and linear mappings, it is often much more easy to analyze them when we think of vectors as linear combinations of their basis vectors, in which case we can write them with coordinates.

Notice that property (3) above allows us to think of matrices as linear maps. We will now pose the following question: does there exist a linear map $T : \mathbb{F}^2 \to \mathbb{F}^3$ such that

$$T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \text{ and } T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}?$$

If yes, how many? The answer is that there is one, and we can represent it with the matrix $\begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & -1 \end{bmatrix}$.

Proposition. Let $\{\vec{u_1}, \ldots, \vec{u_n}\}$ be a basis of V. Let $\vec{w_1}, \ldots, \vec{w_n}$ be any n vectors in W. Then there exists a unique $T : V \to W$ linear such that $T\vec{v_i} = \vec{w_i}$ for all $1 \leq i \leq n$.

Proof.
For all $v \in V$, there exists a unique way to write v as $a_1 v_1 + \ldots + a_n v_n$. Then we define T as $Tv = a_1 w_1 + \ldots + a_n w_n$. It is not hard to verify that this is a linear map.
Definition. Consider $S, T \in \text{Hom}(V, W)$. We define $S + T : V \to W$ to be the mapping $v \mapsto Sv + Tv$. We define λT to be $v \mapsto \lambda Tv$.

Using these facts, we can see that $\text{Hom}(V, W)$ is a vector space.
Compose two linear maps

$$U, V, W/\mathbb{F}$$

$$T \in \text{Hom}(U, V) \quad S \in \text{Hom}(V, W)$$

$$S \circ T : U \to W \quad u \mapsto S(Tu)$$

Properties of Compositions.

(1) $T_1 \circ (T_2 \circ T_3) = (T_1 \circ T_2) \circ T_3$

(2) $T \circ I_u = T = I_v \circ T$ $(T : U \to V)$

(3) $(S_1 + S_2) \circ T = S_1 \circ T + S_2 \circ T$, $S \circ (T_1 + T_2) = S \circ T_1 + S \circ T_2$.

Remark. If $S, T \in \text{Hom}(V, V) = \text{End}(V)$ we have $S \circ T$ and $T \circ S$ linear maps.

Remark. Warning: $S \circ T \neq T \circ S$ (usually).

> **Example.** Consider $V = \mathbb{F}^3$, e_1, e_2, e_3. $T(e_i) = e_{i+1}$, $S(e_1) = e_1$, $S(e_2) = e_3$, $S(e_3) = e_2$. Here, $T \circ S \neq S \circ T$.

Consider $T : V \to W$ a linear map.
Definition (Kernel/Null Space). $\ker T = \{v \in V \mid Tv = 0\} \subseteq V$
Definition (Image/Range). $\text{im} T = \{Tv \mid v \in V\} \subseteq W$.

Proposition. $\ker T$ is a subspace of V and $\text{im} T$ is a subspace of W.

Proposition. T is injective if and only if $\ker T = \{\vec{0}\}$.

Proof Sketch. If we have $Tv = Tw$ then it is the same as saying (there is a sort of equivalence) $T(v - w) = 0$.

Proposition. For $T \in \mathsf{Hom}(\mathbb{F}^n, \mathbb{F}^m)$, we write $Tx = Ax$ where A is an $m \times n$ matrix, and

$$\ker T (= \ker A) = \{x \mid Ax = 0\}$$

$$\mathsf{im}T(=\mathsf{im}A) = \mathsf{span}(\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \dots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}) = \mathsf{span}(Te_1, \dots, Te_n)$$

where $e_i = \begin{bmatrix} \vdots \\ 1 \\ \vdots \end{bmatrix}$ and the ith entry is a 1 and all other entries are 0s.

Definition (Surjective). T is **surjective** if $\mathrm{im}T = W$.

8 Lecture 6: Linear Maps

Lemma. Consider a linear mapping $T \in \mathcal{L}(V, W)$ ($\mathcal{L}$ is the same as Hom).

(1) T is injective if and only if Null$(T) = \{0\}$

(2) T is surjective if and only if Range$(T) = W$

Theorem (Fundamental Theorem of Linear Maps). Assume $\dim(V) < \infty$. Let $T \in \mathcal{L}(V, W)$. Then

$$\dim V = \dim(\text{Null}(T)) + \dim(\text{Range}(T))$$

Proof.
Consider a basis of Null(T): $\{u_1, \ldots, u_k\}$. Applying the extension theorem, there is exists a sequence of linearly independent elements $v_1, \ldots, v_m \in V$ such that

$$\{u_1, \ldots, u_k, v_1, \ldots, v_m\}$$

is a basis of V. We would now like to show that $T(v_1), \ldots, T(v_m)$ form a basis of W.

We will start by showing that Span$(Tv_1, \ldots, Tv_m) =$ Range(T). Consider any $v \in V$. Then $v = \lambda_1 v_1 + \ldots + \lambda_k u_k + \mu_1 v_1 + \ldots + \mu_m v_m$, so

$$Tv = \lambda_1 Tu_1 + \ldots + \lambda_k Tu_k + \mu_1 Tv_1 + \ldots + \mu_m Tv_m$$

Since $u_1, u_2, \ldots, u_k$ is a basis of Null(T),

$$Tv = \mu_1 Tv_1 + \ldots + \mu_m Tv_m$$

so Span$(Tv_1, \ldots, Tv_m) =$ Range(T) as desired.

Two show the second part, we now want to show that $Tv_1 + \ldots, Tv_m$ are linearly independent. If $k_1 Tv_1 + \ldots + k_m Tv_m = 0$, then

$$T(k_1 v_1 + \ldots + k_m v_m) = 0$$

so $k_1 v_1 + \ldots + k_m v_m$ is in the null space of T, so

$$k_1 v_1 + \ldots + k_m v_m = d_1 u_1 + \ldots + d_k u_k \Rightarrow k_1 v_1 + \ldots + k_m v_m + (-d_1) u_1 + \ldots + (-d_k) u_k = 0$$

so $k_1 = k_2 = \ldots = k_m = d_1 = \ldots = d_k = 0$ as desired since these vectors are all independent (they form a basis of V), and $Tv_1, \ldots, Tv_m$ are linearly independent. ∎

Corollary. $T \in \mathcal{L}(V, W)$. Then

(1) If T is injective, then $\dim(V) \leq \dim(W)$

(2) If T is surjective, then $\dim(W) \leq \dim(V)$

(3) If T is bijective, then $\dim(W) = \dim(V)$

Corollary.

(1) $a_{11}x_1 + \ldots + a_{1n}x_n = 0$, $a_{21}x_1 + \ldots + a_{2n}x_n = 0$, ..., $a_{m1}x_1 + \ldots + a_{mn}x_n = 0$. In matrix form, we can write:

$$\begin{bmatrix} a_1 & \ldots & a_n \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = 0$$

$n > m$ implies that this admits nonzero solutions.

(2) $a_{11}x_1 + \ldots + a_{1n}x_n = b_1$, $a_{21}x_1 + \ldots + a_{2n}x_n = b_2$, ..., $a_{m1}x_1 + \ldots + a_{mn}x_n = b_m$. $m > n$ implies that there exists $b_1, \ldots, b_m$ in $\mathbb{F}^m$ such that the above system is not solvable.

Remark. Notice that the matrix in (1) is a mapping from $\mathbb{F}^n$ to $\mathbb{F}^m$.

Proof.

(1)

If $n > m$, then the mapping is not injective, so the null space of T is not $\{0\}$, so there exists nonzero solutions.

(2)

If $m > n$, then the mapping is not surjective, so the range of T is not W, so there exists an element in W with no preimage. ∎

8.1 Matrices

We pose the following question: $\dim(V) = n$, $\dim(W) = m$, $T \in \mathcal{L}(V, W)$. Hww can we realize T as an $m \times n$ matrix? Consider $B_v = \{v_1, \ldots, v_n\}$ basis of V, and $B_w = \{w_1, \ldots, w_m\}$ basis of W. We write:

$$T(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m$$

We denote

$$\mathcal{M}(T, B_v, B_w) = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Lemma. Consider $S, T \in \mathcal{L}(V, W)$. Then

(1) $\mathcal{M}(S) + \mathcal{M}(T)) = \mathcal{M}(S + T)$

(2) $\lambda\mathcal{M}(T) = \mathcal{M}(\lambda T)$ for all $\lambda \in \mathbb{F}$

Lemma. $T_1, \ldots, T_m \in \mathcal{L}(V, W)$ are linearly independent if and only if $\mathcal{M}(T_1) \ldots, \mathcal{M}(T_m)$ are linearly independent.

Proof.

$$k_1T_1 + \ldots + k_mT_m = 0 \Leftrightarrow \mathcal{M}(k_1T_1 + \ldots + k_mT_m) = 0$$
$$\Leftrightarrow k_1\mu(T_1) + \ldots + k_m\mu(T_m) = 0$$

∎

Now, we would like to motivate matrix multiplication. Consider two linear maps $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. We can define matrix multiplication as

$$\mathcal{S} \cdot \mathcal{T} = \mathcal{M}(S \cdot T)$$

Corollary. $\dim(\text{the space of all } m \times n \text{ matrices}) = m \cdot n$.

Proof Sketch. We can simply choose the $n \cdot m$ matrices with 1s in one of the $n \cdot m$ entries and 0s everywhere else.

Corollary. $\dim(V) = m$, $\dim(W) = n$, then $\dim(\mathcal{L}(V, W)) = m \cdot n$.

Example. $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 2 & 1 & 5 \end{bmatrix}$. Compute the dimension of the range of A. Using the Fundamental Theorem of Lienar Maps,

$$\dim(\mathsf{Range}(T)) = \dim V - \dim(\mathsf{Null}(T)) = 3 - \dim(\mathsf{Null}(T))$$

so we would like to find the null space of the matrix. Performing row reduction on the matrix, we get:

$$\mathsf{rref}(\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 2 & 1 & 5 \end{bmatrix}) = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

so the dimension of the null space is 1. Therefore, $\dim(\mathsf{Range}(T)) = 3 - 1 = 2$.

9 Precept 3:

9.1 Matrices

Cpmsoder a linear map: $T \cdot V \to W$. If a basis of V is $(v_1, \ldots, v_n)$ and a basis of W is $w_1, \ldots, w_n$, then we can write

$$Tv_k = A_{1k}w_1 + \ldots + A_{mk}w_m$$

then we can write this as the matrix

$$\begin{bmatrix} A_{1k} \\ \vdots \\ A_{mk} \end{bmatrix}$$

If $V = \mathbb{F}^n$, $w = \mathbb{F}^m$, use standard basis e_1 through e_k to say that Te_j is the jth column of $\mathcal{M}(T)$.

The data of $T : V \to W$ is "equivalent" to the data of $\mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))$.

Example. $T : \mathbb{F}^3 \to \mathbb{F}^2$. Suppose it maps

$$T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}$$

$$T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \end{bmatrix}$$

$$T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -13 \\ 17 \end{bmatrix}$$

so $\mathcal{M}(T)$ is $\begin{bmatrix} 7 & 6 & -13 \\ 11 & 9 & 17 \end{bmatrix}$.

If we want to look at more geometric examples, if we wnat to see what happens to the standard basis under $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, we can just read the columns to find that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ goes to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ goes to $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$.

We can use this idea to come up with the rotation matrix for two dimensions:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Example. Suppose we have two vectors v_1 and v_2 and we have a transformation T that takes v_1 to v_1 and v_2 to $3v_2$. Then, we can write the transformation as the matrix

$$\mathcal{M}(T, (v_1, v_2), (v_1, v_2)) = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

This is basically telling us that the v_2 coordinate gets multiplied by a factor of two and the v_1 coordinate stays the same. Now suppose we have $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in standard basis. Then,

$$\mathcal{M}(T, (v_1, v_2), (e_1, e_2)) = \begin{bmatrix} 3 & 3 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

9.2 Bases for Images

Example. Consider $A = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix}$. What is a basis for $\text{im}(T)$ (T being the transformation associated with A)? Notice that

$$\begin{aligned} A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= x_1 \begin{bmatrix} 2 \\ 6 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 9 \end{bmatrix} \end{aligned}$$

Since the image of A is the span of its column vectors, the image of A is $\text{span}(\begin{bmatrix} 2 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 9 \end{bmatrix})$ which is the same as $\text{span}(\begin{bmatrix} 1 \\ 3 \end{bmatrix})$.

Example (PSet 3/A). Take $T : V \to W$ linear, with $(v_1, \dots, v_n)$ being a basis of V and $(w_1, \dots, w_m)$ being a basis of W.

$$\begin{aligned} A &= \mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m)) \\ B &= \text{rref}(A) \end{aligned}$$

with pivots (leading 1s) in columns $j_1, \dots, j_r$. Then

$$y_k = \sum_{i=1}^{m} a_{i,j_k} w_i, \quad 1 \le k \le r$$

is a basis for $\text{im}(T)$.

Proof will be done for homework.

9.3 Bases for Kernels

Example. Consider $\ker(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix})$. We would like to find a basis for the vectors such that

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

This is equivalent to solving the system of linear equations from the augmented matrix

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

When we find the row reduced echelon form of this above matrix, we get

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

so our solution is given by

$$\left\{ \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} \middle| \, t \in \mathbb{F} \right\}$$

which is given by $\text{span}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right)$, so $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ is a basis.

Notice that the vectors in $\ker(A)$ correspond to relations among the colmn vectors. In the above example, 1 times the first column $+-2$ times the second column $+$ 1 times the third column is 0.

Example (PSset 3/B). Take $T : V \to W$ linear, with $(v_1, \ldots, v_n)$ being a basis of V and $(w_1, \ldots, w_m)$ being a basis of W.

$$A = \mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))$$

$$B = \mathsf{rref}(A)$$

For every j such that the jth column of B has no pivots, consider

$$c_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \text{ and } i\text{th column of } B \text{ does not have a pivot} \end{cases}$$

This is uniquely determined by $Bc = 0$ otherwise.

$$x_j = \sum_{i=0}^{n} c_i v_i$$

for the $(n-r)$ values of j such that jth column of B has no pivot. The x_js form a basis for $\ker(T)$.

Example. Notice that

$$\mathsf{rref}\left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

Then $c_1 = 1$, $c_2 = -2$, $c_3 = 1$

10 Lecture 7: Matrices

Theorem (Rank-Nullity Theorem). Consider $T : V \to W$, with $T \in \mathcal{L}(V, W)$.

$$\dim(V) = \dim(\mathsf{Null}(T)) + \dim(\mathsf{Range}(T))$$

and the dimension of the range of T is also known as the **rank**.

If we have bases $B_V = \{e_1, \ldots, e_n\}$ of V and $V_W = \{f_1, \ldots, f_m\}$ of W, and $T(e_j) = \sum_{i=1}^{m} a_{ij} f_i$ for all j, then the matrix of T is

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Remark. Shouldnt W instead be the image of $T(V)$.

Lemma. $\mathcal{M} \in \mathsf{Hom}(\mathsf{Hom}(V, W), \mathbb{F}^{m,n})$. In other words, $\mathcal{M}$ is linear.

We will now discuss taking the products of matrices. Consider $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. We define

$$\mathcal{M}(S) \cdot \mathcal{M}(T) \stackrel{\text{def}}{=} \mathcal{M}(S \circ T)$$

Consider the bases $\{e_1, \ldots, e_n\}$ for U, $\{f_1, \ldots, f_m\}$ for V, and $\{h_1, \ldots, h_k\}$ for W. Using these, we can write

$$T(e_j) = \sum_{i=1}^{m} a_{ij} f_i$$

$$S(f_i) = \sum_{l=1}^{k} b_{li} h_l$$

$$S \circ T(e_j) = \sum_{k=1}^{k} c_{lj} h_l$$

Also

$$\begin{aligned} S \circ T(e_j) &= S(\sum_{i=1}^{m} a_{ij} f_i) \\ &= \sum_{i=1}^{m} a_{ij} S(f_i) \\ &= \sum_{l=1}^{k} (\sum_{i=1}^{m} a_{ij} b_{li}) h_l \\ c_{li} &= \sum_{i=1}^{m} a_{ij} b_{li} \end{aligned}$$

So we have essentially derived that matrix multiplication is the "row times column" multiplication that we are familiar with.

Definition. $A \in \mathbb{F}^{n,n}$ is **invertible** if and only if there exists a matrix $B \in \mathbb{F}^{n,n}$ such that $AB = BA = I_n$ (the identity matrix is the matrix with the diagonal of ones).

Remark. Notice that B must be linear since $B \in \mathbb{F}^{m,n}$ so by definition we have to check this to verify something is invertible.

Theorem. The set of invertible matrices, denoted by $\mathsf{GL}(n, \mathbb{F})$ is a group with respect to matrix product. GL is an abbreviation for "General Linear".

Notation.

$$A_{ij}: \text{ row } i, \text{ column } j$$
$$A_{i,\cdot}: \text{ entire row } j$$
$$A_{\cdot,k}: \text{ entire column } k$$

Lemma. $A \in \mathbb{F}^{m,n}$, $B \in \mathbb{F}^{n,p}$. Then $AB \in \mathbb{F}^{m,p}$ with

(1) $(AB)_{ij} = A_{i,\cdot} \cdot B_{\cdot,j} = \sum_{k=1}^{n} A_{ik}B_{kj}$

(2) $(AB)_{\cdot,k} = A \cdot B_{\cdot,k}$

Notations. Let $\{v_1, \ldots, v_m\}$ be a basis of V. Then the matrix corresponding to $v \in V$, with $v = \sum \lambda_i v_i$, is

$$\mathcal{M}(v) = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}$$

Consider $T \in \mathcal{L}(v, W)$, with a basis $B_W = \{f_!, \ldots, f_n\}$. Then

$$\mathcal{M}(T(v_k)) = \mathcal{M}(\sum_{i=1}^{n} a_{ik} f_i)$$
$$= \mathcal{M}(T)_{\cdot,k}$$

Also,

$$\mathcal{M}(TV) = \mathcal{M}(T) \cdot \mathcal{M}(T)$$

We will now discuss invertible maps.

Definition (Invertible Linear Map). Consider $T \in \mathrm{Hom}(V, W)$. T is **invertible** if and only if there exists another linear map $S \in \mathrm{Hom}(W, V)$ such that TS is the identity of the space W and ST is the identity of the space V. S is called the **inverse** of T.

Lemma. Any invertible linear map has a unique inverse.

Proof.
Consider linear map T from V to W. Suppose T has two inverses S_1, S_2. Then

$$S_1 = S_1 1_W = S_1 T S_w = 1_V S_2 = S_2$$

where 1_W and 1_V are the identity maps on W and V respectively. ■

Lemma. Consider $T \in \mathrm{Hom}(V, W)$. T is invertible if and only if T is bijective.

Proof.
We will first show the forward direction. We start by showing that T is injective. Let $v \in \mathrm{Null}(T)$. $Tv = 0$, so $T^{-1}Tv = T^{-1}0$, so $v = 0$. Thus T is injective. We now show that T is surjective. For all $w \in W$, taking $v = T^{-1}W$, $Tv = T(T^{-1}w) = w$.

We will now show the reverse direction. Consider bijective linear map $T \in \mathrm{Hom}(V, W)$. For all $w \in W$, w has a unique preimage $\hat{W} \in V$ such that $T(\hat{w}) = w$. Define $S(w) = \hat{w}$.

$$ST(v) = S(w) = v$$

so ST is the identity map on V.

$$TS(w) = T(\hat{w}) = w$$

so TS is the identity map on W. For homework, we verify that S is linear. Then S is the inverse of T as desired.

Definition. $T \in \text{Hom}(V, W)$ is a **isomorphism** if T is invertible. V, W are **isomorphic** if there exists an isomorphism $T : V \to W$.

Remark. If we instead define inverses with S any arbitrary map such that TS and ST are the respective identities, is it necessary that S is linear?

Theorem. V, W : finite dimensional vector spaces. V is isomorphic to W if and only if $\dim(V) = \dim(W)$.

Proof.
We start with the first direction. There exists an isomorphism $T : V \to W$. Then

$$\dim(V) = \dim(\text{Null}(T)) + \dim(\text{Range}(T)) = \dim(W)$$

To show the reverse direction, consider a basis of V, $B_V = \{e_1, \ldots, e_n\}$ and a basis of W, $B_W = \{f_1, \ldots, f_n\}$. There exists a unique linear map $T \in \text{Hom}(V, W)$ such that $Te_i = f_i$

(1) T injective: $0 = Tv = T(\sum_i \lambda_i e_i) = \sum_i \lambda_i T(e_i) = \sum_i \lambda_i f_i$

(2) Surjective:

Corollary. Any V with $\dim V = m$ is isomorphic to $\mathbb{F}^m$.

Corollary. $\mathcal{L}(V, W) \cong \mathbb{F}^{m,n}$, where $\dim V = n$ and $\dim W = m$.

Definition (Linear Operator). $T \in \mathcal{L}(V, V) = \mathcal{L}$ is called a **linear operator**.

Theorem. Consider $T \in \mathcal{L}(V)$. Then

$$T \text{ injective } \Leftrightarrow T \text{ surjective } \Leftrightarrow \text{ bijective}$$

Proof with $\dim(V) = \dim(\textit{Null}(T)) + \dim(\textit{Range}(T))$.
We can now discuss row operations. We have three operations:

- **Swap:** swap two rows.

Consider $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 7 & 7 \end{bmatrix}$. A row swap might yield:

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

and we can represent this with

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

- Multiplication of a row by a scalar
- Replacing rows by a linear combinations including that row

11 Lecture 8: Products and Quotients

Consider V, W vector spaces. Then
Definition (Product). $V \times W = \{(v, w) \mid v \in V, w \in W\}$.
Definition (Addition). $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$.
Definition (Scalar Multiplication). $\lambda(v, w) = (\lambda v, \lambda w)$.

Example. Consider $\mathbb{R} \times \mathbb{R}^3$. This is the set of vectors of the form $(x_1, (x_2, x_3, x_4))$, which is four dimensional.

Lemma. $\dim(V) = n$, $\dim(W) = m$ means that $\dim(V \times W) = n + m$.

Proof.

$$\{(e_i, 0), (0, f_j), 1 \leq i \leq n, 1 \leq j \leq m\}$$

is a basis. ■

Lemma. Consider $V, W \subseteq U$ subspaces with $\dim V = n$, $\dim W = m$. Then

$$\dim(V \oplus W) = n + m$$

In particular $V \times W \cong V \oplus W$.

Lemma. Let $V_1, \ldots, V_m \subseteq V$. We define a linear map $\Gamma : V_1 \times \ldots \times V_m \to V_1 + \ldots + V_m$.

$$(u_1, \ldots, u_m) \mapsto u_1 + \ldots u_m$$

Then Γ is surjective and is injective if and only if the above sum is a direct sum.

Proof.
The surjectivity of Γ follows directly from its definition. Γ is injective if and only if the null space of Γ is $\{0\}$, so $u_1 + \ldots + u_m = 0 \in V$ implies $u_i = 0$ for all i, which happens if and only if $V_1 + \ldots + V_m$ is a direct sum by definition. ■

Theorem. Let $V_1, \ldots, V_m \subseteq V$. $V_1 + \ldots + V_m$ is a direct sum if and only if

$$\dim(V_1 + \ldots + V_m) = \sum_{i=1}^{m} \dim(V_i)$$

Proof.
Applying the rank-nullity theorem,

$$\sum_{i=1}^{m} \dim(V_i) = \dim(\mathrm{Null}(\Gamma)) + \dim(V_1 + \ldots + V_m)$$

Thus $\dim(V_1 + \ldots + V_m) = \sum_{i=1}^{m} \dim(V_i)$ if and only if $\dim(\mathrm{Null}(\Gamma)) = \{0\}$, which is true if and only if $V_1 + \ldots + V_m$ is a direct sum. ■

Corollary.

$$\bigoplus_{i=1}^{m} V_i \cong \prod_{i=1}^{m} V_i$$

$$\prod_{\alpha \in A} V_\alpha = \{(v_1, v_2, \ldots) \mid \forall \alpha \in A, v_\alpha \in V_\alpha\}$$

$$\sum_{\alpha \in A} V_\alpha = \{\sum_{a \in A} v_a \mid \exists \text{ finite } \alpha \in A, v_\alpha \neq 0\}$$

Definition (Affine Subset). $U \subseteq V$ subspace. An **affine subset** may be defined with

$$v + U = \{v + u \mid u \in U\}$$

with $v \in V$.

Example. $V = \mathbb{R}^3$, $U = \{(x, y, 0) \mid x, y \in \mathbb{R}\}$. For any $w \in \mathbb{R}^3$, the affine subset $w + U$ is the plane-containing W and parallel to U.

Definition (Quotient Set). Consider the subspace $U \subseteq V$.

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

We have for all $v, w \in V$, $v \sim w$ if $v - w \in U$.

Lemma. The following are equivalent:

(1) $v + U = w + U$

(2) $v \sim w$, i.e. $v - w \in U$

(3) $((v + U) \cap (w + U) = \varnothing$

Consider $v, w \in V$ and $\lambda \in \mathbb{F}$. Denote equialence classes with brackets.

$$[v] + [w] = [v + w]$$

$$\lambda[v] = [\lambda v]$$

We can also verify well-definedness:

$$[v_1] + [w_1] = [v_1 + w_1]$$

$$[v_2] + [w_2] = [v_2 + w_2]$$

We have $v_1 \sim v_2$ and $w_1 \sim w_2$ means $[v_1 + w_1] = [v_2 + w_2]$. This is since $v_1 - v_2 \in U$ and $w_1 - w_2 \in U$ implies

$$v_1 + w_1 - (v_2 + w_2) = v_1 - v_2 + w_1 - w_2 \in U$$

Lemma. $U \subseteq V$, then V/U is a vector space.

Lemma. We define $\pi_U : V \to V/U$. Then $\pi_U \in \mathrm{Hom}(V, V/U)$ is surjective. $\mathrm{Null}(\pi_U) = U$.

Proof.
We have

$$\pi_U(w) = [w]$$

Note that $[0] = 0 + U \in V/U$ is the zero element. In other words

$$\pi_U(w) = [0] = 0 + U$$

Thus $w \in U$.

Theorem. $\dim(V/U) = \dim(V) - \dim(U)$.

Proof.

$$\dim(V) = \dim(\mathrm{Null}(\pi_U)) + \mathrm{rank}(\pi_U)$$

and since π_U is surjective, $\mathrm{rank}(\pi_U) = \dim(V/U)$. ■

Theorem. $T \in \mathsf{Hom}(V, W)$. $\mathsf{Null}(T) \subseteq V$.

$$T' : V/\mathsf{Null}(T) \to W, [v] \mapsto Tv$$

Proof.
We would like to show that

(1) $\mathrm{Range}(T) = \mathrm{Range}(T')$

(2) $V/\mathrm{Null}(T) \cong \mathrm{Range}(T)$

Using the Rank-Nullity Theorem.

Remark. We can actually show that $(W/U)/(V/U) \cong (W/V)$.

12 Lecture 9:

Remark. I was a bit distracted so my notes for this class were a bit bad.

Consider $T \in \text{Hom}(V, W)$. Let us define $\tilde{T} : V/\text{Null}(T) \to W$, with $U \overset{\text{def}}{=} \text{Null}(T)$ and $V, W \in [v]_U$, $[v]_U \mapsto Tv$.

$$Tw = Tv + T(w - v) = Tv$$

so $\tilde{T}$ is well-defined. $\text{Null}(\tilde{T}) = \varnothing_{V/\text{Null}(T)}$. Let $[v]_U \in V/U$ satisfy $Tv = 0$ whenever $v \in textNull(T)$. Then $[v]_U = \text{Null}(T) \Leftrightarrow [v]_U$ is the zero element in $V/\text{Null}(T)$. Thus, $\tilde{T}$ is injective from $V/\text{Null}(T)$ to W.

(1) $\text{Range}(\tilde{T}) = \text{Range}(T)$

(2) $V/\text{Null}(T) \cong \text{Range}(T)$

Proof 1.
$T : V \to W$, $\dim V = \dim(\text{Null}(T)) + \dim(\text{Range}(T))$. Thus,

$$\dim(V/\text{Null}(T)) = \dim(\text{Range}(T))$$

Proof 2.
$\tilde{T}$ is the bijection from $V/\text{Null}(T)$ onto $\text{Range}(\tilde{T})$.

12.1 Category Theory

In any category, we have objects A and morphisms are the maps between any two given objects.

$$\text{Object}((V, W), f : V \to W)$$

such that $U \subseteq \text{Null}(f)$.

We have a morphism from V to W_1, f_1, a morphism from W_1 to W_2, $g \in \text{Hom}(W_1, W_2)$, and a morphism from V to W_2, f_2, such that

$$f_2 = f_1 \cdot g$$

12.2 Back on Track

Definition (Dual Space). Consider V a vector space over $\mathbb{F}$. We can define V' as $\text{Hom}(V, \mathbb{F})$, where every element in V' is called a linear function on V.

Example. $V = \mathbb{F}^n$. Given any $c_1, \ldots, c_n \in \mathbb{F}$,

$$\varphi(x_1, \ldots, x_n) = c_1 x_1 + \ldots + c_n x_n$$

$$\Rightarrow \varphi \in V$$

However with

$$\psi(x_1, x_2) = x_1 + x_2 + x_1, c_1 \neq 0$$

$$\psi \notin V'$$

Example. Take $V = \mathcal{P}(\mathbb{R})$. Define $\mathcal{D}(p) = p'(1)$. $\mathcal{D}$ is in V.

Example. $V = \mathcal{C}[0, 1]$ the set of continuous functions from 0 to 1. $I(f) = \int_0^1 g(x) f(x)\, dx$. Then $I \in V'$.

Theorem. $\dim(V) = n \Rightarrow V \cong V'$. Moreover, $\{\phi_1, \ldots, \phi_n\}$ is a basis of V'. Let $\{e_1, \ldots, e_n$ be a basis of V.

$$\phi_i(e_j) = \delta_{ij}$$

$$\delta_{ij} \overset{\text{def}}{=} \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

where δ_{ij} is called the Kronecker symbol.

Proof.

(1)

We claim that for all $\phi \in V'$, $\phi = \sum_{i=1}^{n} \lambda_i \phi_i$.

Theorem. $V \cong V'' = (V')'$. $\Phi : V \to V''$ defined by $\Phi_V(\phi) = \phi(v)$, for all $\phi \in V'$ is an isomorphism.

Proof.
Consider $\{e_1, \ldots, e_n\}$ a basis of V, with $v = \sum_{i=1}^{n} \lambda_i e_i$. Taking ϕ_i,

$$0 = \phi_i(v) = \sum_{i=1}^{n} \lambda_i \phi_j(e_i) = \sum_{i=1}^{n} \lambda_i \delta_{ji} = \lambda_j$$

.

13 Row and Column Spaces

Definition (Row Space). The **row space** of a matrix M is the space spanned by its row vectors.

Definition (Column Space). The **column space** of a matrix M is the space spanned by its column vectors.

Definition (Row Rank). The **row rank** of a matrix A is the dimension of its row space.

Definition (Column). The **column rank** of a matrix A is the dimension of its column space.

Theorem. Let $A \in \mathbb{F}^{m,n}$. Then

$$\mathrm{rowrank}(A) = \mathrm{colrank}(A) = \mathrm{rank}(A)$$

Example. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix} \in \mathbb{R}^{1,5}$. Here $\mathrm{rowrank}(A) = \mathrm{colrank}(A) = \mathrm{rank}(A) = 1$.
$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{2,3}$. Here $\mathrm{rowrank}(A) = \mathrm{colrank}(A) = \mathrm{rank}(A) = 2$.

Definition (Duel Map (Operator Adjoint)). Let $T \in \mathcal{L}(V, W)$. Its **duel map** $T' \in \mathcal{L}(W', V')$ is defined by

$$T'(\varphi) \stackrel{\text{def}}{=} \varphi \circ T, \quad \forall \varphi \in W'$$

i.e.

$$T'(\varphi)(V) = \varphi(Tv)$$

for all $v \in V$.

Example. $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$

$$p(x) \mapsto p'(x)$$

$\varphi \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathbb{R})$.

$$\varphi(p(x)) = \int_0^1 xp(x)\, dx$$

$$D'(\varphi)(p(x)) = \varphi \circ D(p(x))$$

$$= \int_0^1 xp'(x)\, dx = p(1) - \int_0^1 p(x)\, dx$$

$$= p(1) - \int_0^1 p(x)\, dx$$

Lemma. $(\ )' : \mathcal{L}(V, W) \to \mathcal{L}(W', V')$ is linear, i.e.

(1) $(T + S)' = T' + S'$

(2) $(\lambda T)' = \lambda T'$

Moreover, $S \in \mathcal{L}(V, W)$, $T \in \mathcal{L}(U, V)$.

$$(ST)' = T'S'$$

Proof.

$$\begin{aligned} (ST)'(\varphi) &= \varphi \circ (ST) \\ &= (\varphi \circ S) \circ T \\ &= (S'\varphi) \circ T \\ &= T' \circ S'(\varphi) \end{aligned}$$

■

Example. Let $A \in \mathbb{F}^{m,n}$
Definition (Transpose of a Matrix). In A^T,

$$a^T_{j,i} \overset{\text{def}}{=} a_{i,j}$$

Lemma. $A, B \in \mathbb{F}^{m,n}$. Then

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

(2) $(\lambda A)^T = \lambda A^T$

(3) $(AB)^T = B^T A^T$

Proof of (3).

$$\begin{aligned}(AB)^T_{i,j} &= (AB)_{ji} \\ &= \sum_{k=1}^n A_{jk}B_{ki} \\ &= \sum_{k=1}^n (A^T)_{kj}(B^T)_{ik}\end{aligned}$$

as desired. ■

Lemma. Let $T \in \mathcal{L}(V, W)$. Then

$$\mathcal{M}(T') = (\mathcal{M}(T))^T$$

Proof.

$$B_V = \{e_1, \ldots, d_n\}, \quad B_W = \{f_1, \ldots, f_m\}$$
$$B_{V'} = \{\varphi_1, \ldots, \varphi_n\}, \quad B_{W'} = \{\psi_1, \ldots, \psi_m\}$$
$$\varphi_i(e_j) = \delta_{ij}, \quad \psi_k(f_l) = \delta_{kl}$$

$$(\mathcal{M}(T))_{ij} = A_{ij}, \quad (\mathcal{M}(T'))_{ij} = B_{ij}.$$

Our goal is to prove that $B_{ij} = A_{ji}$.

$$T'(\psi_j) = \sum_{k=1}^n B_{kj}\phi_k$$

Taking any e_r, $1 \leq r \leq n$,

$$T'(\psi_j)(e_r) = \sum_{k=1}^n B_{kj}\phi_k(e_r) = \sum_{k=1}^n B_{kj}\delta_{kr} = B_{rj}$$

On the other hand,

$$\begin{aligned}T'(\psi_j)(e_r) &\overset{\text{def}}{=\!=} \psi_j(T(e_r)) \\ &= \psi_j(\sum_{p=1}^m A_{pr}f_p) = \sum_{p=1}^m A_{pr}\psi_j(f_p) \\ &= \sum_{p=1}^m A_{pr}\delta_{jp} \\ &= A_{jr}\end{aligned}$$

■

Definition (Annihilator). V is a vector space. Consider U a subset of V. $U^\circ = \{\varphi \in V' \mid \varphi(u) = 0,\ \forall u \in U\}$ is called the **annihilator** of U.

> **Example.** Consider $\mathbb{R}^3$. Suppose U is a line OX. The annihilator of U is all functions that map the x axis to zero.

Lemma. Consider $U \subseteq V$ a subspace. Then

$$\dim(U) + \dim(U^\circ) = \dim(V)$$

Proof.

$$B_U = \{e_1, \dots, e_m\}$$

Extended basis $B_V = \{e_1, \dots, e_m, \hat{e_1}, \dots, \hat{e_k}\}$. Dual basis $B_{V'} = \{\varphi_1, \dots, \varphi_m, \hat{\varphi}_1, \dots, \hat{\varphi}_k\}$.

$$\hat{\varphi}_i(e_j) = 0, \varphi_i(\hat{e_j}) = 0, \varphi_i(e_j) = \delta_{ij}.$$

$$\hat{\varphi}_i(\hat{e_j}) = \delta_{ij}$$

$$\varphi \in U^\circ$$

$$U^\circ = \operatorname{span}(\hat{\varphi}_1, \dots, \hat{\varphi}_k)$$

$$\varphi = \sum_{i=1}^{m} a_i \varphi_i + \sum_{j=1}^{k} b_j \hat{\varphi}_j$$

Lemma. $T \in \mathcal{L}(V, W)$.

(1) $\mathsf{Null}(T') = (\mathsf{Range}(T))^\circ$

(2) $\dim(\mathsf{Null}(T')) = \dim(\mathsf{Null}(T)) + \dim W - \dim V$

Proof.

(1)

Taking any $\varphi \in W'$,

$$0 = T'(\varphi) = \varphi \circ T \Leftrightarrow \forall v \in V, 0 = \varphi \circ T(V)$$

$$(\text{Range}(T))^\circ \subseteq \text{Null}(T')$$

$$\begin{aligned} \dim(\text{Null}(T')) &= \dim((\text{Range}(T))^\circ) \\ &= \dim W - \dim(\text{Range}(T)) \\ &= \dim W - \dim V + \dim(\text{Null}(T)) \end{aligned}$$

Lemma. $T \in \mathcal{L}(V, W)$. Then

(1) $\dim(\mathsf{Range}(T')) = \dim(\mathsf{Range}(T))$

(2) $\mathsf{Range}(T') = (\mathsf{Null}(T))^\circ$

(3) T injective if and only if T' surjective

(4) T surjective if and only if T' injective

Proof.

(1)

$$\begin{aligned}\dim(\text{Range}(T')) &= \dim W - \dim(\text{Null}(T')) \\ &= \dim W - \dim(\text{Null}(T')) \\ &= \dim V - \dim(\text{Null}(T)) \\ &= \dim(\text{Range}(T))\end{aligned}$$

(2)

$$\text{Range}(T') \subseteq (\text{Null}(T))^\circ$$
$$T'(\varphi) = \varphi \circ T$$

Taking $\phi \in \text{Null}(T)$,

$$0 = \varphi \circ T(V)$$

so

$$T'(\varphi) \in (\text{Null}(T))^\circ$$
$$\begin{aligned}\dim(\text{Range}(T')) &= \dim(\text{Range}(T)) \\ &= \dim V - \dim(\text{Null}(T)) \\ &= \dim(\text{Null}(T))^\circ\end{aligned}$$

Now we show our original theorem in one step:

$$\text{colrank}(A) = \dim(\text{Range}(A)) = \dim(\text{Range}(A')) = \dim(\text{Range}(A^T)) = \text{colrank}(A^T) = \text{rowrank}(A)$$

14 Precept 5:

14.1 Unusual Property of Quotient Spaces

I will not include the comparison between 3.E.20 and 3.E.18 from the textbook in these notes. The question we now ask is when does a map $V \to W$ "descend" to a linear map $V/U \to W$?

Theorem (Universal Property of Quotient Spaces). If $U \subset V$ is a subspace, there exists a space V/U and a linear map $\pi : V \to V/U$ satisfying the following universal property:
For every linear map $T : V \to W$ such that $U \in \text{null } T$, there exists a unique linear map $S : V/U \to W$ making the diagram commute:

.

Proof.
Define V/U and $\pi : V \to V/U$ as in 3.83 and 3.88. S exists and is unqieu by 3.E.18 and 3.E.20(b). ■

"Giving a map $T : V \to W$ such that $U \subset \text{null } T$ is the same data as giving a map $V/U \to W$"

Example 3.91d. $V/(\text{null } T) \to W$ induces an isomorphism $V/(\text{null } T) \to \text{range } T$.

Corollary. Consider $U \subset V$ a subspace. Let $\pi_1 : V \to X_1$ and $\pi_2 : V \to X_2$ be two linear maps satisfying the universal property for the quotient $\pi : V \to V/U$. Then, there exists a unique isomorphism $\phi : X_1 \to X_2$ such that for all $T : V \to W$ where $U \subset \text{null } T$, the diagram commutes.

A slogan for this is: "If two objects/maps satisfying the same UP, they are isomorphic in a unique way making all the data compatible"

14.2 Examples of Quotient Spaces

Example. $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$, $f \mapsto \dfrac{df}{dz}$.

15 Eigenvalues and Eigenvectors

Before we start, a quick distraction.

Theorem. Every polynomial $p \in \mathcal{P}(\mathbb{C})$ has a root in $\mathbb{C}$.

Corollary. Every polynomial can be factored as

$$p(z) = A(z - z_1)(z - z_2)\dots(z - z_n)$$

Ok now we go back to the main topic. The word Eigenvalues comes from the word Eigenwart. Consider $V = U \oplus W$ and a transformation T. We want to know what U satisfy $T : U \to U$?

Example. Consider $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $Te_1 = e_2$, $Te_2 = -e_1$, and $Te_3 = e_3$. Then we can write

$$\mathbb{R}^3 = \mathbb{R}^2_{x,y} \oplus \mathbb{R}_z$$

Since $\dim \mathbb{R}_z = 1$ and $T\mathbb{R}_z = \mathbb{R}_z$, any vector along $\mathbb{R}_z$ is an eigenvector. An eigenvector is a vector that only gets scaled under a transformation T.

Definition (Eigenvalue and Eigenvector). Let $T \in \mathcal{L}(V)$. $\lambda \in \mathbb{F}$ is called an **eigenvalue** if $\exists v \in V\backslash\{0\}$ such that

$$Tu = \lambda v$$

In this case, $v \in V\backslash\{0\}$ is called an eigenvector.

Example.

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

$$Ae_1 = \lambda_1 e_1$$

$$Ae_2 = \lambda_2 e_2$$

$$Ae_3 = \lambda_3 e_3$$

so e_1, e_2, e_3 are eigenvectors.

Lemma. Let $\dim(V) < \infty$. $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then the following are equivalent:

(1) λ is an eigenvalue

(2) $T - \lambda I$ is not injective

(3) $T - \lambda I$ is not surjective

(4) $T - \lambda I$ is not bijective

Proof.
λ is an eigenvalue is equivalent to there exists $v \in V\backslash\{0\}$ such that $Tv - \lambda v = 0$ which is equivalent to $(T - \lambda I)v = 0$ which is equivalent to $T - \lambda I$ is not injective. The rest being equivalent is a previous theorem. ∎

Definition (Invariant Subspace). Consider $T \in \mathcal{L}(V)$. A subspace $U \in V$ is called **invariant under** T if $T(U) \subseteq U$, i.e. $Tu \in U$ for all $u \in U$.

Proposition. Let $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues. If $v_1, \ldots, v_m$ are eigenvectors with respect to $\lambda_1, \ldots, \lambda_m$, then $v_1, \ldots, v_m$ are linearly independent.

Proof.
Suppose for contradiction that the are not linearly independent, so $v_m = \sum_{j=1}^{m-1} k_j v_j$. Then

$$Tv_m = T \sum_{j=1}^{m-1} k_j v_j$$

$$Tv_m = \sum_{j=1}^{m-1} k_j T v_j$$

$$\lambda_m v_m = \sum_{j=1}^{m-1} k_j \lambda_j v_j$$

$$0 = \sum_{k=1}^{m-1} k_j(\lambda_m - \lambda_j) v_j$$

Since $\lambda_m - \lambda_j \neq 0$ (since they are distinct), $v_1, \ldots, v_{m-1}$ must also be linearly dependent, and so on, so v_1 must be linearly dependent, a contradiction. ∎

Corollary. Let $T \in \mathcal{L}(V)$, $\dim V = n$. Then there are at most n distinct eigenvalues.

Definition (Quotient Operator). $T \in \mathcal{L}(V)$, $U \subseteq V$ invariant under T. Then

(1) Restriction operator $T|_U \in \mathcal{L}(U)$. This just means

$$T|_U(u) = TU, \; \forall u \in U$$

(2) Quotient operator $T/U \in \mathcal{L}(V/U)$.

$$T/U(v + U) = Tv + U$$

Theorem. Let V be a vector space over $\mathbb{C}$ with $\dim V = n$. Then any $T \in \mathcal{L}(V)$ has an eigenvalue in $\mathbb{C}$.

Example. Consider $P : \mathbb{C} \to \mathbb{C}$. $p(z) = \sum_{j=0}^{n} a_j z^j \;\; z \in \mathbb{C}$. Let $T \in \mathcal{L}(V)$. $p(T) = \sum_{j=0}^{n} a_j T^j$ is still a linear operator on V. $T^0 = I$, $T^3 v = T(T(Tv))$ so everything is linear.

Lemma. Given $p, q \in \mathcal{P}(\mathbb{F})$, $T \in \mathcal{L}(V)$, then

(1) $(pq)(T) = p(T)q(T)$

(2) $p(T)q(T) = q(T)p(T)$ (!! This is not true in general, only here because these are polynomials)

Proof not shown.
Proof.
Let $v \neq 0$. We take $T^0 v$, Tv, $T^2 v$, . . ., $T^n v$. Since the dimension of V is n which is less than $n + 1$, these vectors are linearly dependent. In other words, there exists $a_0, \ldots, a_n \in \mathbb{C}$ such that

$$a_0 v + a_1 Tv + \ldots + a_n T^n v = 0 \Rightarrow p(T)(v)$$

where

$$p(z) = \sum_{j=0}^{n} a_j z^j$$

is a nonconstant polynomial in $\mathcal{P}(\mathbb{C})$. Thus,

$$p(z) = A(z - \lambda_1)\dots(z - \lambda_n)$$

with $A \neq 0$. In other words

$$A(T - \lambda_1 I)\dots(T - \lambda_n I)v = 0$$

Thus, suppose all of these factors are invertible. Then taking each inverse, we have $v = 0$, a contradiction. Thus, there exists j such that $T - \lambda_j I$ is not invertible. so it is not bijective, so λ_j is an eigenvalue. ■

Example. The fact that we used $\mathbb{C}$ in the last theorem is important. Consider $V = \mathbb{R}^2$. $T(z, w) = (-w, z)$. Suppose there exists an eigenvalue λ. Then there exists z, w such that

$$(-w, z) = \lambda(z, w)$$

$$-w = \lambda z, z = \lambda w$$

However, this is impossible. Thus there is no eigenvalue.

Definition (Upper Triangular). $A = (a_{ij}) \in \mathbb{F}^{n,n}$ is called **upper triangular** if $a_{ij} = 0$ for all $i > j$.

Theorem. Let $T \in \mathcal{L}(V)$. V is a vector space over $\mathbb{C}$, with $\dim V < \infty$. Then there exists a basis $B = \{v_1, \dots, v_n\}$ such that $\mathcal{M}(T)$ with respect to B is upper triangular.

Lemma. Suppose $B_V = \{v_1, \dots, v_m\}$ is a basis of V, and $T \in \mathcal{L}(V)$. Then the following are equivalent:

(1) $\mathcal{M}(T)$ with respect to B_v is upper triangular.

(2) $Tv_j \in$ span$\{v_1, \dots, v_j\}$ for all $1 \leq j \leq n$.

(3) span$(v_1, \dots, v_j)$ is invariant under T for all $1 \leq j \leq n$.

Proof.

$$Tv_1 = a_{11}v_1 \in \text{span}\{v_1\}$$
$$Tv_2 = a_{12}v_1 + a_{22}v_2 \in \text{span}\{v_1, v_2\}$$

etc. so (1) and (2) are equivalent.

16 Lecture 11:

Theorem. Consider $T \in \mathcal{L}(V)$, V a vector space over $\mathbb{C}$ with $\dim V < \infty$. Then $\exists$ a basis B of V such that $\mathcal{M}(T, B)$ is upper triangular.

Proof.

Lemma. $B_V = \{v_1, \dots, v_m\}$ basis of V. Pick $T \in \mathcal{L}(V)$. Then the following are equivalent:

(1) $\mathcal{M}(T, B)$ is upper triangular

(2) For all $1 \le j \le m$, $Tv_j \in \text{span}\{v_1, \dots, v_j\}$.

(3) For all $1 \le j \le m$, $\text{span}\{v_1, \dots, v_j\}$ is invariant under T.

Proof done previously.

We will prove it by induction on $\dim V = n$. For our base case, consider $n = 1$. Then we are done because $\mathcal{M}(T)$ is a one by one matrix and we are done with any basis. Assume that the result holds for any V with $1 \le \dim V \le n-1$. We now show that the statement holds for any space V with $\dim V = n$.

By the existence theorem of complex eigenvalue to T, $\exists \lambda \in \mathbb{C}$ an eigenvalue of T. Let us take $U = \text{Range}(T - \lambda I)$.

(1) If $\dim U = 0$, $T \cong \lambda I$.

(2) $\dim U \ge 1$. Then since λ is an eigenvalue, $T - \lambda I$ is not surjective, so $\text{Range}(T - \lambda I)$ is strictly less than n.

We now claim that U is invariant under T. In fact, for every $w \in U$, $Tw = (T - \lambda I)w + \lambda w$. Since $(T - \lambda I)w$ is in U, and λw is in U, Tw is in U as desired.

Thus, $T|_U$ is a linear operator. Applying the induction hypothesis, there exists a basis $B_U = \{v_1, \dots, v_k\}$ of U with $1 \le k = \dim U \le n - 1$ such that $\mathcal{M}(T|_U, B_U)$ is upper triangular. Using the Lemma, $T|_U(v_j) \in \text{span}(v_1, \dots, v_j)$ for all $1 \le j \le k$. U can be extended to $B_V = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ a basis of V.

$$Tv_i = (T - \lambda I)v_i + \lambda v_i \in \text{span}(v_1, \dots, v_k, v_i) \subseteq \text{span}(v_1 m \dots, v_2, \dots, v_i)$$

for all $k + 1 \le i \le n$. ■

Remark. Couldn't we also prove this by doing some kind of column reduction on the matrix of B with respect to an arbitrary basis?

Theorem. $B = \{v_1, \dots, v_n\}$ is a bsis of V. Consider $T \in \mathcal{L}(V)$. Assume that $\mathcal{M}(T, B)$ is upper triangular. T is invertible is equivalent to each diagonal entry not equalling zero.

Proof.

We start with the reverse direction. Consider

$$\mathcal{M}(T, B) = \begin{bmatrix} \lambda_1 & & & \\ 0 & \lambda_2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

with $\lambda_1, \dots, \lambda_n \ne 0$. Since $\lambda_1 \ne 0$,

$$v_1 = \frac{\lambda_1 v_1}{\lambda_1} = \frac{1}{\lambda_1} Tv_1 \in \text{Range}(T)$$

$$Tv_2 = a_1 v_1 + \lambda_2 v_2 \Rightarrow v_2 = \frac{1}{\lambda_2}(Tv_2 - a_1 v_1) \in \text{Range}(T)$$

Repeating the above, $v_j \in \text{Range}(T)$ for all $1 \le j \le n$. Thus, $V_B = \{v_1, \ldots, v_n\}$ basis of V implies that $\text{Range}(T) = V$. Thus, T is surjective, which implies that it is invertible.

We now conisder the forward direction. Since $Tv_1 = \lambda_1 v_1$, T is invertible implies that $\lambda_1 \neq 0$. Since $Tv_2 = a_1 v_1 + \lambda_2 v_2$, if $\lambda_2 = 0$, the n$Tv_@ = a_1 v_1 \in Span(v_1)$. Thus, $\text{span}(Tv_1, Tv_2) = \text{span}(v_1)$, but this means they are not independent, a contradiction of invertibleness. Thus, $\lambda_2 \neq 0$. In general, although we won't show it again here, $\lambda_j \neq 0$ for all $1 \le j \le n$. ■

Corollary. $\mathcal{M}(T, B)$ is upper triangular. Then the eigenvalues of T are precisely the diagonal entries of $\mathcal{M}(T, B)$.

Proof.
Consider

$$\mathcal{M}(T - \lambda I, B) = \begin{bmatrix} \lambda_1 - \lambda & & & \\ 0 & \lambda_2 - \lambda & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \ldots & \lambda_n - \lambda \end{bmatrix}$$

λ is an eigenvalue is equivalent to $T - \lambda I$ is not invertible, which is equivalent to $\lambda = \lambda_i$ for some $1 \le i \le n$. ■

Definition (Diagonal Matrix). We say a matrix $A = (a_{ij}) \in \mathbb{F}^{n,n}$ is **diagonal** if and only if $a_{ij} = 0$ for all $i \neq j$.

Definition (Eigenspace). Let $T \in \mathcal{L}(V)$ with an eigenvalue $\lambda \in \mathbb{F}$.

$$E(\lambda, T) = \text{Null}(T - \lambda I)$$

is called the **eigenspace** of T with respect to λ.

Lemma. $\dim V < \infty$, $T \in \mathcal{L}(V)$ has distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Then $\sum_{j=1}^{k} E(\lambda_j, T)$ is a direct sum.

Proof.
Consider $u_j \in E(\lambda_j, T)$ for all $1 \le j \le k$. Then

$$u_1 + \ldots + u_k = 0$$

implies that $u_j = 0$ for all $1 \le j \le k$ (because of independence). So we have a direct sum. In particular,

$$\sum_{j=1}^{k} \dim E(\lambda_j, T) \le n = \dim V$$

■

Definition. $T \in \mathcal{L}(T)$ is **diagonalizale** if $\exists$ a basis B of V such that $\mathcal{M}(T, B)$ is a diagonal matrix.

Example. If $T \in \mathcal{L}(V)$, $\dim V = n$ has n distinct eigenvalues, then T is diagonalizable. Let $v_j \in V$ satisfy $Tv_j = \lambda_j v_j$ for $\lambda_i \neq \lambda_j$ when $i \neq j$. Then $B = \{v_1, \ldots, v_n\}$ is linearly independent. We have

$$V = E(\lambda_1, T) \oplus E(\lambda_2, T) \oplus \ldots \oplus E(\lambda_n, T)$$

Theorem. $T \in \mathcal{L}(V)$, $\dim V = n < \infty$, consider $\lambda_1, \ldots, \lambda_m$ distinct eigenvalues. Then the following are equivalent:

(1) T is diagonalizable

(2) V has a basis consisting of eigenvectors of T

(3) $\exists$ 1 dimensional subspaces $U_1 \ldots U_n \subseteq V$ such that

$$V = \bigoplus_{j=1}^{n} U_j$$

(4)

$$V = \bigoplus_{j=1}^{m} E(\lambda_j T)$$

(5)

$$\dim V = \sum_{j=1}^{m} \dim E(\lambda_j, T)$$

Proof Sketch.

It is easy to show by definition that 1, 2, and 3 are equivalent. Also, 2 implies 4 by the lemma and definition of basis. 4 is equivalent to 5. We now show 4 implies 2. $E(\lambda_j, T)$ has dimension d_j and basis $B_j = \{v_j^1, v_j^2, \ldots, v_j^{d_j}\}$. We would like to check that all of these base elements are linearly independent.

$$\sum_{j=1}^{m} \sum_{k=1}^{d_j} a_{jk} v_j^k = 0$$

Since $(\sum_{k=1}^{d_j} a_{jk} v_j^k) \in E(\lambda_j T)$ for each j,

$$\sum_{k=1}^{d_j} a_{jk} v_j^k = 0$$

for each j. Thus,

$$a_{jk} = 0 \ \ \forall j, k$$

so all of them are indpendent as desired, so they form a basis for the entire space, as desired. ∎

Corollary. $T \in \mathcal{L}(V)$, $\dim V = n$ has n distinct eigenvalues. Then T is diagonalizable.

16.1 Changing Bases

Lemma. B_1, B_2, B_3 bases of V. $S, T \in \mathcal{L}(V)$.

$$\mathcal{M}(ST, B_1, B_3) = \mathcal{M}(S, B_2, B_3)\mathcal{M}(T, B_1, B_2)$$

Example. $w_1 = \sum_{i=1}^n P_{i1} v_i$, $w_2 = \sum_{i=1}^n P_{i2} v_i$, etc, so $w_n = \sum_{i=1}^n P_{in} v_i$. We have

$$(w_1, \dots, w_n) = (v_1, \dots, v_n) \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

we have $P = \mathcal{M}(I, B_2, B_1)$.

Corollary. Consider B_1, B_2 bases of V. Then

$$\mathcal{M}(I, B_1, B_2) \cdot \mathcal{M}(I, B_1, B_2) = \mathcal{M}(I, B_1, B_1) = I$$

$$\mathcal{M}(I, B_1, B_2) \cdot \mathcal{M}(I, B_2, B_1) = \mathcal{M}(I, B_2, B_2) = I$$

Theorem. U, V two bases of W.

$$A = \mathcal{M}(I, U, V)$$

then $\mathcal{M}(T, U) = A^{-1} \mathcal{M}(T, V) \cdot A$.

Proof.

$$\begin{aligned} \mathcal{M}(T, U, V) &= \mathcal{M}(TI, U, V) \\ &= \mathcal{M}(T, V, V)\mathcal{M}(I, U, V) \\ &= \mathcal{M}(T, V) \cdot A \\ \mathcal{M}(T, U, V) &= \mathcal{M}(IT, U, V) \\ &= \mathcal{M}(I, U, V)\mathcal{M}(T, U, U) \\ &= A\mathcal{M}(T, U) \end{aligned}$$

$$A\mathcal{M}(T, U) = \mathcal{M}(T, V)A \Rightarrow \mathcal{M}(T, U) = A^{-1}\mathcal{M}(T, V)A$$

17 Precept 6:

17.1 Change of Basis and Diagonalization

Example. Consider the reflection about the line spanned by $\begin{bmatrix}1\\2\\3\end{bmatrix} \in \mathbb{R}^3$. We call this the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$. We don't know much about arbitrary vectors, but we do know about the vetors in the plane perpendicular to $\begin{bmatrix}1\\2\\3\end{bmatrix}$. The plane is defined by $x + 2y + 3z = 0$, since $\begin{bmatrix}1\\2\\3\end{bmatrix}$ is the normal vector to the plane. On this plane, any vector gets mapped to its additive inverse. For example, we know immediately that

$$T\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$$

$$T\begin{bmatrix}-3\\0\\1\end{bmatrix} = -\begin{bmatrix}-3\\0\\1\end{bmatrix}$$

$$T\begin{bmatrix}-2\\1\\0\end{bmatrix} = -\begin{bmatrix}-2\\1\\0\end{bmatrix}$$

Notice that the first vector above is an eigenvector with eigenvalue 1 and the second two are eigenvectors with eigenvalue -1. Also, notice that these three vectors form a basis for $\mathbb{R}^3$. We can now write the matrix of the transformation with respect to these bases:

$$\mathcal{M}(T, (\vec{v_1}, \vec{v_2}, \vec{v_3})) = \begin{bmatrix}1 & 0 & 0\\0 & -1 & 0\\0 & 0 & -1\end{bmatrix} = B$$

To find $\mathcal{M}(T, (\vec{e_1}, \vec{e_2}, \vec{e_3}))$, notice that we have a map B from the $\vec{e_i}$ to the $\lambda_i e_i$, and we can use $S = \begin{bmatrix}1 & -3 & -2\\2 & 0 & 1\\3 & 1 & 0\end{bmatrix}$ to transform from the $\lambda_i e_i$ to the Tv_i, and S^{-1} to go from the v_i to the e_i. In other words,

$$\mathcal{M}(T, (e_1, e_2, e_3)) = SBS^{-1}$$

17.2 An Eigenvalue/Eigenvector Example

Example (Exercises 5.C.16). We define the Fibonacci sequence as $F_1 = 1$, $F_2 = 1$, and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$. How do you write down a closed formula for F_n? The idea is that we can use eigenvalues and eigenvectors to solve this problem. (I love this problem)
We are going to use the following linear operator: $T \in \mathcal{L}(R^2)$, $T(x, y) = (y, x + y)$.

$$\mathcal{M}(T) = \begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}$$

We start with the following claim: $T^n(0, 1) = (F_n, F_{n+1})$ for all $n > 0$. We now induce on n (smh apparently its induce but I like induct more). When $n = 1$, $T^1(0, 1) = (1, 1) = (F_1, F_2)$ by definition. Assume $T^{n-1}(0, 1) = (F_n, F_{n+1})$. Then $T^n(0, 1) = T(T^{n-1}(0, 1)) = T((F_n, F_{n+1})) = (F_{n+1}, F_n + F_{n+1}) = (F_{n+1}, F_{n+2})$ as desired. Thus our claim is true. We now ask: what are the eigenvalues of T?

$$T(x, y) = \lambda(x, y)$$

$$(x, x+y) = (\lambda x, \lambda y)$$
$$x = \lambda x$$
$$x + y = \lambda y$$
$$x + \lambda x = \lambda^2 x \Rightarrow 0 = (\lambda^2 - \lambda - 1)x = 0$$

Since $x = y = 0$ does not yield a valid eigenvector, $\lambda^2 - \lambda - 1 = 0$. We thus have

$$\lambda = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

so we have two possible eigenvalues:

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}$$

Consider $v_1 = (1, \frac{1+\sqrt{5}}{2})$. Then

$$Tv_1 = \lambda_1 v_1$$

so λ_1 is an eigenvalue (left to the reader/watcher). Consider $v_2 = (1, \frac{1-\sqrt{5}}{2})$. Then

$$Tv_2 = \lambda_2 v_2$$

so λ_2 is an eigenvalue (left to the reader/watcher). We can now diagonalize T. The reason we want to diagonalize T is that taking power with diagonal matrices is easy (whereas it is hard for arbitrary matrices). For example,

$$\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & \\ & \lambda_2^n \end{bmatrix}$$

For this T, v_1 and v_2 form a basis consisting of eigenvectors.
We now want to compute $T^n(0,1)$ using this basis/diagonalization. We have

$$\mathcal{M}(T, (e_1, e_2)) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = A$$

$$\mathcal{M}(T, (v_1, v_2)) = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = B$$

We have

$$A = SBS^{-1}$$

where $S = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}$. Now,

$$A^n = (SBS^{-1})^n = SBS^{-1}SBS^{-1} \ldots SBS^{-1} = SBIBI \ldots IB = SB^nS^{-1}$$

We now have

$$T^n(0,1) = A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} (\frac{1+\sqrt{5}}{2})^n & \\ & (\frac{1-\sqrt{5}}{2})^n \end{bmatrix} \begin{bmatrix} \frac{5-\sqrt{5}}{10} & \frac{1}{\sqrt{5}} \\ \frac{5+\sqrt{5}}{10} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{\lambda_1^n - \lambda_2^n}{\sqrt{5}} \\ \frac{\lambda_1^{n+1} - \lambda_2^{n+2}}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix}$$

so $F_n = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n)$.

Example. Note that it is sometimes not possible to diagonalize. Consider the matrix

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Since this is an upper triangular matrix, the only possible eigenvalues are the values on the diagonal, so 1 and 3. We now compute $\dim E(1,T) + \dim E(3,T)$. Note that

$$E(1,T) = \ker \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

so $\dim E(1,T) = 1$. Also

$$E(3,T) = \ker \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so $\dim E(3,T) = 1$. Thus, $\dim E(1,T) + \dim E(3,T) = 1 + 1 \neq \dim \mathbb{R}^3$, so the matrix is not diagonalizable. Note that

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

is the smallest example of a non-diagonalizable matrix.

18 Lecture 13: Inner Product Space

Remark. In this section we will be discussing only $\mathbb{C}$ and $\mathbb{R}$ as our fields.

Definition (Normed Space). A vector space V is called a **normed space** if $||\cdot|| : V \to \mathbb{F}$ satisfies the following:

(1) $||v|| \geq 0$ with"=" if and only if $v = 0$.

(2) $||\lambda \cdot v|| = |\lambda| \cdot ||v||$ for all $\lambda \in \mathbb{F}$, $v \in V$.

(3) $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in V$.

Remark. $d(u, v) \stackrel{\text{def}}{=} ||u - v||$ is defined such that

(1) $d(u, v) = d(v, u)$ for all $u, v \in V$

(2) $d(u, v) \geq 0$, "=" if and only if $u = v$

(3) $d(u, v) \leq d(u, v) + d(v, w)$

Example. $V = \mathbb{R}^n$, $||v||_p \equiv (\sum_{j=1}^{n} |v_j|^p)^{1/p}$, $p \geq 1$. You should try to show the result by Minkowski: (Minkowski inequality)

$$||u + v||_p \leq ||u||_p + ||v||_p$$

Definition (Bilinear Form). A function $\varphi : V \times V \to \mathbb{F}$ is called a **bilinear form** if

(1) $\varphi(u, \cdot) \in V'$ for all $u \in V$

(2) $\varphi(\cdot, v) \in V'$ for all $v \in V$

Example. $V = \mathbb{R}^n$.

$$u \cdot v \stackrel{\text{def}}{=} \sum_{j=1}^{n} u_j v_j$$

Definition. Let V be a vector space over $\mathbb{R}$. A bilinear form $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ is called an **inner-product** if

(1) $\langle v, v\rangle \geq 0$ for all $v \in V$, "=" if and only if $v = 0$.

(2) $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

V is called an inner-product space.

Definition. Let V be a vector space over $\mathbb{F}$. A function $\langle \cdot, \cdot\rangle : V \times V \to \mathbb{F}$ is called an **inner-product** if

(!) For any $u \in V$, define $\Phi_u(v) \equiv \langle v, u\rangle \in V'$

(2) $\langle v, v\rangle \geq 0$ for all $v \in V$, "=" if and only if $v = 0$

(3) $\langle v, u\rangle = \overline{\langle u, v\rangle}$ for all $u, v \in V$.

V is called an **inner-product** space.

Example. $V = \mathbb{C}^n$.

$$u \cdot v \stackrel{\text{def}}{=} \sum_{j=1}^{n} u_j \bar{v}_j$$

This is called an Hermitian Product.

We now define define the norm of an inner product space:

$$||v|| = \langle v, v\rangle^{\frac{1}{2}}$$

Lemma. Consider V an inner product space.

(1) $\langle u, 0\rangle = \langle 0, u\rangle = 0$ for all $u \in V$.

(2) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u + w\rangle$ for all $u, v, w \in V$

(3) $(u, \lambda v) = \bar{\lambda}\langle u, v\rangle$, for all $u, v \in V$, $\lambda \in \mathbb{F}$.

(4) $||v|| = 0$ if and only if $v = 0$.

(5) $||\lambda \cdot v|| = |\lambda| \cdot ||v||$

Proof.

(1)

$$\langle 0, u\rangle = \langle v - v, u\rangle = \langle v, u\rangle - \langle v - u\rangle = 0$$

(2)

$$\langle u, v + w\rangle = \overline{\langle v + w, u\rangle} = \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} = \langle u, v\rangle + \langle u, w\rangle$$

Definition (Orthagonality). We define $u \perp v$ **orthogonal** if $\langle u, v\rangle = 0$.

Lemma. $0 \in V$ is orthogonal to any vector, and 0 is the only vector which is orthogonal to itself.

Proof.
The first part follows from property one of the previous Lemma. The second part follows directly from property 4 of the previous Lemma. ■

Theorem (Pythagorean Theorem). For any $u \perp v$,

$$||u + v||^2 = ||u||^2 + ||v||^2$$

Proof.

$$||u + v||^2 = \langle u + v, u + v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle = \langle u, u\rangle + \langle v, v\rangle = ||u||^2 + ||v||^2$$

■

Theorem (Cauchy-Schwarz Inequality). $|\langle u, v\rangle| \leq ||u|| \cdot ||v||$, with "=" if and only if $u = \lambda v$ for some $\lambda \in \mathbb{F}$.

Proof.

$$\langle u - \lambda v, u - \lambda v\rangle \geq 0$$

for all $\lambda \in \mathbb{F}$. The left hand side is

$$\langle u, u\rangle - \lambda\langle v, u\rangle - \bar{\lambda}\langle u, v\rangle + |\lambda|^2\langle v, v\rangle$$

Lemma. Suppose $ax^2 + bx + c \geq 0$, $a > 0$ is rue for all $x \in \mathbb{R}$. Then $b^2 - 4ac \leq 0$.

Using this lemma, if we work in $\mathbb{R}$, then

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle$$

which is the square of our desired result of the Cauchy Schwarz inequality. ■

Remark. Waw another really sweet proof of the Cauchy-Schwarz Inequality :heart_eyes:

Lemma. $u, v \in V$, $\lambda \in \mathbb{F}$, $v \neq 0$. $w = u + \lambda v$ satisfies $\langle w, v\rangle = 0$ if and only if $\lambda = -\frac{\langle u, v\rangle}{||v||^2}$.

Proof.

$$\langle w, v\rangle = 0 \Leftrightarrow 0 = \langle u, v\rangle + \lambda\langle v, v\rangle$$

with $v \neq 0$,

$$\Leftrightarrow \lambda = -\frac{\langle u, v\rangle}{\langle v, v\rangle}$$

We can now show another proof of the Cauchy-Schwarz Inequality:

Proof.

$u = w - \lambda v$ implies that $||u||^2 = ||w||^2 + |\lambda|^2||v||^2 \geq |\lambda|^2|v||^2$. We now simply plug in the result from the previous lemma to get the Cauchy Schwarz inequality. ■

Corollary. $\forall u, v \in V$,

$$|||u|| - ||v||| \leq ||u + v|| \leq ||u|| + ||v||$$

with equality if and only if $u = \lambda v$ for some $\lambda \in \mathbb{F}$.

We only prove the second part.

$$\langle u + v, u + v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle$$

$$\leq \langle u, u\rangle + 2|\langle u, v\rangle| + \langle v, v\rangle$$
$$\leq \langle u, u\rangle + 2||u|| \cdot ||v|| + \langle v, v\rangle$$
$$= (||u|| + ||v||)^2$$

Corollary. Any inner-product space is a normed space if $||v|| = \langle v, v\rangle^{\frac{1}{2}}$ for all $v \in V$.

Remark. If V is a vector space over $\mathbb{R}$ and we have a basis $\{e_1, \ldots, e_n\}$, and we write v, w in components, $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$.

$$\langle v, w\rangle = \langle\sum_{j=1}^{n} v_j e_j, \sum_{k=1}^{n} w_k e_k\rangle = \sum_{j,k=1}^{n} \langle e_j, e_k\rangle v_j w_k$$

$$= (v_1, \ldots, v_n)A \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

$$A = (a_{ij}), a_{ij} = \langle e_i, e_j\rangle$$

In the case of the dot product $a_{ij} = \delta_{ij}$, so A is the identity.

Diagram representing how abstract everything is:

Metric Space

Normed Space

Inner Product Space
A specific example is the dot product for $\mathbb{R}^n$.

Definition. A list of vectors $B = \{v_1, \ldots, v_m\}$ is **orthonormal** if $\langle v_i, v_j \rangle = \delta_{ij}$ for all $1 \leq i, j \leq m$.

Lemma. $e_1, \ldots, e_m$ orthonormal implies that

$$||\sum_{j=1}^{m} a_j e_j||^2 = \sum_{j=1}^{m} |a_j|^2$$

for all $a_j \in \mathbb{F}$.

Corollary. $e_1, \ldots, e_m$ are orthonormal implies that they are linearly independent.

Lemma. $B = \{e_1, \ldots, e_m\}$ is an orthonormal basis of V. Then for all $v \in V$,

(1) $v = \sum_{j=1}^{m} \langle v, e_j \rangle e_j$

(2) $||v||^2 = \sum_{j=1}^{m} \langle v_1, e_j \rangle^2$

Proof.

$$v = a_1 e_1 + a_2 e_2 + \ldots + a_m e_m$$

For any $1 \leq k \leq m$,

$$\langle v, e_k \rangle = \langle \sum_{j=1}^{m} a_j e_j, e_k \rangle = \sum_{j=1}^{m} a_j \langle e_j, e_k \rangle = \sum_{j=1}^{m} a_j \delta_{jk} = a_k$$

Part (2) follows from the Corollary. ∎

19 Lecture 15: Orthogonal Complements/Adjoint Operator

Let $U \subseteq V$.

$$U^\perp = \{v \in V \mid \langle v, u\rangle = 0 \; \forall u \in U\}$$

Proposition. Let $U \subseteq V$ with $\dim U < \infty$. Then $V = U \oplus U^\perp$.

Proof.
Let $\{e_1, \ldots, e_m\}$ be an orthonormal basis of U. Then for all $v \in V$, we define

$$u = \sum_{j=1}^{m} \langle v, e_j\rangle e_j$$

Now note that

$$\langle v - u, e_k\rangle = \langle v, e_k\rangle - \langle u, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle = 0$$

Thus $v = u + (v - u)$, so $V = U + U^\perp$.

Lemma.

(1) If $U \subseteq V$ is a subset, then $U^\perp \subseteq V$ is a subspace.

(2) $\{0\}^\perp = V$

(3) $V^\perp = \{0\}$

(4) If $U \subseteq V$, then $U \cap U^\perp = \{0\}$

(5) $U \subseteq W \Rightarrow W^\perp \subseteq U^\perp$

(1)

$0 \in U^\perp$ by definition. Notice that $U^\perp$ is closed under addition: if $v, u \in U^\perp$, then

$$\langle v + w, u\rangle = \langle v, u\rangle + \langle w, u\rangle = 0 + 0 = 0$$

Notice that $U^\perp$ is closed under scalar multiplication since $v \in U^\perp \Rightarrow \lambda v \in U^\perp$ for all $\lambda \in \mathbb{F}$. Thus, $U^\perp$ is a subspace. ■

Corollary. $\dim V < \infty$, $U \subseteq V$ subspace. Then $\dim V = \dim U = \dim U^\perp$.

Corollary. Let $U \subseteq V$ satisfy $\dim U < \infty$. Then $(U^\perp)^\perp = U$.

Proof.
For all $u \in U$, $\langle u, v\rangle = \langle v, u\rangle = 0$ for all $v \in U^\perp$, so $u \in (U^\perp)^\perp$. Thus $U \subseteq (U^\perp)^\perp$.
Taking any $v \in (U^\perp)^\perp$, $v = u + w$, with $u \in U$ and $w \in U^\perp$.

$$0 = \langle v, w\rangle = \langle u, w\rangle + \langle w, w\rangle = 0 + \langle w, w\rangle = \langle w, w\rangle$$

so $w = 0$. Thus $v = u \in U$, so $(U^\perp)^\perp \subseteq U$ as desired. ■

Corollary. Assume $\dim V < \infty$. $\Phi|_{U^\perp} \in \mathsf{Hom}(U^\perp, U^0)$ is a canonical isomorphism. Here $\Phi \in \mathsf{Hom}(V, V')$. $\Phi_v(w) = \langle w, v\rangle$.

In other words, the isomorphism Φ from V to V' is also an isomorphism between $U^\perp$ and U^0.

Definition. Consider $U \subseteq V$, $\dim U < \infty$. $P_U \in \mathcal{L}(V)$. $P_U(v) = U$, $v = u + w$, with $u \in U$ and $w \in U^\perp$.

Theorem. $U \subseteq V$, $\dim U < \infty$. Then for all $v \in V$, $u \in U$,

$$||v - P_U(v)|| \leq ||v - u||$$

with equality holding if and only if $u = P_U(v)$.

Proof.

$$\begin{aligned} ||v - P_U(v)||^2 &\leq ||v - P_U(v)||^2 + ||P_u(v) - u||^2 \\ &= ||v - u||^2 \end{aligned}$$

■

Example. $\mathcal{Q}([0, 2\pi]) = V$, $U = \text{span}(\sin x, \cos x, \sin 2x, \cos 2x)$. Inner product defined:

$$\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x)g(x)\ dx$$

Given $v(x) = x$, find an element in U which minimizes the distance $||v - u||$ for all $u \in U$.

$$\frac{1}{\pi} \int_0^{2\pi} (\sin x)^2\ dx = \frac{1}{\pi} \int_0^{2\pi} \frac{1 - \cos 2x}{2}\ dx = 1$$

$$\frac{1}{\pi} \int_0^{2\pi} (\cos x)^2\ dx = 1$$

$$||\sin 2x|| = 1 = ||\cos 2x||$$

$$\int_0^{2\pi} \sin(mx)\sin(nx)\ dx = 0 \int_0^{2\pi} \sin(mx)\cos(nx)\ dx = \int_0^{2\pi} \cos(mx)\cos(nx)\ dx$$

$$P_U(v) = a_1 \sin x + b_1 \cos x + a_2 \sin 2x + b_2 \cos 2x$$

Note that

$$\int_0^{2\pi} x\cos(mx)\ dx = 0$$

$$\int_0^{2\pi} x\sin(mx)\ dx = -\frac{2\pi}{m}$$

with $\langle v, \sin mx \rangle = -\dfrac{2}{m}$.

19.1 Adjoint Operators

Definition (Adjoint). Let $T \in \text{Hom}(V, W)$, $T^* : W \to V$ is called the **adjoint** of T if $\langle Tv, w \rangle = \langle v, T^* w \rangle$ for all $v \in V, w \in W$.

T is self adjoint if $T = T^*$.

Lemma. $T \in \text{Hom}(V, W), S \in \text{Hom}(W, U)$, $\lambda \in \mathbb{F}$. Then the following holds:

(1) $T^* \in \text{Hom}(W, V)$

(2) $(S + T)^* = S^* + T^*$

(3) $(\lambda T)^* = \lambda T^*$

(4) $(T^*)^* = T$

(5) $I^* = I$

(6) $(ST)^* = T^* S^*$ (here let $T \in \text{Hom}(U, V)$ and $S \in \text{Hom}(V, W)$)

Lemma. $T \in (V, W)$.

(1) Null $T^* = (\text{Range } T)^\perp$

(2) Range $T^* = (\text{Null } T)^\perp$

(3) Null $T = (\text{Range } T^*)^\perp$

(4) Range $T = (\text{Null } T^*)^\perp$

Definition. $A = (a_{ij}) \in \mathbb{F}^{m,n}$. Its conjugate transpose $B = (b_{ij}) \in \mathbb{F}^{n,m}$ is defined by $b_{ij} = a_{ji}$.

Theorem. $T \in \text{Hom}(V, W)$, $B_V = \{e_1, \ldots, e_m\}$, $B_W = \{f_1, \ldots, f_n\}$ orthonormal. Then $\mathcal{M}(T^*, B_W, B_V)$ is the conjugate transpose of $\mathcal{M}(T, B_V, B_W)$.

Proof.

$$Te_j = \sum_{i=1}^{n} \langle Te_j, f_i \rangle f_i$$

$$a_{ij} = \langle Te_j, f_i \rangle$$

$$b_{ij} = \langle T^* f_j, e_i \rangle = \langle f_j, Te_i \rangle = \langle T\bar{e_i}, f_j \rangle = a_{\bar{j}i}$$

∎

20 Lecture 16: Spectral Theorems

Definition (Self-Adjoint). $T \in \mathcal{L}(V)$ is called **self-adjoint** if $T = T^*$.

Lemma. $T \in \mathcal{L}(V)$ is self-adjoint, then each eigenvalue of T is real.

Proof.
Suppose $v \neq 0$ is an eigenvector such that $Tv = \lambda v$.

$$\lambda\langle v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle.$$

Since $\langle v, v\rangle \neq 0$, $\lambda = \bar{\lambda}$.
Consider $T \in \mathcal{L}(V)$. $T = 0 \Leftrightarrow \langle Tv, w\rangle = 0$ for all $v, w \in V$.
Facts (Polarization).

(1) $\mathbb{F} = \mathbb{R}$, T self-adjoint, then

$$\langle Tv, w\rangle = \frac{\langle T(v+w), v+w)\rangle - \langle T(v-w), v-w\rangle}{4}$$

Proof.

$$\begin{aligned}\langle T(v+w), v+w\rangle &= \langle Tv, v\rangle + \langle Tw, w\rangle + \langle Tv, w\rangle + \langle Tw, v\rangle \\ \langle T(v-w), v-w\rangle &= \langle Tv, v\rangle + \langle Tw, w\rangle - \langle Tw, v\rangle - \langle Tv, w\rangle\end{aligned}$$

difference $= 4\langle Tv, w\rangle$

Remark. Substituting $T = I$ to the above lemma, we get

$$\langle v, w\rangle = \frac{||v+w||^2 - ||v-w||^2}{4}$$

(2) $\mathbb{F} = \mathbb{C}$, $T \in \mathcal{L}(V)$

$$\langle Tv, w\rangle = \frac{\langle T(v+w), v+w\rangle - \langle T(v-w), v-w\rangle}{4} + i\frac{\langle T(v+iw), v+iw\rangle - \langle T(v-iw), v-iw\rangle}{4}$$

Lemma. V : inner product space over $\mathbb{F}$, and $T \in \mathcal{L}(V)$.

(1) Suppose $\mathbb{F} = \mathbb{R}$, T is self-adjoint. Then $\langle Tv, v\rangle = 0 \forall v \in V \Rightarrow T = 0$.

(2) Suppose $\mathbb{F} = \mathbb{C}$, $T \in \mathcal{L}(V)$. Then $\langle Tv, v\rangle = 0 \forall v \in V \Rightarrow T = 0$.

Corollary. Suppose V is an inner product space over $\mathbb{C}$, and $T \in \mathcal{L}(V)$. Then

$$T \text{ self-adjoint } \Leftrightarrow \langle Tv, v\rangle \in \mathbb{R}$$

Proof.

$$\begin{aligned}\langle Tv, v\rangle - \overline{\langle Tv, v\rangle} &= \langle Tv, v\rangle - \langle v, Tv\rangle \\ &= \langle Tv, v\rangle - \langle T^*v, v\rangle \\ &= \langle (T - T^*)v, v\rangle\end{aligned}$$

$$\begin{aligned}\langle Tv, v\rangle \in \mathbb{R} &\Leftrightarrow \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \\ &\Leftrightarrow \langle (T - T^*)v, v\rangle = 0 \; \forall v \in V \\ &\Leftrightarrow T = T^*\end{aligned}$$

■

Definition (Normal). $T \in \mathcal{L}(V)$ is **Normal** if $TT^* = T^*T$.

Lemma. $T \in \mathcal{L}(V)$ is normal. Then $||Tv|| = ||T^*v||$ for all $v \in V$.

Proof.

$$\langle Tv, Tv\rangle = \langle v, T^*Tv\rangle = \langle v, TT^*v\rangle = \langle T^*v, T^*v\rangle.$$

■

Lemma. Suppose $T \in \mathcal{L}(V)$ is normal. Let $v \neq 0$ satisfy $Tv = \lambda v$ for some $\lambda \in \mathbb{F}$. Then $T^*v = \bar{\lambda}v$.

Proof.

$$||(T - \lambda I)v|| = ||(T^* - \bar{\lambda}I)v||$$

v is in an eigenvector of T with respect to λ is equivalent to v being an eigenvector of T^* with respect to $\bar{\lambda}$. ■

Lemma. Suppose $T \in \mathcal{L}(V)$ is normal. Take $\lambda \neq \mu$ different eigenvalues, and let v and w be their corresponding eigenvectors.

$$\begin{cases} Tv = \lambda v, \\ Tw = \mu w \end{cases} \Rightarrow v \perp w$$

Proof.

$$\begin{aligned} &\lambda\langle v, w\rangle - \mu\langle v, w\rangle \\ &= \langle \lambda v, w\rangle - \langle v, \bar{\mu}w\rangle \\ &= \langle Tv, w\rangle - \langle v, T^*, w\rangle \\ &= 0 \end{aligned}$$

Since $\lambda \neq \mu$, we can multiply both sides by $(\lambda - \mu)^{-1}$, so $\langle v, w\rangle = 0$, and v and w are perpendicular as desired. ■

We can now get into Spectral Theorems. Take $T \in \mathcal{L}(V)$. $\mathcal{S} = \{\lambda_1, \ldots, \lambda_m\rangle$ with $\lambda_1, \ldots, \lambda_m$ eigenvalues is the spectrum of T.

Remark. The spectrum of T is like the soul or ghost of T.

Theorem (Complex Spectrum). Consider $T \in \mathcal{L}(v)$, $\mathbb{F} = \mathbb{C}$. Then the following are equivalent:

(1) T is normal

(2) V has an orthonormal basis B consisting of eigenvectors of T

(3) V has an orthonormal basis B such that $\mathcal{M}(T, B)$ is diagonal

We now consider the matrix version.

Definition (Matrix Similarity). We say $T \sim S$ (similar) if $\exists$ an invertible matrix P such that $T = P^{-1}SP$.

Then we are saying that there exists a matrix similar to the matrix of T that is diagonal.

We now complete the proof of this theorem.

Proof.

By Schur's Theorem, there exists an orthonormal basis B such that

$$\mathcal{M}(T,B) = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22} & \dots & \dots & a_{2n} \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & a_{nn} \end{bmatrix}$$

The jth column of the above matrix corresponds to Te_j.

$$||T^*e_1||^2 = ||Te_1||^2, \sum_{j=1}^{n} |a_{1j}|^2 = |a_{11}|^2$$

$$|a_{11}|^2 + |a_{12}|^2 + \ldots + |a_{1n}|^2 = |a_{11}|^2$$

$$||T^*e_2||^2 = ||Te_2||^2$$

$$||T^*e_2||^2 = ||Te_2||^2, \sum_{j=2}^{n} |a_{2j}|^2 = |a_{22}|^2$$

$$|a_{23}|^2 + \ldots + |a_{2n}|^2 = 0$$

so the matrix is diagonal.

Theorem (Real Spectrum). Consider $T \in \mathcal{L}(V)$, $\mathbb{F} = \mathbb{R}$. Then the following are equivalent:

(1) T is self-adjoint

(2) V has an orthonormal basis B consisting of eigenvectors of T

(3) V has an orthonormal basis B such that $\mathcal{M}(T,B)$ is diagonal

Lemma. $T \in \mathcal{L}(V)$ self-adjoint, $b, c \in \mathbb{R}$ such that $b^2 - 4c < 0$. Then $T^2 + bT + cI$ is invertible.

Proof.

$$\begin{aligned} &\langle (T^2 + bT + cI\rangle v, v\rangle \\ &= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\ &\geq ||Tv||^2 - b||Tv||||v|| + c||v||^2 \\ &= ||v||^2 \left(\left(\frac{||Tv||}{||v||} \right)^2 - b \left(\frac{||Tv||}{||v||} \right) + c \right) \end{aligned}$$

Note that $x^2 - bx + c > 0$ for all $x \in \mathbb{R}$ if $b^2 - 4c < 0$. Thus, the above expression is greater than 0. Thus, if the above is zero, then $v = 0$, so $T^2 + bT + cI$ is injective and thus invertible as desired. ■

Proposition. Consider $T \in \mathcal{L}(V)$ self-adjoint. T has an eigenvalue.

Proof.

Considering $v \neq 0$, $v, Tv, T^2v, \ldots, T^nv$ is linearly dependent. Thus, there exist not all zero numbers a_j such that

$$\sum_{j=0}^{n} a_j T^j v = 0$$

In other words, there exists some polynomial p such that

$$p(T)v = 0, p(T) = \sum_{j=0}^{n} a_j T^j$$

Recall that for any real polynomia,

$$p(x) = c\Pi_{j=1}^{k}(x^2 - b_j x + c_j)\Pi_{i=1}^{m}(x - \lambda_i)$$

with $b_j^2 - 4c_j < 0$ for all j. We also have $k + m \geq 1$. Then all the quadratic factors are invertible, so one of the linear factors must not be invertible, giving us an eigenvalue. ■

Lemma. Consider $T \in \mathcal{L}(V)$ a self-adjoint transformation, and consider an invarient subspace U. Then

(1) $U^\perp$ is invariant under T

(2) $T|_U$ is self-adjoint

(3) $T|_{U^\perp}$ is self-adjoint

Proof.

(1) $w \in U^\perp, u \in U$, $\langle Tw, u\rangle = \langle w, Tu\rangle = 0$

(1) $\Rightarrow$ (2) $Te_1 = \lambda_1 e_1$, $U_1 = \text{span}\{e_1\}$, $T|_{U_1^\perp}$ is self-adjoint. This gives us another eigenvector e_2. We continue this process.

21 Lecture 18: Polar/Singular Value Decomposition

Theorem (Polar Decomposition). Suppose $T \in \mathcal{L}(V)$. Then $\exists$ an isometry such that $T = S \cdot \sqrt{T^*T}$.

Remark. Since T^*T is self-adjoint, it is diagonalizable, so it is a dialation. The above theorem says that we can break any transformation T into an isometry S and a dialation $\sqrt{T^*T}$.

Theorem (Singular Value Decomposition). Given $T \in \mathcal{L}(V)$ with singular values $s_1, \ldots, s_n$, $\exists$ orthonormal bases $B_1 = \{e_1, \ldots, e_n\}$ and $B_2 = \{f_1, \ldots, f_n\}$ such that $Tv = \sum_{j=1}^{n} s_j \langle , e_j \rangle f_j$.

$$\hat{T} = \mathcal{M}(T, B_1, B_2) = \begin{bmatrix} s_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & s_n \end{bmatrix}$$

Theorem (Spectral Theorem). For any self-adjoint operator $T \in \mathbb{R}^{m,n}$, there exists an orthogonal matrix $P \in \mathbb{R}^{n,n}$ such that $T = PDP^{-1}$ and

$$D = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}$$

Remark. This is basically saying there is a change of base P from T to D.

Definition (Unitary, Orthogonal). $V = \mathbb{C}^n$, P is called a unitary matrix if $PP^* = P^*P = I$. $V \in \mathbb{R}^n$, P is called an orthogonal matrix if $PP^T = P^TP = I$.

Theorem (Matrix Version of the Singular Value Decomposition). If $(e_1, \ldots, e_n) = (v_1, \ldots, v_n)V$ and $(f_1, \ldots, f_n) = (v_1 \ldots, v_n)U$ then $\hat{T} = U\Sigma V^{-1}$.

Theorem. $T \in \mathcal{L}(V)$. $M_1 = \mathcal{M}(T, B_1)$, $M_2 = \mathcal{M}(T, B_2)$. $B_1 = \{e_1, \ldots, e_n\}$, $B_2 = \{f_1, \ldots, f_n\}$. $B_0 = \{v_1, \ldots, v_n\}$.

$$(e_1, \ldots, e_n) = (v_1, \ldots, v_n)P, \quad (f_1, \ldots, f_n) = (v_1, \ldots, v_n)Q$$

Then

$$M_2 = Q^{-1}M_1P$$

Proof.
We define $Q = \sqrt{T^*T}$ with $Qe_j = s_je_j$. By applying the polar decomposition, there exists an isometry S such that $T = SQ$. We define $f_j = Se_j$.

$$Te_i = SQe_i = S(s_ie_i) = s_iS(e_i) = s_if_i$$

as desired. ∎

Lemma 1. Given $T \in \mathcal{L}(V)$, then $||Tv|| = ||\sqrt{T^*T}v||$.

Proof.

$$\langle \sqrt{T^*T}v, \sqrt{T^*T}v \rangle = \langle v, T^*Tv \rangle = \langle Tv, Tv \rangle$$

■

We have shown that there is an isometry from the range of $\sqrt{T^*T}$ to T given by $\hat{S}(\sqrt{T^*T}v) \stackrel{\text{def}}{=} Tv$. Note that since

$$||\hat{S}(\sqrt{T^*T}v) - \hat{S}(\sqrt{T^*T}w)|| = ||\sqrt{T^*T}v - \sqrt{T^*T}w|| = 0$$

By definition, $\hat{S}$ is also surjective.

Lemma 2. Suppose we have two inner product spaces with $\dim V = \dim W$. Let $V_1 \subset V$, $W_1 \subset W$ with $\dim V_1 = \dim W_1$. Then any isometry $\hat{T}$ from V_1 to W_1 extends to an isometry T from V to W.

Proof.

Consider $B_1 = \{e_1, \ldots, e_m\}$ an orthonormal basis of $V_1^\perp$. Consider $B_2 = \{f_1, \ldots, f_m\}$ an orthonormal basis of $W_1^\perp$. Define $\hat{Q}$ mapping from $V_1^\perp$ to $W_1^\perp$ by $\hat{Q}(e_j) = f_j$. $V = V_1 \oplus V_1^\perp$, $W = W_1 \oplus W_1^\perp$.

$$Tv = T(u + w) = \hat{T}u + \hat{Q}w$$

where $u \in V_1$ and $w \in V_1^\perp$. Then $T : V \to W$ is an isometry.

We now prove the Polar Decomposition Theorem.

Proof.

$$\dim(\text{Range } \sqrt{T^*T}) = \dim(\text{Raneg } T) + \dim(\text{Null } \hat{S})$$

Since $\hat{S}$ is injective (since it is an isometry), $\dim(\text{Range } \sqrt{T^*T}) = \dim(\text{Range } T)$, we can apply Theorem 2 so $\hat{S}$ extends to an isometry S from V to V as desired. ■

22 Linear Operators on Complex Vector Spaces

A blocked diagonal matrix is a matrix that is diagonal but the things on the diagonals are square matrices of the same dimension. Example:

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

We would like to represent any transfrmation T as

$$T = P^{-1}SP$$

where S is a blocked diagonal matrix.

Lemma. $T \in \mathcal{L}(V)$. Then

$$\{0\} = \text{Null } T^0 \subseteq \text{Null } T^1 \subseteq \text{Null } T^2 \subseteq \dots$$

Proof.
We show this via induction. For the base case, Null $T^0 = \{0\} \subseteq$ Null T^1, and for all k, if $T^k(v) = 0$, then $T^{k+1}(v) = T(T^k(v)) = T(0) = 0$ so Null $T^k \subseteq$ Null T^{k+1} so the result follows via induction as desired. ■

Proposition. Consider $T \in \mathcal{L}(V)$. If Null T^m = Null T^{m+1}, then for any $k \in \mathbb{Z}^+$, Null T^m = Null T^{m+k}.

Proof.
We only prove that Null $T^{n+k} \subseteq$ Null T^m. Let $v \in V$ such that $T^{m+k}(v)$.

$$0 = T^{m+k}(v) = T^{m+1}(T^{k-1}(v)) \Rightarrow T^{k-1}(v) \in \text{Null } T^{m+1} \subseteq \text{Null } T^m$$

Thus $T^m(T^{k-1}(v)) = 0 = T^{m+k-1}(v) = 0$. By induction, $T^{m+0}(v) = 0$, so Null $T^{m+k} \subseteq$ Null T^m as desired. ■

Proposition. Let $T \in \mathcal{L}(V)$ with $\dim V = n$. Then Null T^n = Null T^{n+1} = Null $T^{n+2} \dots$.

Proof.
We do a proof by contradiction. Suppose Null $T^n \subset$ Null T^{n+1}. By the previous propositions proof,

$$\{0\} = \text{Null } T^0 \subset \text{Null } T^1 \subset \dots \subset \text{Null } T^n \subset \text{Null } T^{n+1}$$

However, this means that $n \leq \dim \text{Null } T^n < \dim \text{Null } T^{n+1}$ a contradiction. Thus Null $T^n \supseteq$ Null T^{n+1} and we are done by the proposition. ■

Corollary. Consider $T \in \mathcal{L}(V)$ with $\dim V = n$. Then

$$V = \text{Null } T^n \oplus \text{Range } T^n$$

Remark. Do we know when a transformation T has an nth root for $n \in \mathbb{N}$?

Proof.
WE only need to check Null $T^n \cap$ Range $T^n = \{0\}$. Let $u \in$ Null $T^n \cap$ Range $T^n \Rightarrow T^n u = 0$, $u = T^n v$. Then $T^{2n} v = 0$ so $v \in$ Null T^{2n}. However, by the previous proposition, Null T^{2n} = Null T^n,

so $T^n v = 0$. Thus $u = 0$ as desired. ■

Definition (Generalized Eigenvector). Consider $T \in \mathcal{L}(V)$ with an eigenvalue λ. Given any $j \in \mathbb{Z}^+$, a solution $v \neq 0$ of $(T - \lambda I)^j v = 0$ is called **generalized eigenvector**.

Given any arbitrary eigenvalue λ, the set of all generalized eigenvectors of λ and $0 \in V$ is called the **generalized eigenspace** with notation $G(\lambda, T)$.

Lemma. Consider $T \, in \mathcal{L}(V)$ with λ an eigenvalue. Then

$$G(\lambda, T) = \text{Null}\left((T - \lambda I)^{\dim V}\right)$$

Proof.
Taking any $v \in G(\lambda, T)$. Then $\exists j$ such that $v \in \text{Null}\left((T - \lambda I)^j\right)$.
Case 1. $j < \dim V$
Then

$$\text{Null}\left((T-\lambda I)^j\right) \subseteq \text{Null}\left((T-\lambda I)^{j+1}\right) \subseteq \ldots \subseteq \text{Null}\left((T-\lambda I)^{\dim V}\right) = \text{Null}\left((T-\lambda I)^{\dim V + 1})\right) = \ldots$$

Case 2. $j \geq \dim V$
Then $\text{Null}\left((T - \lambda I)\right)^j = \text{Null}\left((T - \lambda I)^{\dim V}\right)$.
In both cases, $\text{Null}\left((T - \lambda I)^j\right) \subseteq \text{Null}\left((T - \lambda I)^{\dim V}\right)$ as desired. ■

Proposition. Consider $T \in \mathcal{L}(V)$. Consider $\lambda_1, \ldots, \lambda_n$ distinct eigenvalues with corresponding $v_1, \ldots, v_n$ generalized eigenvectors. Then $v_1, \ldots, v_m$ are linearly independent.

Proof.

$$0 = a_1 v_1 + \ldots + a_n v_n$$

Let k be the largest integer (which exists from the previous theorems) such that $(T - \lambda I)^k v_1 \neq 0$. Then $(T - \lambda_1 I)^{k+1} v_1 = 0$. We denote $w_1 = (T - \lambda I)^k v_1$.

$$(T - \lambda_1 I) w_1 = 0$$

so $T w_1 = \lambda_1 w_1$. Let $(T - \lambda_1 I)^k \prod_{j=2}^n (T - \lambda_j I)^n$ act on $0 = a_1 v_1 + \ldots + a_n v_n$.

$$a_1 \prod_{j=2}^{n} (T - \lambda_j I)^n w_1 + 0 + \ldots + 0 = 0$$

so $a_1 = 0$. Similarly for the other coefficients. ■

Corollary. Consider $\lambda_1, \ldots, \lambda_n$ distinct eigenvalues. Then $\sum_{j=1}^n G(\lambda_j, T)$ is a direct sum.

Definition (Nilpotent). $T \in \mathcal{L}(V)$ is **nilpotent** if $T^m = 0$ for some $m \in \mathbb{Z}^+$.

Remark. Nilpotent is Prof. Ruobing's favorite word.

Lemma. If $N \in \mathcal{L}(V)$ is nilpotent, then $N^{\dim V} = 0$.

Proof follows pretty quickly from previous lemmas.

Proposition. LEt $N \in \mathcal{L}(V)$ be nilpotent. Then V has a basis B such that $\mathcal{M}(N, B) = (a_{ij})$ satisfies $a_{ij} = 0$ when $i \geq j$.

Remark. The lower triangular part is the zero part.

Proof.
$B_1 = \{v_{11}, \ldots, v_{1k}\}$ basis of Null N. Extend B_1 to B_2 which is a basis of Null N^2. After finite steps, we obtain a basis B of V. Example: suppose $v_1, v_2 \in B_1$, $v_1, v_2, v_3 \in B_2$, $v_1, v_2, v_3, v_4 \in B_3$. Notice that $N(v_1) = N(v_2) = 0$. $N(v_3) \in$ Null N, so v_3 can be generated by v_1, v_2. $N(v_4) \in$ Null N^2, so v_4 can be generated by v_1, v_2, v_3. etc.

Lemma. Consider $T \in \mathcal{L}(V)$, $p \in \mathcal{P}(\mathbb{F})$. Then $\ker p(T), \operatorname{Im} p(T)$ are invariant under T.

Remark. I'm switching to $\ker$ and Im because it's faster.

Proof.
Consider $v \in \ker p(T)$, so $p(T)v = 0$. Then (since $p_1(T)p_2(T) = p_2(T)p_1(T)$)

$$Tp(T)v = p(T)(Tv) = 0$$

■

Proposition. Consider $T \in \mathcal{L}(V)$, and $\lambda_1, \ldots, \lambda_m$ is a spectrum of T.

(1) $V = \bigoplus_{j=1}^{m} G(\lambda_j, T)$

(2) Each $G(\lambda_j, T)$ is invariant under T

(3) Every $(T - \lambda_j I)|_{G(\lambda_j,T)}$ is nilpotent.

Proof.

(1)

We will prove $G(\lambda_j, T_U) = G(\lambda_j, T)$.
We start by showing that $G(\lambda_j, T_U) \geq G(\lambda_j, T)$. Taking any $v \in G(\lambda_j, T) \subseteq V$,

$$v = v_1 + v_2 + \ldots + v_m$$

where $v_i \in G(\lambda_i, T|_U)$. Generalized eigenvactors are linearly independent, so $v_1 = 0$, $v_i = 0$ if $i \neq j$ so We induce on $n = \dim V$. For $n = 1$, it is trivial. Assume the result holds for any V with $\dim V < n$. For λ_1,

$$\begin{aligned} V &= \ker(T - \lambda_1 I)^n \oplus \operatorname{Im}(T - \lambda_1 I)^n \\ &= G(\lambda_1, T) \oplus U \end{aligned}$$

Then $\dim U < n$, so U is invariant under T by the inductive hypothesis.

$$V = G(\lambda_1, I) \oplus (\bigoplus_{j=2}^{n} G(\lambda_j T|_U))$$

$v = v_j$.

(2)

$G(\lambda_j, T) = \text{Null}(T - \lambda_j I)^{\dim V}$ is invariant under T.

(3)

By definition.

23 Lecture : Blocked Diagonal Decomposition

Proposition. Consider V a complex vector space, and $T \in \mathcal{L}(V)$ with spectrum $\{\lambda_1, \ldots, \lambda_m\}$. Then

(1) $V = G(\lambda_1, T) \oplus G(\lambda_2, T) \oplus \ldots \oplus G(\lambda_m, T)$

(2) $G(\lambda_j, T)$ is invariant under T

(3) $(T - \lambda_j I)|_{G(\lambda_j, T)}$ is nilpotent

Definition (Multiplicity). Consider $T \in \mathcal{L}(V)$, with λ_i an eigenvalue of T. Then **algebraic multiplicity** of λ is $\dim G(\lambda, T) = \dim \text{Null}\ (T - \lambda I)^{\dim V}$. The **geometric multiplicity** of $\lambda = \dim E(\lambda, T) = \dim \text{Null}\ (T - \lambda I)$.

Corollary. Consider V with $\dim V = n$. Then the sum of the algebraic multiplicities is equal to n.

Theorem. Consider V a complex space, with $T \in \mathcal{L}(V)$ with spectrum $\lambda_1, \ldots, \lambda_m$. Let the multiplicity of λ_j be d_j. Then V has a basis consisting of generalized eigenvectors such that

$$\mathcal{M}(T, B) = \begin{bmatrix} A_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_m \end{bmatrix}$$

where

$$A = \begin{bmatrix} \lambda_j & & & \\ 0 & \lambda_j & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \lambda_j \end{bmatrix}$$

is a d_j by d_j matrix.

Proof.
To understand teh structure of A_j, Consider $T|_{G(\lambda_j, T)}$, the result when T is restricted to $G(\lambda_j, T)$.

$$T|_{G(\lambda_j, T)} = (T - \lambda_j)|_{G(\lambda_j, T)} + \lambda_j I|_{G(\lambda_j, T)}$$

By the Proposition, $(T - \lambda_j)|_{G(\lambda_j, T)}$ is nilpotent, which immediately gives the result (by the proposition from the last lecture). ∎

Example. Consider

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

with eigenvalues 1, 2.

Remark. Can't we use column reduction? Never mind, no mayyybe not.

$G(1, A) = \text{span}((1, 0, 0), (0, 1, 0))$ $G(2, A) = \text{span}((5, 1, 1))$

$$V = G(1, A) \oplus G(2, A)$$

$$\mathcal{M}(A, B) = P^{-1}AP = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Lemma. If $N \in \mathcal{L}(V)$ is nilpotent, then $I + N$ has a square root.

Motivation: Recall the taylor expansion:

$$\sqrt{1+x} = 1 + \frac{x}{2} + O(x^2)$$

Proof. Suppose $N^m = 0$. We write

$$R = I + a_1 N + a_2 N^2 + \ldots + a_{m-1} N^{m-1}$$

Then

$$R^2 = I + 2a_1 N + (a_1^2 + 2a_2)N^2 + (2a_3 + 2a_1 a_2)N^3 + \ldots + (2a_{m-1} + p(a_1, \ldots, a_{m-2}))$$

We can now inductively solve this system so that $a_1 = \frac{1}{2}$, and each coefficient of further powers of N is zero in the above. ■

Proposition. Consider complex vector space V, and $T \in \mathcal{L}(V)$ invertible. Then T has a square root.

Proof.

$$V = G(\lambda_1, T) \oplus \ldots \oplus G(\lambda_m, T)$$

where $\lambda_j \neq 0$ for all j. On each $G(\lambda_j, T)$,

$$\begin{aligned} T|_{G(\lambda_j,T)} &= (T - \lambda_j I)|_{G(\lambda_j,t)} + \lambda_j I|_{G(\lambda_j,T)} \\ &= \lambda_j(\lambda_j^{-1}(T - \lambda_j I)|_{G(\lambda_j,T)} + I) \end{aligned}$$

(λ_j^{-1} exists since T is invertible) where the expression in the parenthesis is nilpotent. By the previous lemma, $R_j =$ square root of $T|_{G(\lambda_j,T)}$ exists. We define for $v = v_1 + \ldots + v_m$ with $v_j \in G(\lambda_j, T)$ for each j,

$$Rv = R_1 v_1 + \ldots + R_m v_m$$

then R is a square root of T since

$$R^2 v = R_1^2 v_1 + \ldots + R_m^2 v_m = Tv$$

■

Definition (Characteristic Polynomial). Let V be a complex vector space with $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ a spectrum of T with distinct eigenvalues with multiplicities $d_1, \ldots, d_m$. Then polynomial $p(z) = (z - \lambda_1)^{d_1} \cdot \ldots \cdot (z - \lambda_m)^{d_m}$ is called the **characteristic polynomial**.

Remark. :heart_eyes: :heart_eyes: :heart_eyes:

Remark. If $\dim V = n$, then $\deg(p(z)) = n$.

Theorem (Cayley-Hamilton). Consider a complex vector space V with $T \in \mathcal{L}(V)$. If $q(z)$ is the characterisitc polynomial, then $q(T) = 0$.

Proof.

$$V = G(\lambda_1, T) \oplus \ldots \oplus G(\lambda_m, T)$$

For $v \in V$, $v = v_1 + \ldots + v_m$ with $v_j \in G(\lambda_j, T)$,

$$p(T)(v) = p(T)(v_1) + \ldots + p(T)(v_m)$$

But $p(T)v_j = (\ldots)(T - \lambda_j I)^{d_j} v_j = 0$, which gives us the desired result. ■

Lemma. Consider $T \in \mathcal{L}(V)$ with $\dim V = n$. Then there exists a unique monic polynomial p of smallest degree such that $p(T) = 0$.

Proof.
Consider $\{I, T, T^2, \ldots, T^{n^2}\}$. Since $\dim \mathcal{L}(V) = n^2 < n^2 + 1$, the list above is linearly dependent. Let m be the smallest integer such that $\{I, T, T^2, \ldots, T^m\}$ is linearly dependent. Then there exists coefficients not all zero such that

$$\sum_{j=0}^{m-1} a_j T^j + T^m = 0$$

We define $q(z) = \sum_{j=0}^{m-1} a_j z^j + z^m$. Suppose $\hat{q}(z) = \sum_{j=0}^{m-1} b_j z^j + z^m$. Then

$$(q - q^j)(z) = \sum_{j=0}^{m-1} (a_j - b_j) z^j$$

Then

$$\sum_{j=0}^{m-1} (a_j - b_j) T^j = 0$$

however this is a contradiction of m being the lowest degree. Thus, $q = \hat{q}$. ■

The monic polynomial from the above lemma is called the minimal polynomial.

Lemma. Consider $T \in \mathcal{L}(V)$ and $q \in \mathcal{P}(\mathbb{F})$. Then $q(T) = 0$ if and only if q is a polynomial multiple of the minimal polynomial p.

Proof.
The if part is trivial. Otherwise, let $q = r \cdot p + h$ where h has as smaller degree than p. Then $h(T) = 0$, but this is a contradiction. ■

Corollary. Consider $T \in \mathcal{L}(V)$ and p the characteristic polynomial, with q the minimal polynomial. Then $\exists r \in \mathcal{P}(\mathbb{F})$ such that $p = rq$.

Proposition. Let $T \in \mathcal{L}(V)$. Then the zeros of the minimal polynomial p are precisely the eigenvalues of T.

Proof.

(1)

Suppose λ is a zero. Then $p(z) = (z - \lambda)q(z)$, so

$$0 = p(T)v = (T - \lambda I)(q(T)v)$$

But $q(T)v$ must not be zero for some v since that would contradict minimality.

(2)

If λ is an eigenvalue, then

$$p(T)v = \sum_{j=0}^{m} (a_j T^j)v = \sum_{j=0}^{m} a_j (\lambda)^j v = p(\lambda)v$$

24 Precept : Jordan Form

Consider $\mathbb{F} = \mathbb{C}$ and $V = \mathbb{C}^3$.

Example. $T : V \to V$ is linear, $Tv = Av$ where $A = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 2 & -1 \\ -4 & 0 & 3 \end{bmatrix}$. We can write $V = G(\lambda_1, T) \oplus \ldots \oplus G(\lambda_m, T)$.
Notice that 2 is an eigenvalue corresponding to $(0,1,0)$. Observe that $E(2,T) = \ker(A - 2I) = \text{span}(e_2)$.
Now we would like to find other eigenvalues. We will follow 5.21 (and compare this to Problem A from PSet 10 after class).
Notice that with $e_1 = (1,0,0)$, $Te_1 = (-1,0,4)$, $T^2e_1 = (-3,4,-8)$, $T^3e_1 = (-5,16,-12)$. Then

$$0 = -2e_1 + 5Te_1 - 4T^2e_1 + T^3e_1$$

Define $p(x) = -2 + 5x - 4x^2 + x^3$. Then $p(T)e_1 = 0$. Note that $p(x) = (x-1)^2(x-2)$. We now check $\lambda = 1$:

$$\ker(A - I) = \ker \begin{bmatrix} -2 & 0 & 1 \\ 0 & 1 & -1 \\ -4 & 0 & 2 \end{bmatrix}$$

$$= \text{span}(\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix})$$

so 1 is an eigenvalue. (This is an alternative way of checking that 1 is an eigenvalue). We have $E(1,T) = \text{span}(\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix})$.
We now consider generalized eigenspaces. Notice that

$$G(1,T) \oplus G(2,T) \subseteq V \Rightarrow \dim G(1,T) + \dim G(2,T) \leq \dim V = 3$$

Thus each eigenspace has dimension between 1 and 2. In other words,

$$G(1,T) = \ker(A - I)^2$$

$$G(2,T) = \ker(A - 2I)^2$$

Notice that

$$\ker(A - I)^2 = \ker \begin{bmatrix} 0 & 0 & 0 \\ 4 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \text{span}(\begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix})$$

so $\dim G(1,T) = 2$ and $\dim G(2,T) = 1$. Notice that we didn't actually have to notice that $G(1,T) = \ker(A - I)^2$. Even if we had picked 2 arbitrarily, we would have found $\dim G(1,T) \geq 2 \Rightarrow \dim G(1,T) = 2$.
We now find the characteristic polynomial: $(x-1)^2(x-2)$. The minimal polynomial divides the characteristic polynomial. Also, the minimal polynomial and the characteristic polynomial have the same set of roots. Thus, the minimal polynomial is either $(x-1)(x-2)$ or $(x-1)^2(x-2)$.

Example (8.C.12). T being diagonal is equivalent to the minimal polynomial has no repeated roots. Notice that

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

has minimal polynomial $(x-1)^2$ but

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

has minimal polynomial $(x-1)$.

Example. Going back to the original example, the minimal polynomial is $(x-1)^2(x-2)$.

Example.

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

has minimal polynomial x^3

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

has minimal polynomial x^2

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

has minimal polynomial x.
Notice that the following two matrices have the same minimal and characteristic polynomials:

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

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

Example (8.C.15). Consider T, $v \neq 0 \in V$. Then

(1) There exists a smallest monic polynomial p such that $p(T)v = 0$

(2) p is a factor of the minimal polynomial

In the previous example, $(x-1)^2(x-2)$ is the minimal polynomial for T and e_1, so $(x-1)^2(x-2)$ is a factor of the minimal polynomial of T. Thus the characteristic polnomial is $q \cdot (x-1)^2(x-2)$ in the above polynomial, but they m

Theorem. Consider V a complex space, with $T \in \mathcal{L}(V)$ with spectrum $\lambda_1, \dots, \lambda_m$. Let the multiplicity of λ_j be d_j. Then V has a basis consisting of generalized eigenvectors such that

$$\mathcal{M}(T, B) = \begin{bmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_m \end{bmatrix}$$

where

$$A = \begin{bmatrix} \lambda_j & & & \\ 0 & \lambda_j & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \lambda_j \end{bmatrix}$$

is a d_j by d_j matrix.

A matrix of the form with a diagonal containing the eigenvalue and the diagonal above having all 1s is called a Jordan block. A basis as in the above theorem is called a Jordan basis.

Example.

$$M = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad M = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

are Jordan blocks. Notice that in the first case above,

$$\begin{aligned} Tv_1 - \lambda v_1 &= 0 \\ Tv_2 - \lambda v_2 &= v_1 \\ Tv_3 - \lambda v_3 &= v_2 \end{aligned}$$

so $\{v_3, (T - \lambda I)v_3, (T - \lambda I)^2 v_3\}$ is a basis, and notice that $(T - \lambda I)^3 v_3 = 0$.

Theorem. Any matrix $T \in \mathbb{C}^{n,n}$ is similar to its Jordan canonical form, i.e., $\exists$ an invertible matrix P such that

$$P^{-1}AP = \begin{bmatrix} A_1 & \dots 0 & \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_m \end{bmatrix}$$

where A_j is a Jordan block.

Proof (using the Proposition below).
Consider $V = G(\lambda_1, T) \oplus \dots \oplus G(\lambda_m, T)$. We denote

$$N_j = (T - \lambda_j I)|_{G(\lambda_j, T)}$$

nilpotent. Then applying the Proposition, each block with respect to $G(\lambda_j, T)$ is a Nilpotent block. Thus $A_j = N_j + \lambda_j I$ which gives us the desired result. ■

Proposition. $N \in \mathcal{L}(V)$ is nilpotent, then there exists $\{v_1, \dots, v_m\} \subseteq V$ and $\{k_1, \dots, k_m\} \subseteq \mathbb{N}_0$ such that

(1) $B = \{N^{k_1}v_1, \dots, N^2 v_1, Nv_1, v_1, N^{k_2}v_2, \dots, N^2 v_2, Nv_2, v_2, \dots, N^{k_m}v_m, \dots, N^2 v_m, Nv_m, v_m\}$ is a basis of V

(2) $N^{k_j+1}v_j = 0$ for all $1 \leq j \leq m$

Proof.

We show this by induction on $n = \dim V$. For the base case $n = 1$, it is trivial. Suppose the conclusion holds for any V with $\dim V < n$. Consider V with $\dim V = n$. Taking any nilpotent operator $N \in \mathcal{L}(V)$, with N not injective ($\Leftrightarrow N$ is not surjective). Thus dim(Range N) $\leq n-1$, so we can apply the inductive hypothesis on Range N. Notice that Range N is invariant under N. Denote $\hat{N} = N_{\text{Range } N}$. $\hat{N}$ is nilpotent. By the induction hypothesis, there exists a basis

$$\hat{B} = \{N^{k_1}v_1, N^{k_1-1}v_1, \ldots, Nv_1, v_1, N^{k_2}v_2, \ldots N^{k_l}v_l, N^{k_l-1}v_l, \ldots, Nv_l, v_l$$

of Range N. Notice that $v_j = Nu_j$ for some u_j since $v_j \in$ Range N. Then consider the basis:

$$\hat{\hat{B}} = \{N^{k_1+1}u_1, \ldots, Nu_1, u_1, \ldots, \ldots, N^{k_l+1}u_l, \ldots, Nu_l, u_l\}$$

Note that $N\hat{\hat{B}} = \hat{B}$ (by (2)). Now consider a_{ij}.

$$\sum_{j=1}^{l}\sum_{i=0}^{k_j+1} a_{ij}N^i u_j = 0$$

Apply N to both sides. Then

$$\sum_{j=1}^{l}\sum_{i=0}^{k_j+1} a_{ij}N^i v_j = 0$$

Then we have $a_{ij} = 0$ for all $1 \leq j \leq l$ and $0 \leq i \leq k_j$ by the inductive hypothesis since $N^{k_j+1}v_j = 0$ by (2). Thus

$$\sum_{j=1}^{l} a_{k_j+1,j}N^{k_j+1}u_j = 0$$

$$\sum_{j=1}^{l} a_{k_j+1,j}N^{k_j}v_j = 0$$

Thus $a_{k_j+1,j} = 0$ for all j. Thus $a_{ij} = 0$ fir all i, j, so $\hat{\hat{B}}$ is a linearly independent list. Then $\hat{\hat{B}}$ can be extended to a basis of V

$$\tilde{B} = \hat{\hat{B}} \cup \{w_1, \ldots, w_p\}$$

Observe that for all w_j, $Nw_j \in$ Range N, so there exists $x_j \in \text{span}(\hat{\hat{B}})$ such that $Nx_j = Nw_j$. We define $\zeta_j = w_j - x_j$. $N\zeta_j = 0$ for all j. Then

$$\tilde{\tilde{B}} = \hat{\hat{N}} \cup \{\zeta_1, \ldots, \zeta_p\}$$

is also a basis, and it satisfies (1) and (2) as desired. By induction, the result is true. ■

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

We can find that $\dim E(1, A) = 1$, $\dim G(1, A) = 2$, $\dim E(3, A) = \dim G(3, A) = 1$. Then

$$J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

and we can find P such that $P^{-1}AP$ by working out that if the basis is v_1, w, v_2, then $Aw = w + v_1$.

24.1 Trace

Definition (Trace). Consider $A = (a_{ij}) \in \mathbb{F}^{n,n}$. Then $\text{Tr}(A) = \sum_{i=1} 6na_{ii}$.
For a complex vector space this is the sum of the eigenvalues counted with multiplicities.

Lemma. Consider $A, B \in \mathbb{F}^{n,n}$. Then $\text{Tr}(AB) = \text{Tr}(BA)$.

Proof.

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

$$(BA)_{ij} = \sum_{k=1}^{n} b_{ij}a_{kj}$$

Then it follows quickly. ∎

Proposition. Consider $T \in \mathcal{L}(V)$. Consider two bases A, B. Then $\text{Tr}(\mathcal{M}(T, A)) = \text{Tr}(\mathcal{M}(T, B))$.

Proof.
Notice that $\text{Tr}\left(\mathcal{M}(T, B)\right) = \text{Tr}\left(P^{-1}\mathcal{M}(T, A)P\right) = \text{Tr}\left(\mathcal{M}(T, A)PP^{-1}\right) = \text{Tr}\left(\mathcal{M}(T, A)\right)$. ∎

Theorem. Consider $T \in \mathcal{L}(V)$. Then $\text{Tr}\left(T\right) = \text{Tr}\left(\mathcal{M}(T, B)\right)$ for all bases B of V.

25 Lecture : Tensor Product of Vector Spaces

Definition (k-linear). A map $\varphi : V_1 \times \ldots \times V_k \to V$ is called k -linear if $\varphi(v_1, \ldots, v_i, \ldots, v_k)$ is linear for eery v_i as fixing all variables but v_i to be constant.

> **Example.** $\varphi(u, v)$ bilinear. On real vector spaces, inner product spaces are bilinear maps.

Definition (Tensor Product). For all vector spaces U and for all bilinear maps $f : (V \times W) \to U$, then there exists a linear map

$$\overline{f} : V \otimes W \to U$$

such that $\overline{f}(v \otimes w) = f(v, w)$. The diagram formed by $V \times W$, $V \otimes W$, and U with $\otimes$, f, and $\overline{f}$ commutes.

> **Remark.** In category theory, we have a unique element with this universal property such that for all other elements, we can push forward the special element to that. The above definition satisfies that.

The first most natural question is: dess $V \otimes W$ exist?

> **Lemma.** If tensor products exist, then they are unique up to an isomorphism.

Proof.
Suppose there exist two tensor products $V \otimes W, V \tilde{\otimes} W$. By the universal property, there exists a unique linear map $\varphi : V \otimes W \to V \tilde{\otimes} W$ such that

$$\tilde{\otimes} = \varphi \circ \otimes$$

Similarly, there exists a unique linear map $\psi : V \tilde{\otimes} W \to V \otimes W$ such that

$$\otimes = \psi \circ \tilde{\otimes}$$

Then $\otimes = \psi \circ \phi \circ \otimes$. Now,
Thus $\psi \circ \phi = I$, so both are isomorphisms as desired.

26 Precept: Missed

I unfortunately missed this precept (hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhugh I wish I could've gone) and learned the material from the notes posted afterwards.

Therefore, this memory will be absent from my notes . . . ;(

27 Lecture : Wedge Product, Determinant

Proposition. Let $\{e_1, \ldots, e_n\}$ be a basis of V. Then

(1) $\{e_{i_1} \otimes \ldots \otimes e_{i_k} \mid 1 \leq i_1, \ldots, e_k \leq n\}$ basis of $V^{\otimes k}$

(2) $\{e_{i_1} \wedge \ldots \wedge e_{i_k} \mid 1 \leq i_1 <, \ldots, < e_k \leq n\}$ basis of $V^{\otimes k}$

Definition (Tensor Product of Linear Functionals). $\varphi, \psi \in \mathcal{L}(V, \mathbb{R})$, then $\varphi \otimes \psi \in \mathcal{M}_2(V, \mathbb{R})$ with

$$\varphi \otimes \psi(v, w) \stackrel{\text{def}}{=} \varphi(v) \cdot \psi(w)$$

Definition (Wedge Product of Linear Functionals). $\varphi, \psi \in \mathcal{A}_2(V, \mathbb{F})$, then

$$\varphi \wedge \psi(v, w) \stackrel{\text{def}}{=} \varphi(v) \cdot \psi(w) - \varphi(w)\psi(v)$$

Proof.

(1)

For $k = 2$, if $T \in \mathcal{M}_2(V, \mathbb{F})$, then

$$v = \sum_{i=1}^{n} a_i e_i = \sum \varphi_i(v) e_i$$

$$w = \sum_{i=1}^{n} b_i e_i = \sum \psi_i(w) e_i$$

$$T(v, w) = \sum_{i,j=1}^{n} T(\varphi_i(v)e_i, \varphi_j(w)e_j)$$

$$= \sum_{i,j=1}^{n} T(e_i, e_j)\varphi_i(v)\varphi_j(w)$$

$$= \sum_{i,j=1}^{n} T(e_i, e_j)\varphi_i \otimes \varphi_j(v, w)$$

$$T = \sum T(e_i, e_j)\varphi_i \otimes \varphi_j$$

We want to show that

$$\sum_{i,j=1}^{n} a_{ij}\varphi_i \otimes \varphi_j = 0$$

implies that $a_{ij} = 0$ for all i and j.

$$0 = \sum_{i,j=1}^{n} a_{ij}\varphi_i \otimes \varphi_j(e_k, e_l)$$

$$= \sum_{i,j=1}^{n} a_{ij}\delta_{ik}\delta_{jl} = a_{kl}$$

Then $\{\varphi_i \otimes \varphi_j\}_{i,j=1}^{n}$ a basis of $M_2(V, \mathbb{F})$ maps to a basis $\{e_i \otimes e_j\}_{i,j=1}^{n}$ a basis of $V \otimes V$ under the canonical transformation. ■

(2)

For $n = 3$, $k = 2$, $\mathcal{L}(V \wedge V, \mathbb{F}) \cong \mathcal{A}_2(V \times v, \mathbb{F})$. If $T \in \mathcal{A}_2(V \times V, \mathbb{F})$,

$$\sum_{i,j=1}^{3} T(e_i, e_j)\varphi(i)\varphi(j)$$

$$= T(e_1, e_2)\varphi_1(v)\varphi_2(w) + T(e_2, e_1)\varphi_2(v)\varphi_1(w) + \ldots + T(e_2, e_3)\varphi_2(v)\varphi_3(w) + T(e_3, e_2)\varphi_3(v)\varphi_2(w)$$

$$= T(e_1, e_2)(\varphi_1 \wedge \varphi_2) + T(e_1, e_3)\varphi_1 \wedge \varphi_3 + T(e_2, e_3)\varphi_2 \wedge \varphi_3$$

$$\{\varphi_i \wedge \varphi_j\}_{i<j}$$

generates $\mathcal{A}_2(V, \mathbb{F})$.

Definition. Consider $T \in \mathcal{L}(V)$. $(\wedge^k T) \in \mathcal{L}(\wedge^k V)$ with

$$(\wedge^k T)(e_{i_1} \wedge \ldots \wedge e_{i_k}) = T(e_{i_1}) \wedge \ldots \wedge T(e_{i_k})$$

If $k = n$, $\dim(\wedge^n v) = 1$, so $e_1 \wedge \ldots \wedge e_n$ basis of V. $\wedge^n T \in \mathcal{L}(\wedge^n V)$ is the **determinant** of T.

Example.

$$T = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{F}^{2,2}$$

With $e_1 = (1, 0)$ and $e_2 = (0, 1)$. Then

$$\det T = (\lambda^2 T)(e_1 \wedge e_2) = Te_1 \wedge Te_2 = (a_{11}e_1 + a_{21}e_2) \wedge (a_{12}e_1 + a_{22}e_2)$$

$$= a_{11}a_{22}e_1 \wedge e_2 + a_{21}a_{12}e_2 \wedge e_1$$

$$= (a_{11}a_{22} - a_{21}a_{12})(e_1 \wedge e_2)$$

Lemma. $T \in \mathcal{L}(V)$.

(1) $T \in \mathcal{L}(V)$ is invertible if and only if $\det T \neq 0$.

(2) $T, S \in \mathcal{L}(V)$. Then $\det(T \cdot S) = \det(T)\det(S)$.

(1)

$T \in \mathcal{L}(V)$ not being invertible is equivalent to $Tv_1, \ldots, Tv_n$ being linearly dependent for $\{v_1, \ldots, v_n\}$ basis of V. Then WLOG

$$Tv_n = \sum_{j=1}^{n-1} a_j Tv - j$$

$$(\det T)(v_1 \wedge \ldots \wedge v_n) = Tv_1 \wedge \ldots \wedge Tv_n = 0$$

since v_n contains v_i for all i and $v_i \wedge v_i = 0$. ■

(2)

$$\det(T \cdot S)(e_1 \wedge \ldots \wedge e_n) = (T \cdot S)(e_1 \wedge \ldots \wedge e_n) = T(Se_1 \wedge \ldots \wedge Se_n)$$

$$= TS(e_1) \wedge \ldots \wedge TS(e_n)$$

$$= T((\det S)e_1 \wedge \ldots \wedge e_n)$$

$$= (\det S)T(e_1 \wedge \ldots \wedge e_n)$$

$$= (\det S)(\det T)(e_1 \wedge \ldots \wedge e_n)$$

■

Corollary. Consider $T, S \in \mathcal{L}(V)$, $T \cdot S = I$. Then $\det T \det S = 1$.

Lemma. Consider $A, B \in \mathbb{F}^{n,n}$

(1) (Laplace Expansion). $\det(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}M_{ij}$, where M_{ij} is the determinant of the submatrix obtained from taking out row i and column j.

(2) Suppose $A = [Av_1 \quad \dots \quad Av_n] \in \mathbb{F}^{n,m}$ and $B = [Av_1 \quad \dots \quad Av_j \quad \dots Av_i \quad \dots \quad Av_n]$. Then $\det(B) = -\det(A)$.

(3) $A = [Av_1 \quad \dots \quad Av_i \quad \dots Av_i \quad \dots \quad Av_n]$. Then $\det(A) = 0$.

(4) Suppose B has the same columns of A, but the ith column of B is $\lambda Av_i + \mu Av_k$. Then $\det(B) = \lambda \det(A)$.

(5) $B = P^{-1}AP$ for some invertible matrix P. Then $\det(B) = \det(A)$.

We now consider determinants on complex vector spaces. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$.

Proposition. $\det(T) = \prod_{j=1}^{n} \lambda_j$, λ_js eigenvalues counted with multiplicities.

Proof.

Remark. we could use induction.

There exists a basis B such that $\mathcal{M}(T, B)$ is upper triangular.

$$Tv_1 \wedge Tv_2 \wedge \dots \wedge Tv_n = \lambda_1 v_1 \wedge (\lambda_2 v_2 a_1 v - 1) \wedge (\lambda_3 v_3 + b_2 v_2 + b_1 v_1) = (\lambda_1 \lambda_2 \dots \lambda_n) v_1 \wedge \dots \wedge v_n)$$

∎

Proposition. $\det(T - \lambda I) = p(\lambda)$ is the characteristic polynomial of T.

Definition. A permutationi $\tau : \{1, \dots, n\} \to \{1, \dots, n\}$ is bijection. The set of all permutations of $\{1, \dots, n\}$ is denoted by S_n.

Definition. Sign : $S_n \to \{-1, 1\}$ defined as follows:

$$\text{Sign}(\tau) = \begin{cases} +1 & \text{the natural order has been changed for an even number of times} \\ -1 & \text{the natural order has been changed for an odd number of times} \end{cases}$$

Lemma. $\tau, \sigma \in$ Sign.

(1) Sign$(\sigma \cdot \tau) =$ Sign$(\sigma) \cdot$ Sign(τ).

(2) $\tau\sigma = I$ implies that Sign$(\sigma) =$ Sign(τ).

Corollary. $\det(A^t) = \det(A)$.

Corollary. If A has two identical rows, then $\det(A) = 0$.

Corollary. $\det(A_{\tau(1),\cdot}, A_{\tau(2),\cdot}, \dots, A_{\tau(n),\cdot}) =$ Sign$(\tau)\det(A)$.

Lemma. Isometry $S \in \mathcal{L}(V)$ implies that $|\det S| = 1$.

Proof.

$$SS^* = S^*S = I$$
$$\det(S)\det(S^*) = 1$$
$$|\det S|^2 = 1$$

■

28 Lecture : End Credits

28.1 Volume

We define volume intuitively (with Euclidean Geometry intuition) with a unit hypercube having volume 1 and the hypercube with side length a having volume a^n.

Theorem. Consider $\Omega \subseteq \mathbb{R}^n$, and $A\Omega = \{Ax \mid x \in \Omega\}$. Then

$$\text{volume}(A\Omega) = |\det A|\text{volume}(\Omega)$$

We first know that isometry preserves volume since $|\det S| = 1$ for isometry S.
Proof 1 (Gram-Schmidt Process).

Lemma. If $\Omega_1 \subseteq \mathbb{R}^n$, volume$(\Omega_1) \neq 0$, then $|\det A| = \dfrac{\text{volume}(A\Omega_1)}{\text{volume}(\Omega_1)}$.

We notice that the volume of the figure spanned by $v_1, \dots, v_n$ is

$$||v_1|| ||v_2^\perp|| \dots ||v_n^\perp||$$

Now recall that we can do the QR-decomposition of A:

$$\begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} ||v_1|| & & & * \\ & ||v_2^\perp|| & & \\ & & \ddots & \\ 0 & & & ||v_n^\perp|| \end{bmatrix}$$

$$\det(A) = \det(Q) \cdot \det(R)$$

Now since $|\det Q| = 1$,

$$|\det A| = |\det R| = \text{volume(solid spanned by } v_1, \dots, v_n)$$

the result follows? ■

Proof 2 (Spectral Theorem).
Recall the polar decomposition

$$A = S\sqrt{A^t A}$$

Thus

$$\text{volume}(A\Omega) = \text{volume}(S\sqrt{A^t A}\Omega) = \text{volume}(B\Omega)$$

Let $w_1, \dots, w_n$ be an orthonormal eigenbasis for $\sqrt{A^t A}$. Pick Ω to be the solid spanned by $w_1, \dots, w_n$ and let the $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues.

$$\text{volume}(\Omega) = 1$$

$$\sqrt{A^t A}\Omega$$

is the solid spanned by $\lambda_1 w_1, \dots, \lambda_n w_n$. Then

$$\text{volume} = \prod_{i=1}^{n} \lambda_i \Rightarrow \frac{\text{volume}(A\Omega)}{\text{volume}(\Omega)} = \prod \lambda_i = \det \sqrt{A^t A} = |\det A|$$

■

Remark. I didn't completely follow for the above section (I came pretty late trying to find a parking spot) so the notes potentially might not make sense :eyes:. I may update them in the future. I also missed a board :eyes: so I'll probably study this off the textbook.

28.2 Trace

Consider $\dim V < \infty$.

Proposition. For all $\varphi \in V'$ and $v \in V$, $\Phi : V' \otimes V \to \mathcal{L}(V)$ with $\varphi \otimes v \mapsto (V \to V, w \mapsto \varphi(w)v)$ is an isomorphism.

Remark.

$$\mathcal{L}(V) \leftarrow V' \otimes V \to F$$

with $V' \otimes V \to \mathbb{F}$ being defined by $\varphi \otimes v \mapsto \varphi(v)$, the map from $\mathcal{L}(V)$ to $\varphi(v)$ is the trace map.

Proof.

$$\dim(V' \otimes V) = \dim(V')\dim V = (\dim V)^2 = \dim \mathcal{L}(V)$$

Consider $v_1, \ldots, v_n$ a basis of V and $\varphi_1, \ldots, m\varphi_n$ a dual basis of V'. Then $\{varphi_i \otimes v_j\}$ is a basis of $V' \otimes V$.

$$\mathcal{M}(\Phi(\varphi_i \otimes v_j), \{v_1, \ldots, v_n\}) = \begin{bmatrix} & 0 & \\ & 1 & \\ & 0 & \end{bmatrix}$$

where the 1 is in the ith column and jth row.

Exercise: prove that its injective/surjcetive.

Proof that the map from $\mathcal{L}(V)$ to $\varphi(v)$ is the trace. Consider $A = [a_{ij}]_{1 \le i,j \le n}$.

$$\Phi(\sum_{i,j} a_{ij}\varphi_j \otimes v_i) = A$$

$$\mathrm{ev}(\Phi^{-1}(A)) = \mathrm{ev}(\sum_{ij} a_{ij}\varphi_j \otimes v_i) = \sum_{i,j} a_{ij}\delta_{ij} = \sum_{i=1}^{n} a_{ii} = \mathrm{trace}\ A$$

■

28.3 Cayley-Hamilton

Theorem. Consider A, an $n \times n$ matrix. The characteristic polynomial f satisfies

$$f_A(\lambda) = \det(\lambda I - A)$$

and $f_A(A) = 0$.

Proof.
Consider the special case

$$\begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{bmatrix}$$

Using the problem from PSet 10,

$$A^n e_1 + a_{n-1}A^{n-1}e_1 + \ldots + a_0 e_1 = 0$$

The minimal polynomial for A and e_1 is

$$x^n + a_{n-1}x^{n-1} + \ldots + a_0$$

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & & & & a_0 \\ -1 & \lambda & & & a_1 \\ & -1 & \ddots & & a_2 \\ & & \ddots & & \vdots \\ & & & \lambda & a_{n-2} \\ & & & -1 & \lambda + a_{n-1} \end{bmatrix} \overset{\text{Laplace}}{=\!=} \lambda \det \begin{bmatrix} \lambda & & & a_1 \\ -1 & \ddots & & a_2 \\ & \ddots & & \vdots \\ & & \lambda & a_{n-2} \\ & & -1 & \lambda + a_{n-1} \end{bmatrix}$$

Claim: $\det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0$.
We prove this via induction. Using Laplace,

$$\det(\lambda I - A) = \lambda(\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \ldots + a_1) + (-1)^{n+1} a_0 (-1)^{n-1}$$

and the middle parentheses follows from the induction hypothesis. □

In particular, $f_A(A)e_1 = 0$. We now want to reduce to this special caes. We want to prove that for all $v \in V$, $f_A(A)v = 0$. Let W be the smallest invariant subspace of V containing v.

Claim: $W = \operatorname{span}(v, Av, \ldots, A^k v, \ldots) = \operatorname{span}(v, Av, \ldots, A^{m-1}v$ where m is the smallest positive integer such that $v, Av, \ldots, A^m v$ are linearly dependent. In particular, $v, \ldots, A^{m-1}v$ is a basis of W.

Since $v, Av, \ldots, A^m v$ are linearly dependent,

$$A^m v = -a_{m-1}A^{m-1}v - \ldots - a_0 v$$

Apply A^n to both sides with $n \geq 0$. For all $k \geq n$, we can write $A^k v$ as a linear combination of $A^{k-1}v, \ldots, v$. We induct on $k - n$ to conclude that $A^k v$ is a linear combination of $A^{m-1}v, \ldots, v$. □

Now:

$$A|_W = \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{m-2} \\ & & & 1 & -a_{m-1} \end{bmatrix} = B$$

Using our special case, $\det(\lambda I - B) = \lambda^m + a_{m-1}\lambda^{m-1} + \ldots + a_0$.

Claim: The matrix of A with respect to a basis of V extending $\{v, \ldots, A^{m-1}v\}$ is

$$\begin{bmatrix} B & * \\ 0 & * \end{bmatrix}$$

Since f_A is independent of basis,

$$\det(\lambda I - A) = \det(\lambda I - \begin{bmatrix} B & * \\ 0 & * \end{bmatrix}) = \det \begin{bmatrix} \lambda I - B & * \\ 0 & \lambda I - * \end{bmatrix}$$
$$= \det(\lambda I - B)\det(\lambda I - *)$$

Remark. Notice that

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

This comes from the following:

$$Ae_1 = a_{11}e_1 + \ldots + a_{n1}e_n$$
$$\vdots$$
$$Ae_n = a_{1n}e_1 + \ldots + a_{nn}e_n$$
$$\det A = Ae_1 \wedge \ldots \wedge Ae_n$$
$$= \sum_{\{i_1,\ldots,i_n\}=\{1,\ldots,n\}} a_{i_1 1}a_{i_2 2}\ldots a_{i_n n}e_{i_1} \wedge \ldots e_{i_n}$$
$$= \sum \text{sign of } (i_1, \ldots, e_n) a_{i_1 1} \ldots a_{i_n, n} \Rightarrow \det A^t = \det A$$

$$f_B(\lambda)|f_A(\lambda)$$

with | meaning divides and $f_B(A)v = 0$ implies that $f_A(A)v = 0 = s(A)f_B(A)$.

29 Precept : Post-Credits Scene

Remark. The last class . . .. A very nostalgic/indescribable feeling.

29.1 Determinants

Definition (Determinant). Consider an $n \times n$ matrix A with the form:

$$A = \begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{bmatrix}$$

Then

$$\det A = \sum_{(m_1,\dots,m_n)\in \text{perm}(n)} (\text{sign } (m_1, \dots, m_n)) A_{m_1,1} \cdot \dots \cdot A_{m_n,n}$$

The determinant detects whether a matrix is invertible. Specifically, A is invertible if and only if $\det A \neq 0$.

We consider some cases. In the one dimensional case,

$$\begin{bmatrix} a \end{bmatrix}$$

is invertible if and only if $a \neq 0$. In the two dimensional case,

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

is invertible if and only if $ad - bc \neq 0$.

One application of determinants is finding inverses.

Theorem. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

There are multiple proofs but ehre we will do the one by Cayley-Hamilton:

Proof.

Remark. In general, the characteristic polynomial is

$$z^n - (\text{trace } A)z^{n-1} + \dots + (-1)^n \det A$$

The characteristic polynomial is $z^2 - (a + d)z + (ad - bc)$, so by the Cayley-Hamilton theorem,

$$A^2 - (a + d)A + (ad - bc) = 0$$

Since A is invertible,

$$A - (a + d)I + (ad - bc)A^{-1} = 0$$

solving for the inverse we get the desired result. ■

We now wish to consider larger matrices. The formula for $\det A$ has $|\text{perm } (n)| = n!$ terms. For the 3 by 3 case, this is managable:

Example. Consider the matrix

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

A trick we can use is to write the determinant is to with the first two columns again, and all the permutations show up as diagonals with diagonals decreasing from left to right being positive and diagonals going from right to left being negative:

$$\left|\begin{matrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{matrix}\right| \begin{matrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{matrix}$$

we get Sarmr's rule:

$$\det A = A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32}$$
$$-A_{13}A_{22}A_{31} - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33}$$

Example. We now try to find the determinant:

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

Notice that the fourth column only has a 2 that is non zero, so all nonzero terms in the determinant must contain the 2

$$\det \begin{bmatrix} 0 & 0 & 1 & \color{red}{0} & 2 \\ \color{red}{5} & \color{red}{4} & \color{red}{3} & \color{blue}{2} & \color{red}{1} \\ 1 & 3 & 5 & \color{red}{0} & 7 \\ 2 & 0 & 4 & \color{red}{0} & 6 \\ 0 & 0 & 4 & \color{red}{0} & 4 \end{bmatrix}$$

Looking at the uncolored cells, only 3 is nonzero in its column, so all terms in the determinant must contain it:

$$\det \begin{bmatrix} 0 & \color{red}{0} & 1 & \color{red}{0} & 2 \\ \color{red}{5} & \color{red}{4} & \color{red}{3} & \color{blue}{2} & \color{red}{1} \\ \color{red}{1} & \color{blue}{3} & \color{red}{5} & \color{red}{0} & \color{red}{7} \\ 2 & \color{red}{0} & 4 & \color{red}{0} & 6 \\ 0 & \color{red}{0} & 4 & \color{red}{0} & 4 \end{bmatrix}$$

Again, the two in the first column is the only nonzero element so it must be in the determinant

$$\det \begin{bmatrix} \color{red}{0} & \color{red}{0} & \color{orange}{1} & \color{red}{0} & \color{cyan}{2} \\ \color{red}{5} & \color{red}{4} & \color{red}{3} & \color{blue}{2} & \color{red}{1} \\ \color{red}{1} & \color{blue}{3} & \color{red}{5} & \color{red}{0} & \color{red}{7} \\ \color{blue}{2} & \color{red}{0} & \color{red}{4} & \color{red}{0} & \color{red}{6} \\ \color{red}{0} & \color{red}{0} & \color{cyan}{4} & \color{red}{0} & \color{orange}{4} \end{bmatrix}$$

This means that the determinant is just

$$\mathsf{sign}(4,3,5,2,1)72 + \mathsf{sign}(4,3,1,2,5)48 = 72 - 48 = 24$$

In other words, if there are a lot of 0s in a matrix, this formula is useful, but otherwise its not very efficient.

We now discuss the cofactor expansion method of finding the determinant of an $n \times n$ matrix. Let

$$M_{ij} = \det(\text{the } (n-1) \times (n-1) \text{ matrix formed by omitting the } i\text{th row and } j\text{th column of } A)$$

Then we get the cofactor expansion along the jth column:

$$\det A = \sum_{i=1}^{n} A_{ij}(-1)^{i+j} M_{ij}$$

and the cofactor expansion along the ith row:

$$\det A = \sum_{j=1}^{n} A_{ij}(-1)^{i+j} M_{ij}$$

We can also find the determinant using column/row reduction.
Consider an $n \times n$ matrix A.

Step 1. Do column/row operations on A to get a simpler matrix B, keeping track of s the number of row/column swaps and $k_1, \dots, k_r$ what you divide the columns and rows by (r is the rank).

Step 2. $\det A = (-1)^s k_1 \dots k_r \det B$.

Example.

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 2 & 2 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} = B$$

Then $\det A = (-1)^0 \cdot 2 \cdot \det B = 12$ since $\det b = 1 \cdot 2 \cdot 3$ (since it is upper-triangular).

Example. Consider

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

Note that the characteristic polynomial of A is $\det(zI - A) = \det(zI - A^T)$. Thus, the characteristic polynomial is $(z-2)^3(z+1)$.

Example. Consider

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

We would like to find the Jordan form. First, we find the eigenvalues. We can use the characteristic polynomial to do this. Notice that

$$\det(zI - A) = \det \begin{bmatrix} z+1 & 0 & -1 \\ 3 & z & -1 \\ 4 & 0 & z-3 \end{bmatrix}$$

We do expansion along the second column. We get:

$$= z \det \begin{bmatrix} z+1 & -1 \\ 4 & z-3 \end{bmatrix}$$

$$= z((z+1)(z-3)+4) = z(z^2 - 2z - 3 + 4) = z(z^2 - 2z + 1) = z(z-1)^2$$

thus the eigenalues are 0 and 1.

Remark. Expansion is best when we are trying to find the characteristic polynomial because we don't want to have to do weird things like dividing by z.

Now we want to find bases for $G(0, A)$ and $G(1, A)$. For $\lambda = 1$,

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

this has rank 2. Also,

$$(A - I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 5 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Notice that $(-1, 5, 0)$ is in $\ker(A - I)^2$ but not $\ker(A - I)$. Then $(A - I)(-1, 5, 0) = (2, -2, 4)$ is in $\ker(A - I)$ in addition to $(-1, 5, 0)$.
For $\lambda = 0$, $(0, 1, 0)$ is in the kernel by inspection. Then

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -2 & 5 & 1 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ -2 & 5 & 1 \\ 4 & 0 & 0 \end{bmatrix}^{-1}$$