MATH 110: Linear Algebra

Jad Damaj

Spring 2022

Contents

Chapter 1

Vector Spaces

1.1 January 19

1.1.1 Overview

Linear Algebra is about:

- Vector spaces (also called linear spaces)
- Linear maps
- Extra structures on spaces/maps (inner products, etc.)

Motivation:

- Physics - we live in a 3D space
- Geometry - even for a curved object, locally it looks like a vector space (tangent space)
- Representation Theory and Group Theory - the set of all matrices has a rich structure and interesting subsets (subgroups)
- Solving Differential Equations - natural tool and solution spaces
- Normal Operators - guaranteed good bases
- Statistics - square matrices, ...
- Applied Math - designing of algorithms, ...

1.1.2 Ch1 - Vector Spaces

$\mathbb{R}$- set of reals, $\mathbb{R}^2$ - plane, $\mathbb{R}^3$ - 3D space

Key feature: Have addition and scalar multiplication by $\mathbb{R}$

Generalizations: Vector spaces over $\mathbb{R}$ (or a general $\mathbb{F}$)

1.1.3 1.A: $\mathbb{R}^n$ and $\mathbb{C}^n$

Definition 1.1.1 ($\mathbb{C}$)**.**
Introduced i such that $i^2 + 1 = 0$
$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$
Addition: $(a + bi) + (c + di) = (a + c) + (b + d)i$
Multiplication: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$

eg: $(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i)^2 = 0 + i \cdot i = i$
$\mathbb{R} \subset \mathbb{C}$: view x as $x + 0i$

Theorem 1.1.2 (Properties of $\mathbb{C}$)**.**
Commutativity: $\alpha + \beta = \beta + \alpha, \quad \alpha\beta = \beta\alpha \quad \forall \alpha, \beta \in \mathbb{C}$
Associativity: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma), \quad (\alpha\beta)\gamma = \alpha(\beta\gamma) \quad \forall \alpha, \beta, \gamma \in \mathbb{C}$
Additive Identity: $\alpha + 0 = \alpha \quad \forall \alpha \in \mathbb{C}$
Additive Inverse: $\forall \alpha \in \mathbb{C}$, $\exists! \beta \in \mathbb{C}$ such that $\alpha + \beta = 0$
Multiplicative Identity: $\alpha \cdot 1 = \alpha \quad \forall \alpha \in \mathbb{C}$
Multiplicative Inverse: $\forall \alpha \neq 0 \in \mathbb{C}$ $\exists! \beta \in \mathbb{C}$ such that $\alpha\beta = 1$
Distributive Properties: $\lambda(\alpha + \beta) = \lambda\alpha + \lambda + \beta \quad \forall \lambda, \alpha, \beta \in \mathbb{C}$

1.2 January 24

Example 1.2.1. Show existence and uniqueness of the multiplicative inverse of $\forall a \neq 0$
Idea: Assume $\alpha = a + bi$ want $(a + bi)(? + ?i) = 1 \to ? + ?i = \frac{1}{a+bi}$ "$=$" $\frac{a-bi}{(a+bi)(a-bi)} = \frac{1-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$

Proof. Assume $\alpha = a + bi$, $a, b \in \mathbb{R}$, not both zero. We see that $\beta = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$ satisfies $(a + bi)(\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i) = \frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2} = 1$. Similarly, $(\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i)(a + bi) = 1$. $\to$ existence
Moreover, if there exists $\tilde{\beta}$ such that $\alpha\tilde{\beta} = 1$, then $\beta = \beta\alpha\tilde{\beta} = \tilde{\beta}$. $\to$ uniqueness □

Definition 1.2.2.

- For $\alpha \in \mathbb{C}$, let $-\alpha \in \mathbb{C}$ define the unique element such that $\alpha + (-\alpha) = 0$
- For $\alpha \in \mathbb{C}$, let $1/\alpha \in \mathbb{C}$ define the unique element such that $\alpha(1/\alpha) = 1$
- Subtraction: $\alpha - \beta = \alpha + (-\beta)$
- Division: $\beta/\alpha = \beta \cdot (1/\alpha)$, $\alpha \neq 0$

$\mathbb{F}$: field(In the book, $\mathbb{R}$ or $\mathbb{C}$)

- In general, generalization of $\mathbb{R}$ or $\mathbb{C}$

Definition 1.2.3. A set $\mathbb{F}$(with addition "+" and multiplication "$\times$") is a field if:

(i) $\exists 0, 1 \in \mathbb{F}$, $0 \neq 1$

(ii) $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ by $(a, b) \mapsto a + b$

(iii) $\times : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ by $(a, b) \mapsto a \cdot b$

Satisfying:

(a) Commutivity: $a + b = b + a$, $ab = ba$

(b) Associativity: $a + (b + c) = (a + b) + c$, $a(bc) = (ab)c$

(c) Inverses: $\forall a$, $\exists - a$ such that $a + (-a) = 0$
$\forall a$, $\exists 1/a$ such that $a \cdot (1/a) = 1$

(d) Distributive: $c(a + b) = ca + cb$

Example 1.2.4.

1. $\mathbb{R}, \mathbb{C}$
2. $\{0, 1\}$ $+, \times \bmod 2$
3. $\mathbb{F}_p = \{0, \ldots, p - 1\}$ $+, \times \bmod p$, p prime
4. $\mathbb{Q}$: rationals
5. $\{a + b\sqrt{2} : a, b, \in \mathbb{Q}$
6. $\{P(x)/Q(x) : P, Q$ are polynomials over $\mathbb{F}$, $Q \neq 0\}$

We will define $\cdot$ for $\mathbb{F}$. Elements of $\mathbb{F}$ are known as scalars (as opposed to vectors)

Definition 1.2.5. An n-tuple of elements of $\mathbb{F}$ is $(x_1, \ldots, x_n)$ where each $x_i \in \mathbb{F}$

Definition 1.2.6. $\mathbb{F}^n = \{$all n-tuples of elements in $\mathbb{F}\}$

Definition 1.2.7.

- Addition "+": $\mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$ by $((x_1, \ldots, x_n), (y_1, \ldots, y_n)) \mapsto (x_1 + y_1, \ldots, x_n + y_n)$
- Scalar Multiplication "$\cdot$": $\mathbb{F} \times \mathbb{F}^n$ by $(\lambda, (x_1, \ldots, x_n)) \mapsto (\lambda x_1, \ldots, \lambda x_n)$

Theorem 1.2.8 (Properties of $\mathbb{F}^n$).

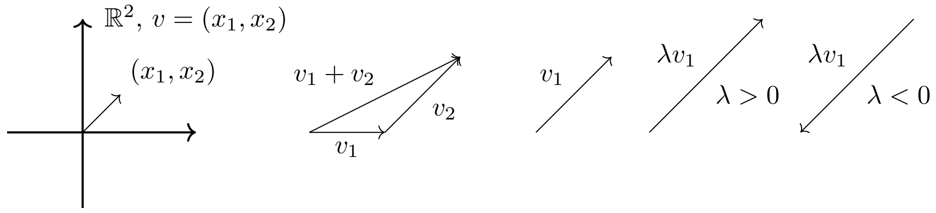
- Addition is commutative: $v_1 + v_2 = v_2 + v_1 \quad \forall v_1, v_2 \in \mathbb{F}^n$

  Proof. Assume $v_1 = (x_1, \ldots, x_n)$, $v_2 = (y_1, \ldots, y_n)$ then
  $v_1 + v_2 = (x_1 + y_1, \ldots, x_n + y_n) = (y_1 + x_1, \ldots, y_n + x_n) = v_2 + v_1$

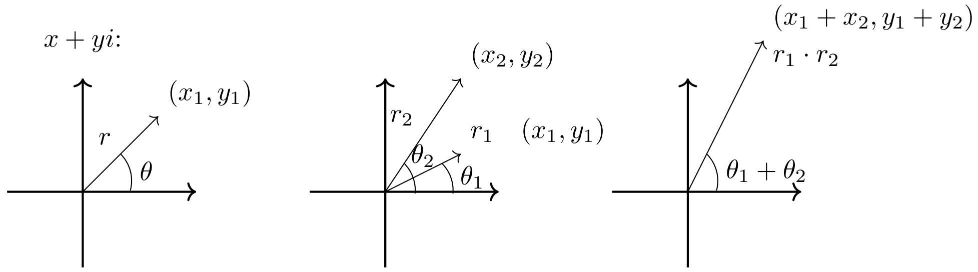
- Existence of $0 \in \mathbb{F}^n$: Denote $0 = (0, \ldots, 0)$. Then $v + 0 = v \ \forall v \in \mathbb{F}^n$
- Additive Inverse: $\forall v \in \mathbb{F}^n$, $\exists!(-v) \in \mathbb{F}^n$ such that $v + (-v) = 0$

Geometric Meaning for $\mathbb{F} = \mathbb{R}$

Descartes Coordinate System:



Geometric Meaning of Multiplication on $\mathbb{C}$



1.2.1 1B - Vector Spaces

Definition 1.2.9. Fix a field $\mathbb{F}$. A vector space over $\mathbb{F}$ is a set V with addition "+" and scalar multiplication "$\cdot$ " denoted as $+ : V \times V \to V$ by $(v_1, v_2) \mapsto v_1 + v_2$, $\cdot : \mathbb{F} \times V \to V$ by $(\lambda, v) \mapsto \lambda v$
Satisfies:

(1) $u + v = v + u$, $\forall u, v \in V$

(2) $(u + v) + w = u + (v + w)$, $a(bv) = (ab)v$ $\forall u, v \in \mathbb{V}, a, b \in \mathbb{F}$

(3) $\exists 0 \in \mathbb{V}$ such that $v + 0 = v$, $\forall v \in V$

(4) $\forall v \in V$, $\exists w \in V$ such that $v + w = 0$. (we will show w is unique and denote it as $-v$)

(5) $1 \cdot v = v$, $\forall v \in V$

(6) $a(u + v) = au + av$, $(a + b)v = av + bv$, $\forall a, b \in \mathbb{F}, u, v \in V$

Definition 1.2.10. Elements in a vector space V are called points or vectors

Definition 1.2.11. A vector space over $\mathbb{F}(/\mathbb{F})$ is also called an $\mathbb{F}$-vector space

Example 1.2.12.

(1) $\mathbb{R}^n, \mathbb{C}^n, \mathbb{F}^n$ are all vector spaces

(2) $\mathbb{C}$ is a vector space over $\mathbb{R}$

(3) Let S be a set. Define F^s = the set of all functions from S to $\mathbb{F}$. $\mathbb{F}^S$ is a vector space $/\mathbb{F}$ under the operations $(f + g)(s) = f(s) + g(s)$, $(\lambda f)(s) = \lambda \cdot f(s)$. Each element has additive inverse $(-f)(s) = -f(s)$

$\mathbb{F}^\infty = \mathbb{F}^{\{1,2,3,\ldots,\}}$, consists of $(a_1, a_2, a_3, \ldots)$ $\forall a_n \in \mathbb{F}$

(4) the set of all sequences of real numbers that converge to 0

(5) the set of all polynomials over $\mathbb{F}$, with deg $\leq n$ in k variables is a vector space $/\mathbb{F}$

Theorem 1.2.13. A vector space V has a unique additive identity

Proof. Assume 0 and $0'$ are both additive inverses. Then $0 = 0 + 0' = 0'$

Theorem 1.2.14. $\forall v \in V$ has a unique additive inverse.

Proof. If w_1, w_2 are both additive inverses of v, then $w_1 = w_1 + (v + w_2) = (w_1 + v) + w_2 = w_2$

Definition 1.2.15. Let $w - v = w + (-v)$

Notation 1.2.16. V will be used to denote a vector space over $\mathbb{F}$

Theorem 1.2.17. $0 \cdot v = 0$, $\forall v \in V$

Proof. $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$ so by the existence of additive inverses $0 = 0 \cdot v$

Theorem 1.2.18. $a \cdot 0 = 0$, $\forall a \in \mathbb{F}$

Proof. $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$ so $0 = a \cdot 0$

Theorem 1.2.19. $(-1) \cdot v = -v$, $\forall v \in V$

Proof. $0 = 0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v$ so by def $(-1)v = -v$

1.3 January 26

1.3.1 1.C - Subspaces

Definition 1.3.1. Assuming V is a vector space $/\mathbb{F}$. $U \subset V$ is called a subspace of V if U is also a vector space $/\mathbb{F}$ under $+$ and $\cdot$ in V.

Example 1.3.2. $\{(x_1, x_2, 0) : x_1, x_2, \in \mathbb{F}\}$ is a subspace of $\mathbb{F}^3$

Proposition 1.3.3. U is a subspace iff

(i) $0 \in V$

(ii) $u_1, u_2 \in U \to u_1 + u_2 \in U$

(iii) $a \in \mathbb{F}$, $u \in U \to a \cdot u \in U$

Proof. $\to$) Suppose conditions hold. Then properties of $+, \cdot$ follow from V, U has identity by (i) and additive inverses by (iii). Finally, $+, \cdot$ well defined by (ii), (iii) so U is a subspace.
$\leftarrow$) Suppose U is a subspace. Then U is nonempty so $0 \cdot u = 0 \in U$ so (i) holds. Also, $+, \cdot$ well defined so (ii), (iii) hold.

Example 1.3.4.

(a) $\{0\}$ is a subspace

(b) $\{(x_1, x_2, x_3, x_3) \in \mathbb{F}^4 : x_3 = 5x_4 + b\}$ is a subspace iff $b = 0$

(c) $C[0,1] = \{\text{continuous real valued functions on } [0,1]\}$ is a subspace of $\mathbb{R}^{[0,1]}$ (over $\mathbb{R}$)

(d) $C^\infty[0,1] = \{\text{smooth real-valued functions on } [0,1]\}$ is a subspace $\mathbb{R}^{[0,1]}$

(e) The real sequences that converge to zero form a subspace of $\mathbb{R}^\infty$

(f) The only subspaces of $\mathbb{F}^1$ are $\{0\}$ and $\mathbb{F}$ (over $\mathbb{F}$)

(g) If U is a subspace of V, W is a subspace of U, then W is a subspace of V

(h) We will show the only subspace of $\mathbb{R}^3$ are $\{0\}$, lines through the origin, planes through the origin, and $\mathbb{R}^3$

Definition 1.3.5. For $U_1, \ldots, U_n$ subspaces of V, define the sum

$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m : u_1 \in U_1, \ldots, u_m \in U_m\}$$

also denoted as $\sum_{j=1}^m U_j$.

Example 1.3.6. In $\mathbb{F}^3$, what is $\{(x, x, 0)\} + \{(0, y, y)\}$?

Proof. $\{(x, y, z) : y = x + z\}$ ☐

Theorem 1.3.7. For subspaces $U_1, \ldots, U_m \subset V$, $\sum_{j=1}^m U_j$ is a subspace. Moreover, it is the smallest subspace containing $U_1, \ldots, U_n$ in the sense that if W contains $U_1, \ldots, U_n$, then $W \supset U$.

Proof. Subspace: (i) $0 \in U_i$ for $i = 1, \ldots, n$ so $0 = 0 + \cdots + 0 \in W$
(ii)/(iii): follow from closedness of each U_j
Containing $U_1, \ldots, U_n$: Consider the sum $0 + \cdots + 0 + u_j + 0 + \cdots + 0$ for $j = 1, \ldots, m$
Smallest Subspace: Suppose W contains $U_1, \ldots, U_m$ then W contains $u_1, \ldots, u_m$ $\forall u_j \in U_j$ so $u_1 + \cdots + u_m \in$

W.

1.3.2 Direct Sums

Definition 1.3.8. If $U_1, \ldots, U_m$ are subspaces of V then the sum $U_1 + \cdots + U_m$ is a direct sum if each element in $U_1 + \cdots + U_m$ can be written as $u_1 + \cdots + u_m$ in a unique way with $u_j \in U_j$. In this case, we also use $U_1 \oplus \cdots \oplus U_m$ to denote $U_1 + \cdots + U_m$.

Example 1.3.9.

(1) If $U_1 = \{(x_1, x_2, 0) x_1, x_2 \in \mathbb{F}\}$, $U_2 = \{(0, 0, x_3) x_3 \in \mathbb{F}\}$, then $\mathbb{F}^3 = U_1 \oplus +U_2$.

(2) Let $U = \{(x, x, \ldots) \in \mathbb{R}^\infty$, $V = \{(x_1, x_2, \ldots) \in \mathbb{R}^\infty : \sum |x_n| < \infty, \sum x_n = 0\}$. Then $U + V$ is a direct sum.
(ex): Prove $U + V \neq \mathbb{R}^\infty$

Theorem 1.3.10. $U_1 + \cdots + U_m$ is a direct sum iff $\exists!$ way to write 0 as a sum of $u_1 + \cdots + u_m$, $\forall u_j \in U_j$ (which is $0 = 0 + \cdots + 0$).

Proof. $\rightarrow$) by def
$\leftarrow$) For $u \in U_1 + \cdots + U_m$, assume $u = u_1 + \cdots + u_m = \tilde{u_1} + \cdots + \tilde{u_n}$, $u_j, \tilde{u_j} \in U_j$. Then $(u_1 - \tilde{u_1}) + (u_2 - \tilde{u_2}) + \cdots + (u_m - \tilde{u_m}) = 0$. Hence $u_1 - \tilde{u_1} = u_2 - \tilde{u_2} = \cdots = 0$. Thus there is only one way to write u as $\sum_{j=1}^m$, $\forall u_j \in U_j$.

Theorem 1.3.11. For subspaces $U_1, U_2 \in V$, $U_1 + U_2$ is a direct sum iff $U_1 \cap U_2 = \{0\}$.

Proof. $\rightarrow$) If $v \in U_1 \cap U_2$, $\underbrace{v}_{\in U_1} + \underbrace{(-v)}_{\in U_2} = 0$ so $v = (-v) = 0$
$\leftarrow$) Take $u \in U_1 + U_2$ assume $u = u_1 + u_2 = \tilde{u_1} + \tilde{u_2}$. Then $\underbrace{u_1 - \tilde{u_1}}_{\in U_1} = \underbrace{-(u_2 - \tilde{u_2}}_{\in U_2}$ so by assumptions, $u_1 = \tilde{u_1}$ and $u_2 = \tilde{u_2}$.

Example 1.3.12. For subspaces $U_1, \ldots, U_m$ of V, TFAE:

(i) $U_1 + \cdots + U_m$ is a direct sum

(ii) $\forall j\ U_j \cap (\sum_{k \neq j} U_k) = \{0\}$

(iii) $\forall j > 1$, $U_j \cap (U_1 + \cdots + U_{j-1}) = \{0\}$

(iv) If $u_1 + \cdots + u_m = 0$, $u_j \in U_j$ then $u_1 = u_2 = \cdots = u_m = 0$

1.3.3 Chapter 2: Finite Dimensional Vector Spaces

$\mathbb{F}$: field, V : Vector space $/\mathbb{F}$

1.3.4 2.A: Span and Linear Independence

Motivation: In some V(such as $\mathbb{F}^n$), we can find vectors $e_1 = (1,0,\ldots,0), e_2 = (0,1,0,\ldots,0),\ldots,e_n = (0,\ldots,0,1)$ such that every $v \in V$ can be written as $v = \sum_{j-1}^n a_j e_j$ and the choice of a_j is unique.

We will work with such vectors in a general setting.

1.4 January 31

1.4.1 Chapter 2: Finite Dimensional Vector Spaces

- Main motivation: Find "coordinate systems" in a vector space
- Recall in $\mathbb{F}^n$, $(x_1,\ldots,x_n) = x_1(1,0\ldots,0) + x_n(0,\ldots,0,1) = x_1e_1 + \cdots + x_ne_n$.

1.4.2 2.A: Span and Linear Independence

Definition 1.4.1. A linear combination of vectors $v_1,\ldots,v_m \in V$ is a vector of the form

$$v = \sum_{j=1}^m a_j v_j, \quad a_1,\ldots,a_m \in \mathbb{F}.$$

Example 1.4.2. $(1,2,-3) = (1,0,-1) + 2(0,1,-1)$

Example 1.4.3. Is $(1,2,3)$ a linear combination of $(1,0,-1)$ and $(0,1,1)$?
No, if $(1,2,-3) = a_1(1,0,-1) + a_2(0,1,1)$ then $a_1 = 1, a_2 = 2$ but $1(1,0,-1)+2(0,1,1) = (1,2,1) \neq (1,2,-3)$.

Definition 1.4.4. The set

$$\{\sum_{j=1}^m a_j v_j, a_i \in \mathbb{F}, \forall 1 \leq j \leq m\}$$

is the span of $v_1,\ldots,v_m$, denoted by $\text{span}(v_1,\ldots,v_m)$. Note $\text{span}() = \{0\}$.

Example 1.4.5. $(1,2,-3) \in \text{span}((1,0,-1),(0,1,-1))$.

Theorem 1.4.6. $\text{span}(v_1,\ldots,v_m)$ is the smallest subspace of V that contains $v_1,\ldots,v_m$.

Proof. Subspace: $0 = 0v_1 + \cdots, 0v_n \in \text{span}(v_1,\ldots,v_m)$
Closed under addition: $(a_1v_1 + \cdots + a_mv_m) + (b_1v_1 + \cdots + b_mv_m) = (a_1+b_1)v_1 + \cdots + (a_m+b_m)v_m$.
Closed under multiplication: $\lambda(a_1v_1 + \cdots + a_mv_m) = (\lambda a_1)v_1 + \cdots + (\lambda a_m)v_m$.
So it is a subspace.
Smallest: If $v_1,\ldots,v_m \in W$ for some subspace W, then $\forall a_1,\ldots,a_n \in \mathbb{F}$, $a_1v_1,\ldots,a_mv_m \in V$ so $a_1v_1 + \cdots + a_mv_m \in W$. Thus, $\text{span}(v_1,\ldots,v_m) \subseteq W$.

Definition 1.4.7. If $V = \text{span}(v_1,\ldots,v_m)$, then we say the list $v_1,\ldots,v_m$ spans V.

Example 1.4.8. $e_1,\ldots,e_n$ spans $\mathbb{F}^n$

Definition 1.4.9. V is called finite dimensional if some (finite) list of vectors spans V.

Example 1.4.10. $\mathbb{F}^n$ is finite dimensional.

Definition 1.4.11. A finite expression

$$p(z) = a_0 + a_1 z^1 + \cdots + a_m z^m, \quad a_0, \ldots, a_m \in \mathbb{F}, a_m \neq 0, \tag{*}$$

also written as

$$\sum_{j=0}^{\infty} a_j x_j, a_{n+1} = a_{n+2} = \cdots = 0,$$

is called a polynomial with coefficients in $\mathbb{F}$. (By definition $p = 0$ is a polynomial.)

- Each polynomial over $\mathbb{F}$ gives rise to a function from $\mathbb{F} \to \mathbb{F}$ defined by $p : \mathbb{F} \to \mathbb{F}$ by $z \mapsto p(z)$
- m is the degree of p if p has the form $(*)$. The zero polynomial has degree $-\infty$ by definition.
- $\mathcal{P}(\mathbb{F}) = \{\text{all polynomials over } \mathbb{F}\}$
- $\mathcal{P}_m(\mathbb{F}) = \{\text{all polynomials of deg} \leq m \text{ over } \mathbb{F}\}$

Example 1.4.12. $\mathcal{P}_m(\mathbb{F})$ and $\mathcal{P}(\mathbb{F})$ are vector spaces over $\mathbb{F}$ (also subspaces of $\mathbb{F}^{\mathbb{F}}$ if viewed as functions.)

Example 1.4.13.

(a) $\mathcal{P}_m(\mathbb{F})$ is finite dimensional

(b) $\mathcal{P}(\mathbb{F})$ is infinte dimensional

Proof.

(a) $1, z, \ldots, z^m$ spans $\mathcal{P}_m(\mathbb{F})$

(b) For any $p_1, \ldots, p_m \in \mathcal{P}(\mathbb{F})$, assume N is larger than $\deg p_j$ for $1 \leq j \leq m$. Then every $\sum_{j=1}^m a_j p_j$ is not equal to z^N.

□

Definition 1.4.14. $v_1, \ldots, v_m$ is called linearly independent if whenever $0 = \sum_{j=1}^m a_j v_j$, $a_1, \ldots, a_m \in \mathbb{F}$, we must have $a_1 = \cdots = a_m = 0$. Otherwise, the list is linearly dependent. The empty list is linearly independent by definition.

Example 1.4.15.

(a) v is linearly independent iff $v \neq 0$

(b) $e_1, \ldots, e_n$ is linearly independent in $\mathbb{F}^n$

(c) v_1, v_2 is linearly independent iff neither vector is a scalar multiple of the other.

(d) $1, z, \ldots, z^m$ is linearly independent in $\mathcal{P}_m(\mathbb{F})$.

(e) $(1, *, *), (0, 1, *), (0, 0, 1)$ where each $*$ is arbitrary is linearly independent in $\mathbb{F}^3$

(f) $(1, 1, \ldots, 1), (a_1, a_2, \ldots, a_n), (a_1^2, a_2^2, \ldots a_n^2), \ldots, (a_1^{n-1}, a_2^{n-1}, \ldots, a_n^{n-1})$ is linearly dependent iff at least two of the a_j's are the same.

(g) Any subset of a linearly independent list is linearly independent.

Example 1.4.16.

(a) If some vector in a list is a linear combination of the others, then the list is linearly dependent.

(b) Every list containing 0 is linearly dependent.

1.5 February 2

1.5.1 2.A: Span and Linear Independence

Notation 1.5.1. $\mathcal{P}(\mathbb{F})$ can also be written as $\mathbb{F}[x]$

Lemma 1.5.2. For $v_1, \ldots, v_n \in V$, TFAE:

(a) $v_1, \ldots, v_n$ is linearly dependent.

(b) $\exists\, 1 \leq j \leq n$ such that $v_j \in \text{span}(v_1, \ldots, v_{j-1})$

(c) $\exists\, 1 \leq j \leq n$ such that $v_j \in \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$ (Note: here $\hat{v}_j$ means v_j is excluded from the list)

(d) $\exists\, 1 \leq j \leq n$ such that $\text{span}(v_1, \ldots, v_n) = \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$.

Proof. a $\to$ b) By def, $\exists\, a_1, \ldots, a_n \in \mathbb{F}$, not all 0, such that $a_1 v_1 + \cdots + a_n v_n = 0$. Take the largest j such that $a_j \neq 0$. Then, $a_1 v_1 + \cdots + a_j v_j = 0$. Hence, $v_j = -\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1}$ so $v_j \in \text{span}(v_1, \ldots, v_{j-1})$.
b $\to$ c) Notice $\text{span}(v_1, \ldots, v_{j-1}) \subset \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$ so $v_j \in \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$.
c $\to$ d) By assumption $v_j \in \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$. Also $v_k \in \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$ for $k \neq j$ so $\text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$ contains $v_1, \ldots, v_n$. Thus, it contains $\text{span}(v_1, \ldots, v_n)$. Since $\text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n) \subset \text{span}(v_1, \ldots, v_n)$, the two are equal.
d $\to$ a) By assumption, $\exists\, b_k \in \mathbb{F}$, $1 \leq k \leq n$, $k \neq j$ such that $v_j = \sum_{j \neq k} b_k v_k$. So $\sum_{j \neq k} b_k v_k - v_j = 0$ so the set is linearly dependent.

Theorem 1.5.3. If $v_1, \ldots, v_m$ spans V, and $u_1, \ldots, u_n \in V$ are linearly independent, then $n \leq m$.

Idea. If $m = 2$, why can't $n = 3$?
Consider the functions:

$$\begin{aligned} u_1 &= a_{1,1} v_1 + a_{1,2} v_2 \\ u_2 &= a_{2,1} v_1 + a_{2,2} v_2 \\ u_3 &= a_{3,1} v_1 + a_{3,2} v_2 \end{aligned}$$

We must rearrange u_1, u_2, u_3 to show they are linearly dependent (3 equations in 2 variables.) $\square$

Proof. We will proceed by induction on m.
Note that for $m = 0$, span() $= \{0\}$ so this is trivially true.
Basis: If $m = 1$, $n \geq 2$. Let v_1 span V and let $u_1, u_2 \in V$ be arbitrary. Then $u_1 = \lambda_1 v_1$ and $u_2 = \lambda_2 v_2$. If $\lambda_1 = 0$, then $u_1 = 0$ and the set is linearly dependent so assume $\lambda_1 \neq 0$. Then $\lambda_2 u_1 - \lambda_1 u_2 = 0$ so the list is linearly dependent.
Inductive Step: Assume that the theorem holds for $m = k$. It suffices to show the $m = k + 1$ case. Let $v_1, \dots, v_{k+1}$ be a spanning list of V. If $n \geq k + 2$, let

$$u_i = \sum_{j=1}^{k+1} a_{i,j} v_j, \quad 1 \leq i \leq k+2, \quad a_{i,j} \in \mathbb{F},$$

be a list of $k + 2$ vectors.
If all $a_{i,k+1} = 0$, then the list of vectors can be represented using only the vectors $v_1, \dots, v_k$ so they would be linearly independent by the IH.
Otherwise, WLOG, assume $a_{k+2,k+1} \neq 0$ (if not we can relabel the vectors to achieve this). Then

$$u_i - \frac{a_{i,k+1}}{a_{k+2,k+1}} u_{k+2} \in \mathrm{span}(v_1, \dots, v_k)$$

for $1 \leq i \leq k + 1$.
By IH, $\exists\, b_1, \dots, b_{k+1} \in \mathbb{F}$, not all 0, such that

$$b_1(u_1 - \frac{a_{1,k+1}}{a_{k+2,k+1}} u_{k+2}) + \cdots + b_{k+1}(u_{k+1} - \frac{a_{k+1,k+1}}{a_{k+2,k+1}} u_{k+1} u_{k+2}) = 0$$

so

$$b_1 u_1 + \cdots + b_{k+1} u_{k+1} - (b_1 \frac{a_{1,k+1}}{a_{k+2,k+2}} + \cdots + b_{k+1} \frac{a_{k+1,k+1}}{a_{k+1,k+2}}) u_{k+2} = 0$$

so the list $u_1, \dots, u_{k+2}$ is linearly dependent.

Example 1.5.4. $e_1, \dots, e_n$ spans $\mathbb{F}^n$ and is linearly independent so:

- $(1, 2, 3), (4, 5, 8), (4, 6, 7), (-3, 2, 8)$ are linearly dependent in $\mathbb{F}^3$
- $(1, 2, 3, -5), (4, 5, 8, -3), (4, 6, 7, -1)$ does not span $\mathbb{F}^4$

Proposition 1.5.5. Every subspace of a finite dimensional vector space is finite dimensional.

Proof. Assume V is spanned by $v_1, \dots, v_m$, and U is a subspace of V.
Start from the empty list () in U and add vectors 1 by 1 so that no vector is in the span if the previous vectors in the list. This produces a linearly independent list in U.
By the thm, this process must terminate since the length of a list of linearly independent vectors in V cannot be greater than m.
Assume we have $u_1, \dots, u_n$. Then each $u \in U$ is a linear combination of $u_1, \dots, u_n$, otherwise we could add it to the list and it would produce a longer linearly independent list. Thus, $u_1, \dots, u_n$ spans U.

1.5.2 2.B - Bases

Definition 1.5.6. A basis of V is a list of vectors in V that is linearly independent and spans V.

Theorem 1.5.7. Every finitely dimensional vector space has a basis.

Proof. Take $U = V$ in the proof of proposition 5.5. Then we can generate a linearly independent list in V that spans V. Thus V has a basis.

Example 1.5.8.

(a) $e_1, \ldots, e_n$ forms a basis of $\mathbb{F}^n$ (standard basis)

(b) $(1, 2, 3), (3, 4, 6), (0, 0, 1)$ is a basis of $\mathbb{F}^3$ unless char $\mathbb{F} = 3$

(c) $(1, -1, 0), (0, 1, -1)$ is a basis of $\{(x, y, z) \in \mathbb{F}^3 : x + y + z = 0\}$

(d) $1, z, \ldots, z^m$ is a basis of $\mathcal{P}_m(\mathbb{F})$

(e) $f_0, f_1, \ldots, f_m$ is a basis of $\mathcal{P}_m(\mathbb{F})$ if $\deg f_j = j$, $0 \leq j \leq m$

Proposition 1.5.9. $v_1, \ldots, v_m$ forms a basis of V iff $\forall v \in V$ can be uniquely represented as $v = \sum_{j=1}^{n} a_j v_j$, $a_j \in \mathbb{F}$.

Proof. If $v_1, \ldots, v_n$ forms a basis of V, then they span V so all vectors can be represented in the desired form. Suppose $\exists\, a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{F}$ such that $a_1v_1 + \cdots + a_nv_n = v = b_1v_1 + \cdots + b_nv_n$, then $(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0$. Since the set is linearly independent, $a_1 - b_1 = \cdots = a_n - b_n = 0$ so $a_i = b_i$ for all i, thus the representation is unique.
If the stated conditions hold, then the list spans v. Also, 0 has a unique representation so the list is linearly independent and hence a basis.

Proposition 1.5.10. Every spanning list in a finite dimensional vector space contains a basis.

Proof (Proof 1)**.** Starting from (), we can create a linearly independent list consisting of vectors in the spanning list by the same procedure as proposition 5.5. This process must terminate since the spanning list is finite. So we produce a linearly independent list that spans V, eg. a basis.

Proof (Proof 2)**.** We can also start with the spanning list $v_1, \ldots, v_m$ and at each step, if the list is linearly dependent, we can choose v_j such that $\text{span}(v_1, \ldots, v_n) = \text{span}(v_1, \ldots, \hat{v}_j, \ldots, v_n)$. This process must terminate in a linearly independent list since the spanning list is finite. So we produce a linearly independent list that spans V, eg. a basis.

Chapter 2

Linear Maps

2.1 February 2

2.1.1 Ch3 - Linear Maps

Notation 2.1.1. U, V, W will represent subspaces.

2.1.2 3.A - Linear Maps as a Vector Space

Definition 2.1.2. $T : V \to W$ is called a linear map if $\begin{cases} T(u+v) = Tu + Tv & \forall u, v \in V \\ T(\lambda v) = \lambda T v & \forall \lambda \in \mathbb{F}, v \in V \end{cases}$. Note: V is called the domain of T.

Definition 2.1.3. $\{$linear maps from V to $W\}$ is denoted by $\mathrm{Hom}(V, W)$ $(\mathcal{L}(V, W))$. $\mathrm{Hom}(V, V) = \mathrm{End}(V)$.

Example 2.1.4.

(1) Zero map: $0 \in \mathrm{Hom}(V, W)$ $\quad 0 : V \to W$ by $v \mapsto 0$

(2) Identity: $I \in \mathrm{End}(V)$ $\quad I : V \to W$ by $v \mapsto v$

(3) Inclusion: "i''. If $V \subseteq W$, $i : V \to W$ by $v \mapsto v$

(4) Differentiation: $D : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ by $\sum_{j=0}^{\infty} a_j x^j \mapsto \sum_{j=1}^{\infty} j a_j x^{j-1}$. $D \in \mathrm{End}(\mathcal{P}(\mathbb{F}))$

(5) Integration from 0 to $1 \in \mathrm{End}(\mathcal{P}(\mathbb{R}))$

(6) "Multiplication by f": Fix $f \in \mathcal{P}(\mathbb{F})$. Let $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ by $g \mapsto g \cdot f$. $[(\sum_j a_j x^j)(\sum_j b_j x^j) = \sum_{k=0}^{\infty} (\sum_{j_1 + j_2 = k} a_{j_1} b_{j_2}) x^k]$. $T \in \mathrm{End}(\mathcal{P}(\mathbb{F}))$.

(7) For a matrix

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \cdots & \\ a_{m,1} & \cdots & a_{m_n} \end{pmatrix}$$

$T : \mathbb{F}^m \to \mathbb{F}^m$ by $(x_1, \ldots, x_n) \mapsto (a_{1,1} x_1 + \cdots + a_{1,n} x_n, \ldots, a_{1,m} x_1 + \cdots + a_{n,m} x_n)$. $T \in \mathrm{Hom}(\mathbb{F}^n, \mathbb{F}^m)$.

2.2 February 7

2.2.1 2.B - Bases

Proposition 2.2.1. Every linearly independent list of vectors in a finite dimensional vector space can be extended to a basis.

Proof. Starting with a linear independent list we can continue to add vectors that are not in the span of the previous vectors of the list to product a linearly independent list. This will eventually terminate since a linearly independent list cannot be longer than the spanning list. Thus, it will produce a linearly independent spanning list, eg. a basis.

Proposition 2.2.2. If V is finite dimensional and U is a subspace of V, then there exists a subspace $W \subset V$ such that $V = U \oplus W$.

Proof. U is finite dimensional so take a basis $u_1, \ldots, u_n$ of U. Extend this to a basis $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$ of V. We will show $W = \text{span}(u_{m+1}, \ldots, u_n)$ suffices.
Since $u_1, \ldots, u_n$ is a basis of V, every $v \in V$ can be written as $\underbrace{a_1u_1 + \cdots + a_mv_m}_{\in U} +$ $\underbrace{a_{m+1}u_{m+1} + \cdots + a_nu_n}_{\in W}$ so $U + W = V$.
Moreover, if $w \in U \cap W$, then $w = \sum_{j=1}^m b_jv_j$ and $w = \sum_{j=m+1}^n b_jv_j$ for $b_1, \ldots, b_n \in \mathbb{F}$. Hence, since $\sum_{j=1}^m b_jv_j - \sum_{j=m+1}^n b_jv_j = 0$, all $b_j = 0$ so $w = 0$.

2.2.2 2C - Dimension

Theorem 2.2.3. Any two bases of a finite dimensional vector space have the same length.

Proof. Bases are spanning lists and linearly independent lists so for two bases B_1, B_2, len$B_1 \leq$ lenB_2 and len$B_2 \leq$ lenB_1 so lenB_1 =lenB_2.

Definition 2.2.4. The dimension of a finite dimensional vector space is the length of every basis, denoted $\dim V$

Example 2.2.5.

(a) $\dim \mathbb{F}^n = n$

(b) $\mathbb{C}$ as a vector space over $\mathbb{R}$ has dimension 2. eg. $\dim_{\mathbb{R}} \mathbb{C} = 2$

(c) $\dim \mathcal{P}_m(\mathbb{F}) = m + 1$

(d) $\dim\{(x_1, \ldots, x_n) \in \mathbb{F}^n : x_1 + \cdots + x_n = 0\} = n - 1$.
A basis is $(1, -1, 0, \ldots, 0), (0, 1, -1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1, -1)$.

(e) Every subspace $U \subset V$ such that $U \neq V$ has $\dim U < \dim V$.

Proof. Take a basis of U and extend to a basis of V. We must add ≥ 1 element, otherwise $U = V$. □

(f) Every vector space $\neq \{0\}$ has $\dim \geq 1$.

Proof. Take a nonzero element (linearly independent) and extend to a basis. Thus $\dim \geq 1$. □

Theorem 2.2.6. If V is fin dim with $\dim V = n$, then if a list of n vectors is linearly independent it is a basis.

Proof. Extend the list to a basis. Since the basis has length n no vectors were added so the list is already a basis.

Theorem 2.2.7. If V is finite dimensional with $\dim V = n$, then if a list of n vectors spans V, it must be a basis.

Proof. Refine the list to a basis. The basis has n vectors so no vectors were removed. Thus, the list is already a basis.

Example 2.2.8. $U = \{p \in \mathcal{P}_3(\mathbb{R}) : p'(5) = 0\}$, [for $p(x) = \sum_{j=0}^{\infty} a_j x_j$, define $p'(x) = \sum_{j=1}^{\infty} j a_j x^{j-1}$], has $\dim \leq 3$. $1, (x-5)^2, (x-5)^3$ are linearly independent so $\dim U \geq 3$. Thus, $\dim U = 3$.

Theorem 2.2.9. If U_1, U_2 both subspaces of V, $\dim V < \infty$. Then

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

Proof. Find a basis $u_1, \ldots, u_n$ of $U_1 \cap U_2$. Extend to a basis $u_1, \ldots, u_n, v_1, \ldots, v_m$ of U_1 and a basis $u_1, \ldots, u_n, w_1, \ldots, w_k$ of U_2. We claim $u_1, \ldots, u_n, v_1, \ldots, v_n, w_1, \ldots, w_k$ is a basis of $U_1 + U_2$.
First $\forall v \in U_1 + U_2$, $v = u_1 + u_2$ for $u_1 \in U_1$, $u_2 \in U_2$. Consider $u_1 = \sum_{j=1}^{n} a_j u_j + \sum_{j=1}^{m} b_j v_j$, $u_2 = \sum_{j=1}^{n} c_j u_j + \sum_{j=1}^{k} d_j w_j$. Then, $v = u_1 = u_2 + u_1 = \sum_{j=1}^{n} (a_j + c_j) u_j + \sum_{j=1}^{m} b_j v_j + \sum_{j=1}^{k} d_j w_j$. Hence $u_1, \ldots, u_n, v_1, \ldots, v_n, w_1, \ldots, w_k$ spans $U_1 + U_2$.
Moreover, if $\sum_j \alpha_j u_j + \sum_j \beta_j v_j + \sum_j \gamma_j w_j = 0$ for $\alpha_j, \beta_j, \gamma_j \in \mathbb{F}$, then

$$\underbrace{(\sum_j \alpha_j u_j + \sum_j \beta_j v_j)}_{\in U_1} = \underbrace{-\sum_j \gamma_j w_j}_{\in U_2}$$

so both in $U_1 \cap U_2$. So $\sum_j \gamma_j w_j = \sum_j \delta_j u_j$ for $\delta_1, \ldots, \delta_n \in \mathbb{F}$ so $\gamma_1 = \cdots = \gamma_n = 0$. Hence $\sum_j \alpha_j u_j + \sum_j \beta_j v_j = 0$ so all $\alpha_j, \beta_j = 0$. Hence, $u_1, \ldots, u_n, v_1, \ldots, v_m, w_1, \ldots, w_k$ is linearly dependent and the claim holds.
Now, $\dim(U_1 + U_2) = n + m + k$, $\dim U_1 = n + m$, $\dim U_2 = n + k$, $\dim(U_1 \cap U_2) = n$ so theorem follows by a direct computation.

2.3 February 9

2.3.1 3.A- Linear Maps a Vector Space

Theorem 2.3.1. $\text{Hom}(V, W)$ is a vector space with respect to:
$+ : (T_1 + T_2)v = T_1v + T_2v$
$\cdot : (\lambda T_1)v = \lambda \cdot T_1v$

Theorem 2.3.2. If $T \in \text{Hom}(V, W)$, then $T0 = 0$.

Proof. $T0 = T(0 + 0) = T0 + T0$ so $0 = T0$.

Product of linear maps defined by composition

Definition 2.3.3. If $T \in \text{Hom}(U, V)$, $S \in \text{Hom}(V, W)$. Then the product (defined by composition) $ST \in \text{Hom}(U, W)$ is defined as $ST : U \to W$ by $v \mapsto S(Tv)$

Proof that ST is linear.
$(ST)(v_1 + v_2) = S(T(v_1 + v_2)) = S(Tv_1 + Tv_2) = S(Tv_1) + S(Tv_2) = (ST)v_1 + (ST)v_2$
$(ST)(\lambda v) = S(T(\lambda v)) = S(\lambda Tv) = \lambda S(Tv) = \lambda(ST)v$ □

Proposition 2.3.4.

(1) $(T_1T_2)T_3 = T_1(T_2T_3)$ as long as everything is defined

(2) $TI = IT = T$

(3) $(S_1 + S_2)T = S_1T + S_2T$, $S(T_1 + T_2) = ST_1 + ST_2$ as long as everything is defined.

- Assuming $S : U_1 \to U_2$, $T : V_1 \to V_2$ where ST makes sense (ie. $V_2 = U_1$). TS may not make sense
- Even if TS also makes sense (ie. $U_2 = V_1, V_2 = U_1$), $TS : U_1 \to U_1$ but $ST : U_2 \to U_2$
- Even if $U_1 = U_2 = V_1 = V_2$, TS might not equal ST.
  eg. $U_1 = U_2 = V_1 = V_2 = \mathcal{P}(\mathbb{R})$, S : Differentiation, T: multiply by x.
  Then $(ST)(p) = S(T(p)) = S(xp) = p + xp'$ but $(TS)(p) = T(S(p))' = T(p') = xp'$.

Theorem 2.3.5. If $v_1, \ldots, v_m$ is a basis of V and $w_1, \ldots, w_m \in W$ then $\exists!$ linear map $T : V \to W$ such that $Tv_j = w_j$, $1 \leq j \leq n$.

Proof.
Existence: $\forall a_1, \ldots, a_n \in \mathbb{F}$ define $T(\sum a_jv_j) = \sum a_jw_j$
Well defined: only one way to write $\forall v \in V$ as some $\sum a_jv_j$
Linear: For $\lambda \in \mathbb{F}$, $u_1, u_2 \in V$ write $u_1 = \sum_{j=1}^n b_jv_j$, $u_2 = \sum_{j=1}^n c_jv_j$,
$b_j, c_j \in \mathbb{F}$. Then $T(u_1 + u_2) = T(\sum_j(b_j + c_j)v_j) = \sum_j(b_j + c_j)w_j = \sum_j b_jw_j + \sum_j c_jw_j = T(\sum b_jv_j) +$

$T(\sum c_j v_j) = Tu_1 + Tu_2$.
$T(\lambda v_1) = T(\sum_j (\lambda b_j) w_j = \lambda(\sum_j b_j w_j)\lambda T u_1$
Uniqueness: If $T_1 v_j = T_2 v_g = w_j$, $\forall 1 \leq j \leq n$, then $\forall v \in V$, write $v = \sum_{j=1}^n d_j v_j$, $d_j \in \mathbb{F}$, $1 \leq j \leq n$ so $T_1 v = T(\sum d_j v_j) = \sum (T d_j v_j) = \sum d_j T_1(v_j) = \sum d_j w_j$ and $T_2 v = \sum d_j v_j$ for the same reason so $T_1 v = T_2 v$.

2.3.2 3.B - Kernels and Images

Definition 2.3.6. For $T \in \text{Hom}(V, W)$, the kernel (or null space) of T is $\ker T = \{v \in V : Tv = 0\}$.

Example 2.3.7.

(1) $0 : V \to W \quad \ker 0 = V$

(2) If $V \subset W$, $i : V \to W \quad \ker i = \{0\}$

(3) $D : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$, char $F = 0 \quad \ker D = \{\text{constants}\}$

Proposition 2.3.8. $\forall T \in \text{Hom}(V, W)$, $\ker T$ is a subspace

Definition 2.3.9. A map $f : S_1 \to S_2$ is called injective if $f(x_1) = f(x_2) \to x_1 = x_2$.

Proposition 2.3.10. If $T \in \text{Hom}(V, W)$, then T is injective iff $\ker T = \{0\}$

Proof. $\to$) $0 \in \ker T$. By injectivity, nothing else is mapped to 0.
$\leftarrow$) If $Tv_1 = Tv_2$, then $T(v_1 - v_2) = 0$. Thus with $\ker T = \{0\}$ implies that $v_1 - v_2 = 0$ so $v_1 = v_2$

Definition 2.3.11. If $T \in \text{Hom}(V, W)$, then image (or range) of T is defined as $\text{im}T = \{w \in W : \exists v \in V \text{ such that } w = Tv\}$

Example 2.3.12.

(1) $\text{im}0 = \{0\}$

(2) $V \subset W$, $i : V \to W$ has image V

(3) $D : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$, char $\mathbb{F} = 0 \quad \text{im}D = \mathcal{P}(\mathbb{F})$

Proposition 2.3.13. $\forall T \in \text{Hom}(V, W)$, $\text{im}T$ is a subspace.

Proof. $\forall w_1, w_2 \in \text{im}T$, find $v_1, v_2 \in V$ such that $Tv_1 = w_1$, $Tv_2 = w_2$. Then $T(v_1 + v_2) = w_1 + w_2$, $T(\lambda v_1) = \lambda w_1$.

Definition 2.3.14. A map $f : S_1 \to S_2$ is surjective if $\{f(s) : s \in S_1\} = S_2$.

Observation: $\forall T \in \text{Hom}(V, W)$, T is sujective iff $\text{im}T = W$

Theorem 2.3.15 (Fundamental Theorem of Linear Maps)**.** Assume V is finite dimensional and $T \in \text{Hom}(V, W)$, then $\dim V = \dim(\text{im}T) + \dim(\ker T)$

Proof. If $v_1, \dots, v_n$ is a basis of $\ker T$, extend it to a basis $v_1, \dots, v_n, v_{n+1}, \dots, v_m$ of V. We claim: $Tv_{n+1}, \dots, Tv_m$ is a basis of $\text{im}T$.
Spans: $\forall w \in \text{im}T$, $\exists v \in V$ such that $Tv = w$. Write $v = \sum_{j=1}^m a_j v_j$. Then $Tv = \sum_{j=1}^m a_j Tv_j = \sum_{n<j\leq m} a_j Tv_j$. Hence $Tv_{n+1}, \dots, Tv_m$ spans $\text{im}T$.
Lin. Indpendent: If $b_{n+1}, \dots, b_m \in \mathbb{F}$ such that $b_{n+1}Tv_{n+1} + \cdots + b_m Tv_n$. Then $T(\sum_{n<j\leq m} b_j v_j) = 0$ so $\sum_{n<j\leq m} b_j v_j \in \ker T$. So $\exists a_1, \dots, a_n$ such that $\sum_{n<j\leq m} = \sum_{j=1}^n c_j v_j$ so all $b_j = 0$.
Hence the claim is verified. Thus, $\dim V = m, \dim(\ker V) = n, \dim(\text{im}V) = m - n$.

2.4 February 14

2.4.1 3.B - Kernels and Images

Corollary 2.4.1. If $\dim V > \dim W$, then no $T \in \text{Hom}(V, W)$ is injective.

Corollary 2.4.2. If $\dim V < \dim W$, then no $T \in \text{Hom}(V, W)$ is surjective.

Corollary 2.4.3. $D : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ is not surjective

Theorem 2.4.4. A homogeneous system of linear equations

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ \cdots \\ f_m(x_1, \dots, x_n) = 0 \end{cases} \qquad \text{where } f_j(x_1, \dots, x_n) = \sum_{j=1}^n A_{j,k} x_k$$

with more variables than equations has a nonzero solution.

Proof. Construct a linear map $T : \mathbb{F}^n \to \mathbb{F}^m$ by $(x_1, \dots, x_n) \mapsto (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$. Then, $\dim \ker T = \dim \mathbb{F}^n - \dim \text{im}T \geq n - m \geq 1$. Take a nonzero element in the kernel and that is a nonzero solution.

Theorem 2.4.5. An inhomogenous system of linear equations

$$\begin{cases} f_1(x_1, \dots, x_n) = a_1 \\ \cdots \\ f_m(x_1, \dots, x_n) = a_m \end{cases} \qquad \text{where } f_j(x_1, \dots, x_n) = \sum_{j=1}^{n} A_{j,k} x_k$$

with more equations than variables has no solutions for some choice of constant terms.

Proof. Define T as in the proof above. Then T is not going to be surjective so there exists $(a_1, \dots, a_n)$ not in the image of T so take that vector as the choice of constants.

2.4.2 3.C - Matrices

A linear map can be represented by a matrix.

Definition 2.4.6. An $m \times n$ matrix is an array of scalars in the from

$$A = \underbrace{\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ A_{2,1} & \cdots & A_{2,n} \\ \cdots & \cdots & \cdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}}_{n \text{ columns}} \left.\vphantom{\begin{pmatrix} A \\ A \\ A \\ A \end{pmatrix}}\right\} m \text{ rows}$$

Also written as $(A_{i,j})_{m \times n}$. $\mathbb{F}^{m,n} = \{\text{all } m \times n \text{ matrices}\}$.

Definition 2.4.7 (Matrix of a Linear Map)**.** If $T \in \text{Hom}(V, W)$, $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_n$ is a basis of W. Assume $Tv_k = \sum_{j=1}^{m} A_{j,k} v_j$. Then $(A_{j,k})_{m \times n}$ is called the matrix of T with respect to the bases $(v_1, \dots, v_n)$ and $(w_1, \dots, w_m)$, denoted by $\mathcal{M}(T)$.

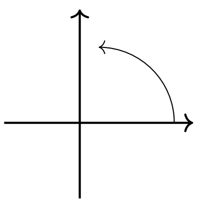
Digest:

$$\begin{array}{cc} & \begin{matrix} \mathrm{v}_1 & \cdots & \mathrm{v}_\mathrm{n} \end{matrix} \\ \begin{matrix} w_1 \\ \vdots \\ w_m \end{matrix} & \begin{pmatrix} A_{1,1} & \cdots & \\ \vdots & & \vdots \\ A_{1,n} & \cdots & \end{pmatrix} \end{array} \qquad \begin{aligned} \text{columns} &\leftrightarrow \text{element in basis of domain} \\ \text{rows} &\leftrightarrow \text{element in basis of target space} \end{aligned}$$

Motivation: Matrix Multiplication

Example 2.4.8. In $\mathbb{R}^2$

(a) Rotation about 0.

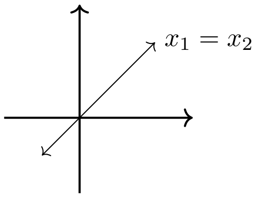


Rotate by $\frac{\pi}{2}$ counterclockwise.

Matrix with respect to (e_1, e_2) is $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

More generally, rotation by θ with respect to (e_1, e_2) is $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

(b) Orthogonal projection to L but then included into $\mathbb{R}^2$.



Matrix with respect to (e_1, e_2): $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$

Matrix with respect to $((1,1),(1,-1))$: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

(c) $i : V \to W$ (assume $V \subset W$) with respect to $(v_1, \ldots, v_n)$, $(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$.

$$\mathcal{M}(i) = \begin{pmatrix} 1 & 0 & & 0 \\ 0 & 1 & & \vdots \\ \vdots & 0 & \ddots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \end{pmatrix} \leftarrow n\text{th row}$$

Definition 2.4.9. If $A, B \in \mathbb{F}^{m,n}$, $\lambda \in \mathbb{F}$, $A + B$, λA are defined as entrywise addition and scalar multiplication.

Proposition 2.4.10. If $T_1, T_2 \in \mathrm{Hom}(V, W)$. Fix a basis of V and a basis of W. Then $\mathcal{M}(T_1 + T_2) = \mathcal{M}(T_1) + \mathcal{M}(T_2)$, $\mathcal{M}(\lambda T_1) = \lambda \mathcal{M}(T_1)$.

Proposition 2.4.11. $\mathbb{F}^{m,n}$ is a vector space with dimension mn.

Proof. The list of all possible $m \times n$ matrices with 0 in all entries except one (where the entry is 1) form a basis.

2.4.3 Matrix Multiplication

- Motivated by looking for matrix of ST.

Definition 2.4.12. For $A \in \mathbb{F}^{m,n}$, $B \in \mathbb{F}^{n,p}$, define $AB \in \mathbb{F}^{m,p}$ such that $(AB)_{i,k} = \sum_{j=1}^{n} A_{i,j} B_{j,k}$.

Proposition 2.4.13. If $T \in \text{Hom}(V, W)$, $S \in \text{Hom}(V, W)$, $u_1, \dots, u_p$ is a basis of U, $v_1, \dots, v_n$ is a basis of V, and $w_1, \dots, w_m$ is a basis of W, then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

Proof. Assume $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$. $\forall k \in \{1, \dots, p\}$

$$\begin{aligned}
(ST)u_k &= S(Tu_k) \\
&= S(\sum_{j=1}^{n} B_{j,k} v_j) \\
&= \sum_{j=1}^{n} B_{j,k}(Sv_j) \\
&= \sum_{j=1}^{n} B_{j,k}(\sum_{i=1}^{m} A_{i,j} w_i) \\
&= \sum_{i=1}^{m}(\sum_{j=1}^{n} A_{i,j} B_{j,k}) w_i
\end{aligned}$$

Hence $(\mathcal{M}(S(T))_{i,k} = \sum_{j=1}^{m} A_{i,j} B_{j,k} = (AB)_{j,k}$.

Example 2.4.14. $\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix} = \begin{pmatrix} 6 & 7 \\ 26 & 31 \end{pmatrix}$

Proposition 2.4.15. $(AB)_{i,j} =$(ith row of A) $\cdot$ (jth column of B), here "$\cdot$" is the dot product.

Proposition 2.4.16. The jth column of $AB = A$(jth column of B).

Proposition 2.4.17. If $A \in \mathbb{F}^{m,n}$, $c \in \mathbb{F}^{n,1} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, then Ac is a linear combination of the columns of A: $Ac = c_1 A_{\cdot,1} + \dots + c_n A_{\cdot,n}$.

2.5 February 23

2.5.1 3.D- Invertibility and Isomorphic Vector Spaces

Definition 2.5.1. $T \in \text{Hom}(V, W)$ is called invertible if $\exists S \in \text{Hom}(W, V)$ such that $ST = I, TS = I$. Such an S is called an inverse of T.

Proposition 2.5.2. If T has an inverse, then the inverse is unique.

Proof. If T_1, T_2 are inverses, $T_2 = T_2 S T_1 = T_1$.

We use T^{-1} to denote the inverse of T.

Theorem 2.5.3. A linear map T is invertible iff it is injective and surjective.

Proof. $\to$) True by set theory.
$\leftarrow$) T has a set theoretic inverse S. It suffices to show S is linear.
Assume $T \in \mathrm{Hom}(V, W)$, $\forall w_1, w_2 \in W$, $\forall \lambda \in \mathbb{F}$, there is v_1, v_2 such that $Tv_1 = w_1$, $Tv_2 = w_2$. Then $T(v_1 + v_2) = w_1 + w_2$ so $S(w_1 + w_2) = w_1 + w_2 = Sw_1 + Sw_2$. Similarly, $S(\lambda w_1) = \lambda S w_1$.

Example 2.5.4.

(1) Multiplication by $(x+1)$ is not invertible (viewed as map from $\mathcal{P}(\mathbb{F})$ to itself)

(2) Multiplication by $(x+1)$ and discarding terms of deg $> n$ is invertible in $\mathrm{End}(\mathcal{P}_n(\mathbb{F}))$.

Definition 2.5.5. An invertible $T \in \mathrm{Hom}(V, W)$ is called an isomorphism. If such a T exists, we say V and W are isomorphic and write $V \cong W$.

Example 2.5.6. $T : \mathcal{P}(\mathbb{F}) \to \mathbb{F}^3$ by $a_0 + a_1 x + a_2 x^2 \mapsto (a_0, a_1, a_2)$ is an isomorphism.

Theorem 2.5.7. If V and W are finite dimensional vector spaces, then $V \cong W \leftrightarrow \dim V = \dim W$.

Proof. $\to$) If $V \cong W$, then there is an isomorphism $T : V \to W$ so $\dim V = \dim \mathrm{im} T = \dim \ker T = \dim W + 0 = \dim W$.
$\leftarrow$) Take bases $v_1, \ldots, v_n$ of V and $w_1, \ldots w_n$ of W. Then $T \in \mathrm{Hom}(V, W)$ is an isomorphism if $Tv_j = w_j$, $\forall j$.

Corollary 2.5.8. If $\dim V = n$, then $V \cong \mathbb{F}^n$.

Theorem 2.5.9. If $\dim V = n, \dim W = m$ with bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$, respectively. Then $\mathcal{M} : \mathrm{Hom}(V, W) \to \mathbb{F}^{m,n}$ is an isomorphism.

Proof. We have already shown $\mathcal{M}$ is linear.
If $\mathcal{M}(T) = 0$, then $Tv_j = 0\ \forall j$ so $T = 0$ so $\mathcal{M}$ is injective.
For $A \in \mathbb{F}^{m,n}$, define $T \in \mathrm{Hom}(V, W)$ such that $Tv_k = \sum_{j=1}^{m} A_{j,k} w_j$, then $\mathcal{M}(T) = A$. So $\mathcal{M}$ is surjective.

Corollary 2.5.10. $\dim \mathrm{Hom}(V, W) = (\dim V)(\dim W)$, if $\dim V, \dim W$ are finite.

We can think of linear maps as matrix multiplication.

Definition 2.5.11. For a basis $v_1, \ldots, v_n$ of V, the matrix of $v = \sum_j a_j v_j$, $a_j \in \mathbb{F}$ is $\begin{pmatrix} a_1 \\ \vdots \end{pmatrix}$, denoted $\mathcal{M}(v)$.

Example 2.5.12. In $\mathbb{F}^n$, with respect to $e_1, \ldots, e_n$, the matrix of $(a_1, \ldots, a_n)$ is $\begin{pmatrix} a_1 \\ \vdots \end{pmatrix}$.

Proposition 2.5.13. $\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$

Proposition 2.5.14. $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$.

2.6 February 28

2.6.1 3.D - Invertibility and Isomorphic Vector Spaces

Proposition 2.6.1. $\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$

Proof. Let $v = \sum_{k=1}^n a_k v_k$. Assume $\mathcal{M}(T) = A$. Then

$$\begin{aligned} Tv &= \sum_{k=1}^n a_k \sum_{j=1}^m A_{j,k} w_j \\ &= \sum_{j=1}^m (\sum_{k=1}^n A_{j,k} a_k) w_j \end{aligned}$$

so the jth entry of a linear map can be thought of as a matrix multiplication.

Example 2.6.2. $\mathcal{M}(\cdot)$ is an isomorphism from V to $\mathbb{F}^{n,1}$. Recall $F^{n,1}$ is canonically {linear maps from $\mathbb{F}$ to $\mathbb{F}^n$}. What is $\mathcal{M}(\cdot)$ in this context?

Solution. $M(\cdot)$ is the linear map which maps $a_1 v_1 + \cdots + a_n v_n$ to $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ or equivalently, the linear map from $\mathbb{F}$ to $\mathbb{F}^n$ which sends 1 to $(a_1, \ldots, a_n)$. □

Operators: $T \in \mathrm{End}(V)$ (or $\mathcal{L}(V)$ in the book), also called linear transformations.

Theorem 2.6.3. If V is finite dimensional, $T \in \mathrm{End}(V)$, then T is invertible $\iff$ T is injective $\iff$ T is surjective.

Proof. Since $\dim V = \dim(\ker T) + \dim(\mathrm{im} T)$, the theorem follows from the fact that $\dim(\ker T) = 0$ iff T is injective and $\dim(\mathrm{im} T) = \dim V$ iff T is surjective.

Example 2.6.4. Find a counterexample $T \in \mathrm{End}(V)$ such that

(1) T is injective but not surjective.

(2) T is surjective but not injective.

Solution.

(1) Consider $T \in \mathrm{End}(\mathbb{R}^\infty)$ defined by $T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$.

(2) Consider $T \in \mathrm{End}(\mathbb{R}^\infty)$ defined by $T(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots)$.

□

2.6.2 3.E- Products and Quotient Spaces

Definition 2.6.5. For vector spaces $V_1, \dots, V_m/\mathbb{F}$, the product $V_1 \times \cdots \times V_m$ is defined as $V_1 \times \cdots \times V_m = \{(v_1, \dots, v_m) : v_j \in V_j, 1 \le j \le m\}$.

Proposition 2.6.6. $V_1 \times V_2 \times \cdots \times V_n$ is a vector space $/\mathbb{F}$ with respect to:

$$\begin{aligned} \text{addition}: &\quad (v_1, \dots, v_m) + (u_1, \dots, u_m) = (v_1 + u_1, \dots, u_m + v_m) \\ \text{scalar multiplication}: &\quad \lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m) \end{aligned}$$

Example 2.6.7. $\mathbb{R} \times \mathbb{R} \cong \mathbb{R}^2, \quad \mathbb{R}^2 \times \mathbb{R}^3 \cong \mathbb{R}^5$

Proposition 2.6.8. If $V_1, \dots, V_m$ are finite dimensional, then $\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m$.

Proof. For each V_j, choose a basis $v_{j,1}, \dots, v_{j,d_j}$ where $d_j = \dim V_j$. Then $(v_{1,1}, 0, \dots, 0), (v_{1,2}, 0, \dots, 0), \dots, (v_{1,d_1}, 0, \dots, 0)$ $(0, v_{2,d_2}, 0, \dots, 0), (0, \dots, 0, v_{m,1}), \dots, (0, \dots, 0, v_{m,d_m})$ is a basis of $V_1 \times \cdots \times V_m$. Hence $\dim(V_1 \times \cdots \times V_m) = d_1 + d_2 + \cdots + d_m$.

Theorem 2.6.9. If $U_1, \dots, U_m$ are subspaces of V, then

(1) $\Gamma : U_1 \times \cdots \times U_m \to U_1 + \cdots + U_m$ by $(u_1, \dots, u_m) \mapsto u_1 + \cdots + u_m$ is a linear map and surjective. Moreover, Γ is injective iff $U_1 + \cdots + U_m$ is a direct sum.

(2) If V is finite dimensional, then $U_1 + \cdots + U_m$ is a direct sum iff $\dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m$.

Proof.

(1) For injectivity: $U_1 + \cdots + U_m$ is a direct sum $\leftrightarrow \forall v \in U_1 + \cdots + U_m$, $\exists!$ way to represent v as a sum of $u_1 + \cdots + u_m \leftrightarrow \Gamma$ is injective (and surjective).

(2) By surjectivity of Γ, $\dim(U_1 \times \cdots \times U_m) = \dim\ker\Gamma + \dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m$. Note that $\dim(\ker\Gamma)) = 0$ iff Γ is injective iff $U_1 + \cdots + U_m$ is a direct sum.

2.6.3 Quotient Spaces

Motivation:

1. How to define a "3rd dimension" in $\mathbb{R}^3$ if we have "2 defined dimensions"

2. To construct new vector spaces from a known vector space.

Definition 2.6.10. If $U \subset V$ is a subspace, $x \in V$, define the affine subset $x + U = \{x + u : u \in U\}$. We say that $x + U$ is parallel to U.

Example 2.6.11. $V = \mathbb{R}^3$, $U =$"the plane of the floor".
Many affine subsets are the same.

Example 2.6.12. Any two affine subsets are identical or disjoint.

Definition 2.6.13. If $U \subset V$ is a subspace, then $V/U = \{$all lines in V parrallel to $U\}$.

Theorem 2.6.14. V/U is a vector space with respect to:
$+ : \quad (v_1 + U) + (v_2 + U) = (v_1 + v_2) + U$
$\cdot : \quad \lambda(v + U) = \lambda v + U$

Proof. We first prove a lemma.
Lemma 2.6.15. $v + U = \tilde{v} + U$ iff $v - \tilde{v} \in U$.

Proof. $\to)$ $v \in \tilde{v} + U$. Hence $\exists v \in U$ such that $v = \tilde{v} + U$, so $v - \tilde{v} = u \in U$.
$\leftarrow)$ $\forall x \in v + U$, take $u_1 \in U$ such that $x + u_1$. Then $x = \tilde{v} + (v - \tilde{v} + u_1)$. Hence $x \in \tilde{v} + U$. Hence $v + U \subset \tilde{v} + U$.
By an entirely similar argument $\tilde{v} + U \subset v + U$. Hence the lemma holds.

Now, we preove "+", "·" are well defined. If $v_1 + U = \tilde{v}_1 + U$, $v_2 + U = \tilde{v}_2 + U$, then $v_1 - \tilde{v}_1 \in U$, $v_2 - \tilde{v}_2 \in U$. Hence $v_1 + v_2 - \tilde{v}_1 - \tilde{v}_2 \in U$ so $(v_1 + v_2) + U = (\tilde{v}_1 + \tilde{v}_2) + U$ so $+$ is well defined.
Similarly, $\forall \lambda \in \mathbb{F}$, v_1, $\tilde{v}_1$ as above, $v_1 - \tilde{v}_1 \in U$ so $\lambda(v_1 - \tilde{v}_1) \in U$ so $\lambda v_1 + U = \lambda \tilde{v}_1 + U$.
Now V/U is a vector space as properties "carried down" from V to the quotient space.

Alternate way to construct quotient space: Use equivalence classes, $v_1 - v_2 \in V \leftrightarrow v_1 \sim v_2$. Quotient space $V/\sim$ defines the set theoretic quotient.

Definition 2.6.16. For $U \subset V$ subspace, define the quotient map

$$\pi : V \to V/U \text{ by } v \mapsto v + U.$$

π is linear.

2.7 March 2

2.7.1 3.E - Product and Quotient Spaces

Set Theoretic Quotient: Set + "equivalence relation" $\rightarrow$ quotients

Given a set S- an equivalence relation on S is "$\sim$" (binary relations) such that:

(1) $x \sim x$, $\forall x \in S$

(2) $x \sim y \rightarrow y \sim x$

(3) $x \sim y$, $y \sim z \rightarrow x \sim z$

Given S and equivalence relation "$\sim$", natural $S = \sqcup$ of "equivalence classes." such that $x_1 \sim x_2$ x_1 and x_2 are in the same class.

Example 2.7.1. {people with permanent addresses in US}: For $p_1, p_2 \in S$, $p_1 \sim p_2$ iff their permanent address is in the same state.

Quotient Set $(S/\sim) = \{$all equivalence classes$\}$
In above example: $S/\sim = \{\{$people with permanent addresses in CA$\}, \{\ldots$ in WI$\}, \{\ldots$ in NJ $\}, \ldots\}$.

If $f : S \rightarrow X$, there is a natural equivalence relation in S "$\sim_f$" defined as: "For $x_1, x_1 \in S$, $x_1 \sim_f x_2$ off $f(x_1) = f(x_2)$." Forms the quotient $S/\sim_f$ where $S/\sim_f \cong \text{im} f$.

Example 2.7.2. Is $S/\sim$ isomorphic to a subset of S? (yes)
Is $S/\sim$ canonically isomorphic to a subset of S? (no)

Quotient also makes sense in quotient space context: $U \subset V$ subspace, then V/U is the set theoretic quotient with respect to the $\sim$ (defined by "$x_1 \sim x_2$ iff $x_1 - x_2 \in U$"). V/U is a vector space.

Theorem 2.7.3. With π as in definition 2.6.16, $\ker \pi = U$, $\dim(V/U) = \dim V - \dim U$ if $\dim U, \dim V < \infty$.

Proof. $\ker \pi \{v : v + U = 0 + U\} = \{v : v - 0 \in U\} = U$. Second claim follows.

Definition 2.7.4. If $T \in \text{Hom}(V, W)$, define the induced map $\tilde{T} : V/\ker T \rightarrow W$ by $v + \ker T \mapsto Tv$.

Note: $\tilde{T}$ is well defined since if $v_1 + \ker T = v_2 + \ker T$, $v_1 - v_2 \in \ker T$ so $Tv_1 - Tv_2 = 0$ so $Tv_1 = Tv_2$. Also, $\tilde{T}$ is linear.

Theorem 2.7.5.

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

(2) $\tilde{T}$ is an isomorphism from $V/\ker T$ to $\text{im}T$.

Proof.

(1) $\text{im}\tilde{T} = \{Tv : v \in V\} = \text{im}T$

(2) Surjective by (1)
For injectivity, if $\tilde{T}(v + \ker T) = 0$, then $Tv = 0$ so $v \in \ker T$ so $v + \ker T = 0 + \ker T$.

Example 2.7.6. Is V/U isomorphic to a subspace of V? (yes if V is finite dimensional)
If V/U canonically isomorphic to a subspace of V? (no in general)

2.7.2 3.F - Duality and Rank

Definition 2.7.7. A linear functional (or linear function) on V is a map in $\mathrm{Hom}(V, \mathbb{F})$. Also denoted as V^*.

Example 2.7.8. $(\mathbb{F}^3)^*$ contains all functions of the form $(x_1, x_2, x_3) \mapsto a_1x_1 + a_2x_2 + a_3x_3$, $a_j \in \mathbb{F}$.
V^* called the dual space of V.

Theorem 2.7.9. $\dim V^* = \dim V$ if $\dim V < \infty$

Example 2.7.10. Is there a canonical map between V and V^*?

Definition 2.7.11. For a basis $v_1, \ldots, v_n$ of V (if $\dim V < \infty$) define the dual basis $\varphi_1, \ldots, \varphi_n$ in V^* as:

$$\varphi_i(v_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

This is indeed a basis of V^* (linearly independent, right length)

Caution: "Globally defined"- need to know $v_1, \ldots, v_n$ to define φ_i.

Definition 2.7.12. If $T \in \mathrm{Hom}(V, W)$, define the dual map $T^* \in \mathrm{Hom}(W^*, V^*)$ by $T^*(\varphi)(v) = \varphi(Tv)$.

Example 2.7.13. Let $D \in \mathrm{End}(\mathcal{P}(\mathbb{R}))$ be the differentiation map, $\forall$ linear functions on $\mathcal{P}(\mathbb{R})$, $(D^*\varphi)(f) = \varphi(Df)$ eg. first differentiate f then act on by φ.

Proposition 2.7.14.

(1) If $S, T \in \mathrm{Hom}(v, W)$, $(S + T)^* = S^* + T^*$. If $\lambda \in \mathbb{F}, (\lambda S)^* = \lambda S^*$.

(2) If $S \in \mathrm{Hom}(V, W)$, $T \in \mathrm{Hom}(V, W)$, then $(ST)^* = T^* \circ S^*$.

Proof.

(1) $\forall \varphi \in W^*$, $(ST)^*(\varphi) = \varphi \circ (ST) = \varphi \circ S \circ T = (\varphi \circ S) \circ T = T^*(\varphi \circ S) = T^*S^*\varphi$.

What is $\ker T^*$? $\mathrm{im} T^*$?
$\varphi \in \ker T^* \leftrightarrow \varphi(Tv) = 0 \forall v \in V \leftrightarrow \varphi(0)$ on $\mathrm{im} T$ (aka. $\forall w \in \mathrm{im} T, \varphi(w) = 0$)

Definition 2.7.15. $\forall U \subset V$ subspace, the annihilator of U is defined as $U^0 = \{\varphi \in V^* : \varphi = 0 \text{ on } U\}$.

Example 2.7.16. If $U \subset \mathcal{P}(\mathbb{R})$ defined as $\{cx : c \in \mathbb{F}\}$. Then $\varphi : p \mapsto p'(0)$ is in U^0.

Proposition 2.7.17.

(1) U^0 is a subspace of V^*.

(2) If $\dim V < \infty$, then $\dim U^0 = \dim V - \dim U$.

Proof. (1) $\forall \varphi_1, \varphi_2 \in V^*$, $\lambda, \mathbb{F}$, $\forall u \in U$, $(\varphi_1 + \varphi_2)(u) = \varphi_1(u) + \varphi_2(u) = 0 + 0 = 0$ and $(\lambda\varphi_1)(u) = \lambda\varphi_1(u) = \lambda \cdot 0 = 0$.

(2) Consider the inclusion $i : U \to V$. i^* is a restriction of φ to U. $\ker i^* = \{\varphi \in V^* : \varphi = 0 \text{ on } U\} = U^0$. Also, $\mathrm{im} i^* = U^*$. Hence, thm follows since $\dim V^* = \dim \ker i^* + \dim \mathrm{im} i^*$.
Alternate Solution: Choose a basis of U and extend to a basis of V then consider the dual basis.

Theorem 2.7.18.

(a) $\ker T^* = (\mathrm{im} T)^0$ if $T \in \mathrm{Hom}(V, W)$

(b) $\dim \ker T^* = \dim \ker T + \dim W - \dim V$. If $\dim V, \dim W < \infty$.

Proof.

(a) by previous discussion.

(b) $\dim(\ker T^*) = \dim(\mathrm{im} T)^0 = \dim W - \dim(\mathrm{im} T) = \dim W - (\dim V - \dim \ker T)$.

2.8 March 7

2.8.1 3.F- Duality and Rank

Q1: What does canonical mean?
There is a unique choice that is much better than every other choice.
Canonical isomorphism: {States of US} $\leftrightarrow$ {Quarters with states on them}
Non-canonical isomorphism: {Everyone in Class I of 100 people} $\leftrightarrow$ {Everyone in Class II of 100 people}

Q2: Why $\mathrm{Hom}(V, \mathbb{F})$, not $\mathrm{Hom}(\mathbb{F}, V)$?
$\mathrm{Hom}(\mathbb{F}, V)$ is canonically isomorphic to V.

Corollary 2.8.1. If $\dim V, \dim W < \infty$, $T \in \mathrm{Hom}(V, W)$. Then T is surjective $\leftrightarrow$ T^* is injective.

Proof. T surjective $\leftrightarrow \dim(\mathrm{im} T) = \dim W \leftrightarrow \dim \ker T + \dim W - \dim V = 0 \leftrightarrow \dim \ker T^* = 0$

Theorem 2.8.2. If $\dim V, \dim W < \infty$, $T \in \mathrm{Hom}(V, W)$, then

(a) $\dim(\mathrm{im} T^*) = \dim(\mathrm{im} T)$

(b) $\text{im}T^* = (\ker T)^0$

Proof.

(a) $\dim(\text{im}T^*) = \dim W^* - \dim\ker T^* = \dim W - (\dim W + \dim\ker T - \dim V) = \dim V - \dim\ker T = \dim \text{im}T$.

(b) $\psi \in \text{im}T^* \leftrightarrow \psi = \varphi \circ T, \varphi \in W^* \rightarrow \psi(v) = 0, \forall v \in \ker T \leftrightarrow \psi \in (\ker T)^0$.
Hence, $\text{im}T^* \subset (\ker T)^0$, but $\dim(\text{im}T^*) = \dim(\text{im}T)$, so $\dim(\ker T)^0 = \dim V - \dim\ker T = \dim(\text{im}T) = \dim(\text{im}T^*)$ so $\text{im}T^* = (\ker T)^0$

Corollary 2.8.3. If $T \in \text{Hom}(V, W)$, $\dim V, \dim W < \infty$, then T is injective $\leftrightarrow$ T^* is surjective.

Proof. T^* is surjective $\leftrightarrow \dim \text{im}T^* = \dim V^* \leftrightarrow \dim \text{im}T = \dim V \leftrightarrow \dim\ker T = 0 \leftrightarrow T$ is injective.

Definition 2.8.4. For $A \in \mathbb{F}^{m,n}$, define its transpose $A^t \in F^{n,m}$ by $(A^t)_{i,j} = A_{j,i}$.

Example 2.8.5. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

Theorem 2.8.6. If $T \in \text{Hom}(V, W)$, $\dim V, \dim W < \infty$, take a basis $v_1, \dots, v_n$ of V, its dual basis $\varphi_1, \dots, \varphi_n$, and $w_1, \dots, w_m$ of W, its dual basis $\psi_1, \dots, \psi_m$. Then $\mathcal{M}(T^*) = \mathcal{M}(T)^t$.

Proof. $(T^*\psi_j)(v_k) = \psi_j(Tv_k) = \mathcal{M}(T)_{j,k}$ so $T^*\psi_j = \sum_{k=1}^n \mathcal{M}(T)_{j,k}\varphi_k$ so $\mathcal{M}(T^*)_{k,j} = \mathcal{M}(T)_{j,k}$

Theorem 2.8.7. If $A \in \mathbb{F}^{m,n}, B \in \mathbb{F}^{n,p}$ then $(AB)^t = B^t A^t$.

Proof (Proof 1)**.** Direct Computation.

Proof (Proof 2)**.** View A canonically in $\text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$ (w.r.t standard basis) and B in $\text{Hom}(\mathbb{F}^p, \mathbb{F}^n)$. Follows from $(AB)^* = B^* A^*$.

Definition 2.8.8. IF $A \in \mathbb{F}^{m,n}$, the row rank of $A = \dim\text{span}\{$ rows of $A\}$, the column rank of $A = \dim\text{span}\{$ columns of $A\}$.

Theorem 2.8.9. $\dim(\text{im}T) =$ column rank of $\mathcal{M}(T)$ w.r.t any basis. (Assuming $\dim V, \dim W < \infty, T \in \text{Hom}(V, W)$)

Proof. Take a bases $v_1, \dots, v_m$, $w_1, \dots, w_m$. Columns $\mathcal{M}(T)$ are the coefficients of the expression of Tv_j

$(1 \leq j \leq n)$ into $w_k's$, equivalent to the matrices of $Tv_1, \ldots, Tv_n$ so span of columns $\cong \mathrm{im}T$. □

Theorem 2.8.10. For $A \in \mathbb{F}^{m,n}$, row rank of $A =$ column rank of A.

Proof. View A in $\mathrm{Hom}(\mathbb{F}^n, \mathbb{F}^m)$ canonically. Then RHS $= \mathrm{im}A$, LHS $= \dim \mathrm{im}A^*$ so thm follows.

Definition 2.8.11. For $A \in \mathbb{F}^{m,n}$, rank $A =$ row rank of A. $T : V \to W$, $\dim V, \dim W < \infty$, rank $T =$ rank $\mathcal{M}(T)$.

Chapter 3

Polynomials

3.1 March 7

3.1.1 Ch4 - Polynomials

3.1.2 More on Complex Numbers

Definition 3.1.1. For $z = a + bi \in \mathbb{C}$, with $a, b \in \mathbb{R}$, the real part of z is a ($\operatorname{Re} z = a$), and the imaginary part of z is b ($\operatorname{Im} z = b$).
The norm/absolute value of $z = |z| = \sqrt{a^2 + b^2}$.
The complex conjugate of z is $\overline{z} = a - bi$.

Theorem 3.1.2.

(1) $z \mapsto \overline{z}$ is a field automorphism of $\mathbb{C}$. ie. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$.

(2) $z + \overline{z} = 2\operatorname{Re} z$.

(3) $\frac{z - \overline{z}}{i} = 2\operatorname{Im} z$.

(4) $z - \overline{z} = |z|^2$.

(5) $\overline{\overline{z}} = z$.

(6) $|\operatorname{Re} z|, |\operatorname{Im} z| \leq |z|$.

(7) $|z_1 z_2| = |z_1||z_2|$
(use $(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) = \cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)$).

(8) $|z_1 + z_2| \leq |z_1| + |z_2|$ (triangle inequality)

Proof.

$$\begin{aligned}|z_1+z_2|^2 &= (z_1+z_2)(\overline{z}_1+\overline{z}_2)\\ &= |z_1|^2+|z_2|^2+\operatorname{Re} z\\ &\le |z_1|^2+|z_2|^2+2|z_1||z_2|\\ &= (|z_1|+|z_2|)^2\end{aligned}$$

3.1.3 Polynomials

Commutative: $p_1, p_2 \in \mathcal{P}(\mathbb{F})$, $p_1+p_2=p_2+p_1$, $p_1p_2=p_2p_1$.
The division algorithm: Assume $p, s \in \mathcal{P}(\mathbb{F})$, $s \neq 0$, then $\exists!$ pair $(q,r) \in \mathcal{P}(\mathbb{F})$ such that $sq+r$ and $\deg r < \deg s$.
Example 3.1.3. $(\underbrace{x^4+2x^3+3x^2+4x+5}_{p}) = (\underbrace{x^2+x+1}_{q})(\underbrace{x^2+x+1}_{s}) + \underbrace{2x+4}_{r}$

Definition 3.1.4. $\lambda \in \mathbb{F}$ is called a zero (or a root) of $p \in \mathcal{P}(\mathbb{F})$ if $p(\lambda)=0$.

Definition 3.1.5. If $p, s \in \mathcal{P}(\mathbb{F})$, $s \neq 0$, s is called a factor of p if $\exists q \in \mathcal{P}(\mathbb{F})$ such that $p = qs$.

Theorem 3.1.6. A polynomial $p \neq 0$ of degree m has $\le m$ distinct roots.

Theorem 3.1.7 (Fundamental Theorem of Algebra)**.** Given $p \neq 0$, $p \in \mathcal{P}(\mathbb{C})$, $\deg p = m$, $\exists c \neq 0$, $\lambda_1, \dots, \lambda_m \in \mathbb{C}$, unique up to permutation of $\lambda_1, \dots, \lambda_m$ such that $p(z) = c(z-\lambda_1)\cdots(z-\lambda_m)$.

Theorem 3.1.8. If $p \in \mathcal{P}(\mathbb{C})$ with real real coefficients, then $\lambda \in \mathbb{C}$ is a root of $p \leftrightarrow \overline{\lambda} \in \mathbb{C}$ is a root of p.

Theorem 3.1.9. Given $p \in \mathcal{P}(\mathbb{R})$, $p \neq 0$, $\deg p = n$, then $\exists m, M > 0$ such that $m + 2M = n$, $\exists \lambda_1, \dots, \lambda_m$, $b_1, c_1, \dots, b_M, c_M$ unique up to permutation of λ's, (b,c)'s such that $p(x) = c(x-\lambda_1)\cdots(x-\lambda_m)(x^2+b_1x+c_1)\cdots(x^2+b_Mx+c_M)$, $b_j^2 < 4c_j$.

3.2 March 9

3.2.1 Polynomials

Corollary 3.2.1. If $p \in \mathcal{P}(\mathbb{F})$ is a zero function, and $\operatorname{char}(\mathbb{F}) = 0$, then p is the zero polynomial. (Not true over finite fields).

Chapter 4

Invariant Subspaces

4.1 March 9

4.1.1 Ch5: Eigenvalues, Eigenvectors, and Invariant Subspaces

.
Good Viewpoint: Given $A \in \mathbb{F}^{m,n}$, canonically A corresponds to a linear map $T_A \in \text{Hom}(\mathbb{F}^n, \mathbb{F}^m)$ such that T_A with respect to $((e_1, \ldots, e_n), (e_1, \ldots, e_m))$ is A. In other words $T_A e_j = \sum_{i=1}^n A_{i,j} e_j$.

We will now begin our discussion of linear operators.
Q: $T \in \text{End}(V)$. What is a good baisis $(v_1, \ldots, v_n)$ such that $\mathcal{M}(T)$ with respect to $(v_1, \ldots, v_n), (v_1, \ldots, v_n)$ is "simple"?

Let $T \in \text{End}(V)$. There may be subspaces $U \subset V$ that are invariant under T, we can study $T|_U$.

Definition 4.1.1. If $T \in \text{End}(V)$, $U \subset V$, subspace, is called invariant under T if $Tu \in U$, $\forall u \in U$.

Example 4.1.2. $\{0\}, T, \ker T, \text{im} T$ are invariant under T.

Example 4.1.3. "rotation counterclockwise around 0". Matrix with respect to e_1, e_2 is $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. $\exists$ an invariant subspace iff $\theta = k\pi$, $k \in \mathbb{Z}$.

Example 4.1.4. $\mathcal{P}_m(\mathbb{F}) \subset \mathcal{P}(\mathbb{F})$ is invariant under differentiation.

Definition 4.1.5. If $T \in \text{End}(V)$, $v \neq 0$ is an eigenvector of T corresponding to eigenvalue λ if $Tv = \lambda v$. (directly relates to 1D invariant subspaces)

Theorem 4.1.6. Assume $\dim V < \infty$. TFAE:

(a) $\lambda \in \mathbb{F}$ is an eigenvalue of T

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

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

(d) $T - \lambda I$ is not invertible

Proof. $(b) \leftrightarrow (c) \leftrightarrow (d)$ by fundamental thm of linear maps.
Moreover, λ is an eigenvalue of $T \leftrightarrow \exists v \neq 0$ such that $(T - \lambda I)v = 0 \leftrightarrow T - \lambda I$ is injective.

Theorem 4.1.7. Let $T \in \text{End}(v)$, if $v_1, \ldots, v_m$ are eigenvectors of T corresponding to eigenvalues $\lambda_1, \ldots, \lambda_m$ respectively and $\lambda_1, \ldots, \lambda_m$ are pairwise distinct. Then $v_1, \ldots, v_m$ are linearly independent.

Proof. If $\sum_{j=1}^m a_j v_j = 0$, $a_j \in \mathbb{F}$. Apply T, $T(\sum a_j v_j) = \sum a_j T v_j = \sum \lambda_j a_j = 0$. Apply T again, $\sum_{j=1}^m \lambda_j^2 a_j v_j = 0$. Hence $\sum_{j=1}^m \zeta a_j v_j = 0$ whenever $(\zeta_1, \ldots, \zeta_m)$ is in the span of $(1, \ldots, 1), (\lambda_1, \ldots, \lambda_m), \ldots, (\lambda_1^{m-1}, \ldots, \lambda_m^{m-}$ so by Hw 7-2(b) they span $\mathbb{F}^m$. In particular $a_1 v_1 = 0$ (take $(\zeta_1, \ldots, \zeta_m) = (1, 0 \ldots, 0)$) so $a_1 = 0$. For the same reason all a_j's=0. Hence $v_1, \ldots, v_m$ are linearly independent.

Corollary 4.1.8. If $\dim V < \infty$, the number of distinct eigenvalues of $T \in \text{End}(V)$ is $\leq \dim V$.

Proof. Since the list with 1 egienvector corresponding to each eigenvalue is linearly independent it must have $\leq \dim V$ eigenvalues.

Definition 4.1.9. Assume $T \in \text{End}(V)$ and $U \subset V$, subspace, is invariant under T. Define:
The restriction operator $T|_U \in \text{End}(U)$ by $T|_U = Tu$
The quotient operator $T/U \in \text{End}(V/U)$ by $T/U(u + U) = Tu + U$.

Quotient operator is well defined: If $v_1 + U = \tilde{v}_1 + U$, $v_1, \tilde{v}_1 \in V$, the $v_1 - \tilde{v}_1 \in U$. $Tv_1 - T\tilde{v}_1 = T(v_1 - \tilde{v}_1) \in U$. Hence $Tv_1 + U = T\tilde{v}_1 + U$.

Example 4.1.10. Given the matrices $T|_U$, T/U, find a basis of V such that the matrix of T with respect to the basis relates to the other two.
Soln: Consider a basis of U and extend it to a basis of V. The matrix of T with respect to this basis can be seen as a 2×2 block diagonal matrix with the matrices of $T|_U$ and T/U on its diagonal.

Example 4.1.11. Let $T \in \text{End}(\mathbb{F}^2)$ such that $T(x, y) = (y, 0)$.
Matrix is $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Let $U = \{(x, 0) : x \in \mathbb{F}\}$ invariant under T. $T|_U = 0$, $T/U = 0$.

4.1.2 5.B - Eigenvectors cont. and Upper Triangular Matrices

Definition 4.1.12. For $T \in \text{End}(V)$, define:

$$T^m = \underbrace{T \circ \cdots \circ T}_{m \text{ times}} \quad m \geq 1, m \in \mathbb{Z}$$

$$T^0 = I$$

$$T^{-m} = (T^{-1})^m, \quad m \in \mathbb{Z} \text{ if } T \text{ is invertible}$$

For $p \in \mathcal{P}(\mathbb{F})$, define $p(T)$ plugging T into p.
eg. $p(x) = x^3 + x + \frac{1}{2} \to p(T) = T^3 + T + \frac{1}{2}I$.

Theorem 4.1.13. If $p, q \in \mathcal{P}(\mathbb{F})$, $T \in \text{End}(V)$, then
$(pq)(T) = p(T)q(T) = q(T)p(T)$.

Theorem 4.1.14. Let V be over $\mathbb{C}$, $\dim V < \infty$. Then any $T \in \text{End}(V)$ has an eigenvalue.

Proof. Take arbitrary $v \in V$ such that $v \neq 0$. Assuming $\dim V = n$, consider $v, Tv, T^2v, \ldots, T^nv$, linearly dependent in V. Assuming $a_0v + a_1Tv + \cdots + a_nT^nv = 0$, not all $a_j = 0$. Then let the polynomial $p(x) = a_0 + a_1x + \cdots a_nx^n$. Also, not constant otherwise $a_0 = 0$. Since $p(T) = 0$, factorize p to get $c(T - \lambda_1 I)\cdots(T - \lambda_n I)v = 0$ with $c \neq 0$ so $(T - \lambda_1 I)\cdots(T - \lambda_n I)v = 0$. Hence $T - \lambda_j I$ is not injective for some j so some λ_j is an eigenvalue.

Warning: not true for $\mathbb{R}-$vector spaces.

Definition 4.1.15. The diagonal of $A \in \mathbb{F}^{n,n}$ consists of $A_{1,1}, A_{1,2}, \ldots, A_{n,n}$. $A \in \mathbb{F}^{n,n}$ is called upper triangular if $A_{i,j} = 0$ $\forall i < j$.

Theorem 4.1.16. Let $\dim V < \infty$. For $T \in \text{End}(v)$. Fix a basis $v_1, \ldots, v_n$ of V. TFAE:

(a) $\mathcal{M}(T)$ is upper triangular.

(b) $Tv_j \in \text{span}(v_1, \ldots, v_j)$ $\forall i \leq j \leq n$.

(c) $\text{span}(v_1, \ldots, v_j)$ is invariant over T, $\forall i \leq j \leq n$.

Proof. $(a) \iff (b)$ by def.
$(b) \to (c)$. $\forall j$ $i \leq j$, $Tv_i \in \text{span}(v_1, \ldots, v_i) \subset \text{span}(v_1, \ldots, v_j)$.
$(c) \to (b)$. by (c) $Tv_j \in \text{span}(v_1, \ldots, v_j)$.

Theorem 4.1.17. $\mathbb{F} = \mathbb{C}$, $\dim V < \infty$, $T \in \text{End}(V)$ then $\exists$ basis $v_1, \ldots, v_n$ such that $\mathcal{M}(T)$ is upper triangular.

4.2 March 14

subsection5.B - Eigenvectors cont. and Upper Triangular Matrices

Proof (Proof of Theorem 4.1.17)**.** Assume $\dim V < \infty$. We prove that we can find a linearly independent set of vectors $v_1, \ldots, v_n$ such that (b) of them 13.17 holds by induction. More precisely, we will find v_1, then v_2, then v_3 such that $Tv_j \in \text{span}(v_1, \ldots v_j)$.
Since T is over $\mathbb{C}$ is has an eigenvector v_1. Let $V_1 = \text{span}(v_1)$.

Assuming we have found linearly independent vectors such $v_1, \dots, v_m$ $(m > n)$ such that $Tv_j \in \text{span}(v_1, \dots, v_j)$ $\forall 1 \leq i \leq m$, and denote $V_m = \text{span}(v_1, \dots, v_n)$. Goal in the inductive step: Find v_{m+1} such that $v_1, \dots, v_{m+1}$ linearly independent and $Tv_{m+1} \in \text{span}(v_1, \dots, v_{m+1})$. By assumptions V_m invariant under T so consider the map $T/V_m : V/V_m \to V/V_m$, $\dim V/V_m = n - m > 0$. It has an eigenvector $v_{m+1} + V_m$. Now $v_{m+1} + V_m \neq 0 + V_m$ so $v_{m+1} \notin V_m$. Hence since $v_1, \dots, v_m$ are linearly independent by assumption, $v_1, \dots, v_{m+1}$ are also linearly inpdependent. Moreover $(T/V_m)(v_{m+1}) = \lambda_{m+1} v_{m+1} + V_m$ so $Tv_{m+1} = \lambda_{m+1} v_{m+1} + \tilde{v}$, $\tilde{v} \in V_m$ so $Tv_{m+1} \in \text{span}(v_1, \dots, v_j)$, $\forall 1 \leq j \leq n$, implying the rest of the thm.

Theorem 4.2.1. If $T \in \text{End}(V)$, $\dim V < \infty$, and $\mathcal{M}(T)$ upper triangular with respect to a basis. Then T is invertible iff all the diagonal entries of $\mathcal{M}(T)$ are nonzero.

Proof. $\to$) If $\mathcal{M}(T)_{j,j}$ is zero, then im$T \subset \text{span}(v_1, \dots, v_{j-1}, Tv_{j+1}, \dots, Tv_n$ dim $< n$ so T is not invertible. $\leftarrow$) Suppose $\mathcal{M}(T)$ is upper triangular entries with nonzero diagonal entries. If $T(a_1 v_1 + \dots + a_m T v_m) = 0$, $m \leq n$, then since $Tv_1, \dots, Tv_m \in \text{span}(v_1, \dots, v_m)$ and $T(a_m v_m) = a_m \mathcal{M}(T)_{m,m} v_m + \tilde{v}$, $\tilde{v} \in \text{span}(v_1, \dots, v_{m-1})$ this implies $a_m = 0$. So $T(a_1 v_1 + \dots + a_{m-1} v_{m-1}) = 0$. Repeating this argument we see that $a_1 = \dots = a_m = 0$ so T is injective.

Theorem 4.2.2. If $T \in \text{End}(V)$ is upper triangular with respect to $v_1, \dots, v_n$ then {diagonal entries of $\mathcal{M}(T)$} = {eigenvalues of T}

Proof. λ is not an eigenvalue $\leftrightarrow$ $T - \lambda I$ is invertible $\leftrightarrow$ $\lambda \neq$ any diagonal element of $\mathcal{M}(T)$. (Since $T_\lambda I$ is also upper triangular wrt $v_1 k \dots, v_n$) .

4.2.1 Change of Basis

Theorem 4.2.3. Assume $\dim V, \dim W < \infty$

(1) $T \in \text{End}(V)$ is invertible $\iff$ $\mathcal{M}(T)$ is invertible with respect to some matrix $\leftrightarrow$ $\mathcal{M}(T)$ is invertible with respect to every matrix.

(2) If $T \in \text{Hom}(V, W)$, and $S \in \text{Hom}(V, W)$ are inverses of each other then the matrices of T and S are inverses of each other.

Theorem 4.2.4. If $v_1, \dots, v_m$ and $u_1, \dots, u_m$ are bases of V. $A = \mathcal{M}(I, (u_1, \dots, u_m), (v_1, \dots, v_m))$. Then $\mathcal{M}(T, (u_1, \dots, u_m)) = A^{-1} \mathcal{M}(T, (v_1, \dots, v_m)) A$.

Proof. View $T = ITI$, with 1rst I wrt. $(v_1, \dots, v_n), (u_1, \dots, u_n)$, T wrt $(v_1, \dots, v_n)$, second I wrt. $(u_1, \dots, u_n) + (v_1, \dots, v_m)$. Then we see that T is wrt. $(u_1, \dots, u_n)$.

Corollary 4.2.5. $\forall B \in C^{n,n}$, $\exists$ invertible $A \in C^{n,n}$ such that $A^{-1} B A$ is upper triangular.

4.2.2 Eigenvalues and Diagonal Matrices

Definition 4.2.6. A diagonal matrix is a matrix whose off-diagonal entries are all 0. If A if diagonal we can write $A =$diag$(A_{1,1}, A_{2,2}, \ldots, A_{n,n})$

Definition 4.2.7. For $\lambda \in \mathbb{F}$, the eigenspace of λ wrt. $T \in \text{End}(V)$ is $\ker(T - \lambda I)$, denoted as $E(\lambda, T)$.

Example 4.2.8. If the matrix of T wrt. (v_1, v_2, v_3) is $\begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}$, then $E(1,T) = \text{span}(v_1)$, $E(2,T) = \text{span}(v_2)$, $E(3,T) = \text{span}(v_3)$, $E(\lambda, T) = \{0\}$ for $\lambda \notin \{1, 2, 3\}$.

Theorem 4.2.9. If $\lambda_1, \ldots, \lambda_n$ are pairwise distinct, then $E(\lambda_1, T) + E(\lambda_n, T)$ is a direct sum.

Proof. WLOG, assume all $v_i \neq 0$. If $v_1 + \cdots + v_n = 0$ with $v_i \in E(\lambda_i, T)$ then since eigenvectors of distinct eigenvalues are linearly independent, all $v_i = 0$

Definition 4.2.10. $\dim V < \infty$, $T \in \text{End}(V)$ is diagonalizable if $\exists$ basis $v_1, \ldots, v_n$ such that $\mathcal{M}(T)$ is diagonal.

Theorem 4.2.11. Assume $\dim V < \infty$, $T \in \text{End}(V)$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ (finitely many), TFAE:

(a) T is diagonalizable.

(b) There is a basis whose vectors are all eigenvectors.

(c) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$

(d) $\dim V - \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$

Proof. $(a) \leftrightarrow (b)$ by def. Also $(d) \iff (e)$.
$(d) \to (b)$: Take a basis from each $E(\lambda_j, T)$ and add them all together.
$(b) \to (d)$: Take a basis of $v_1, \ldots, v_n$. Each v_k has to be in some $E(\lambda_j, T)$. Hence in $E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.
Thus, $V \subset E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

Example 4.2.12. $T : (x_1, x_2) \mapsto (x_2, 0)$ in $\mathbb{F}^2$ is not diagonalizable.
$T - \lambda I$: $(x_1, x_2) \mapsto (x_2 - \lambda x_1, \lambda x_2)$ is invertible iff $\lambda \neq 0$. But $E(0,T)$ is 1 dimensional so it is not diagonalizable.

Theorem 4.2.13. If $T \in \text{End}(V)$ has $\dim V (< \infty)$ distinct eigenvalues T is diagonalizble.

Proof. Note if λ is an eigenvalue, then $\dim E(\lambda, 1) \geq 1$ so all have dimension 1 since n of them.

Example 4.2.14. If $\dim V = 3$, $T \in \mathrm{End}(V)$ has a matrix $\begin{pmatrix} 2 & ? & ? \\ 0 & 5 & ? \\ 0 & 0 & 8 \end{pmatrix}$. T is diagonalizable. (when 5 is replaced by 2, T can be nondiagonalizable)

Chapter 5

Inner Product Spaces

5.1 March 14

5.1.1 Ch 6 - Inner Products

Motivation: Euclidean Geometry,
Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Definition 5.1.1. Over $\mathbb{F}^n$, the Euclidean inner product (dot product if $\mathbb{F} = \mathbb{R}$) of $(w_1, \dots, w_n)$ of $(z_1, \dots, z_n)$ is $w_1\overline{z}_1 + \cdots + w_n\overline{z}_n$.

5.2 March 16

5.2.1 Ch 6 - Inner Products

Motivation: can talk about angles, lengths, orthogonality, etc.

Definition 5.2.1. An inner product on V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that

(a) Positive Definiteness: $\langle v, v \rangle \geq 0\ \forall v \in V$. Equality iff $v = 0$.

(b) Linear in First Spot: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle, \quad \langle \lambda v, w \rangle = \lambda \langle v, w \rangle$

(c) Conjugation Symmetry: $\langle u, v \rangle = \overline{\langle v, u \rangle}$

Example 5.2.2. This implies: $\langle u, \lambda_1 v_1 + \lambda_2 v_2 \rangle = \overline{\lambda}_1 \langle u, v_1 \rangle + \overline{\lambda}_2 \langle u, v_2 \rangle$
Example 5.2.3. of inner products:

(a) Euclidean inner product on $\mathbb{F}^n$

(b) For $f, g \in C[0,1]$ (continuous functions from $[0,1]$ to $\mathbb{C}$), define
$\langle f, g \rangle = \int_0^1 f\overline{g}dx$. It is an inner product on $C[0,1]$

(c) For $f, g \in C[0,1]$, $\langle f, g, \rangle = \int_0^1 f\overline{g}e^x dx$ is also an inner product.

(d) For $f, g \in C[0,1]$, $\langle f, g \rangle = \int_0^{\frac{1}{2}} f\overline{g}dx$ is not an inner product.

Definition 5.2.4. An inner product space is a vector space $/\mathbb{R}$ or $\mathbb{C}$ equipped with an inner product.

Theorem 5.2.5.

(a) $\langle \cdot, c\rangle$ is linear if u is fixed. ie. $\langle \lambda_1 v_1 + \lambda_2 v_2, u\rangle = \lambda_1\langle v_1, u\rangle + \lambda_2\langle v_2, u\rangle$.

(b) $\langle 0, v\rangle = \langle v, 0\rangle$, $\forall v \in V$.

Definition 5.2.6. u and v are orthogonal if $\langle u, v\rangle = 0$. In this case, we say $u \perp v$.

Proposition 5.2.7.

(a) 0 is orthogonal to v, $\forall v \in V$

(b) If $v \perp v$, then $v = 0$ (by positive definiteness)

Definition 5.2.8. For v, V, the norm is $||v|| = \sqrt{\langle v, v\rangle}$

Proposition 5.2.9.

(a) $||v|| = 0 \leftrightarrow v = 0$

(b) $||\lambda v|| = |\lambda| \cdot ||v||$, $\forall \lambda \in \mathbb{F}, v \in V$.

Proof. $\langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2 \cdot ||v||^2$

Theorem 5.2.10 (Pythagorean Theorem)**.** $u \perp v \to ||u + v||^2 = ||u||^2 + ||v||^2$.

Proof. $||u + v||^2 = \langle u + v, u + v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle = ||u||^2 + ||v||^2 + 0 + 0$.

For $u, v \in V$, want to be able to describe u as some scalar multiple of v and some vector orthogonal to it. If $u_1 = ?v$, $u_2 = u - u_1$, what is "?". $\langle u_1, v\rangle = \langle u, v\rangle - \langle u_2, v\rangle = \langle u, v\rangle$. Also $\langle u_1, v\rangle = ?||v||^2$ so $? = \frac{\langle u,v\rangle}{||v||^2}$.

Theorem 5.2.11 (Vector Projection)**.** For $u, v \in V$, let $u_1 = \frac{\langle u,v\rangle}{||v||^2}v$ and $u_2 = v - u_1$, then $v \perp u_2$, u is "along the direction of v", called the vector projection of u onto v.

Proof. Compute $\langle u_2, v\rangle = \langle u - u_1, v\rangle = \langle u, v\rangle - \langle u_1, v\rangle = \langle u, v\rangle - \langle \frac{\langle u,v\rangle}{||v||^2}v, v\rangle = \langle u, v\rangle - \frac{\langle u,v\rangle}{||v||^2}\langle u, v\rangle = 0$.

Theorem 5.2.12 (Cauchy-Schwarz)**.** $|\rangle u, v\langle| \leq ||u||||v||$, $\forall u, v \in V$.

Proof. WLOG, assume $v \neq 0$. Form u_1 as above, $v = u_1 + u_2$. $u_1 = \frac{\langle u,v\rangle}{||v||^2}v$, $u_2 \perp v$. $||u_2||^2 \geq 0 \leftrightarrow ||u||^2 \geq ||u_1||^2 \leftrightarrow ||u||^2 \geq \frac{|\langle u,v\rangle|^2}{||v||^4}||v||^2$. Thm follows by taking square roots.

Example 5.2.13. Use $||v - \lambda v||^2 \geq 0$ to give another proof.

Corollary 5.2.14.

(a) For $x_1, \dots, x_n$, $y_1, \dots, y_n \in \mathbb{C}$. $|\sum_{j=1}^n x_j \overline{y}_j| \leq \sqrt{\sum_{j=1}^n |x_j|^2}\sqrt{\sum_{j=1}^n |y_j|^2}$

(b) For $f, g \in C[0,1]$, $|\int_0^1 f\overline{g}dx|^2 \leq (\int_0^1 |f|^2 dx)(\int_0^1 |g|^2 dx)$, and $|\int_0^1 f\overline{g}e^x dx|^2 \leq (\int_0^1 |f|^2 e^x e dx)(\int_0^1 |g|^2 e^x dx)$

Theorem 5.2.15 (Triangle Inequality)**.** $||u+v|| \leq ||u|| + ||v||$, $\forall u, v \in V$.

Proof.

$$\begin{aligned}||u+v||^2 &= \langle u+v, u+v\rangle \\ &= ||u||^2 + ||v||^2 + \langle u,v\rangle + \langle v,u\rangle \\ &= ||u||^2 + ||v||^2 + 2\mathrm{Re}(\langle u,v\rangle) \\ &\leq ||u||^2 + ||v||^2 + 2||u||\,||v|| \\ &= (||u|| + ||v||)^2\end{aligned}$$

Example 5.2.16 (Parallelogram Identity)**.** $||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2)$.
For $u, v \neq 0$, the angle between them is $\arccos \frac{\langle u,v\rangle}{||u||\,||v||}$

Norms in General:
A norm on a vector space U is $||\cdot|| : U \to \mathbb{R}^{\geq 0}$ such that:

- Positive Definiteness: $||u|| = 0 \leftrightarrow u = 0$, $\forall u \in U$
- Absolute Homogeneity: $||\lambda u|| = |\lambda|\,||u||$
- Triangle Inequality: $||u+v|| \leq ||u|| + ||v||$

We proved an inner product gives rise to a norm.

5.2.2 6.B - Orthogonal Bases

Definition 5.2.17. A list $v_1, \dots, v_m$ in V is orthogonal if each $||v_j|| = 1$, $1 \leq j \leq m$ and $v_{j_1} \perp v_{j_2}$, $\forall j_1 \neq j_2 \in \{1, \dots, m\}$.

Example 5.2.18. Standard basis in $\mathbb{F}^n$ is normal.
usepackage(Note: we will not use $e_1, \dots, e_n$ to denote the standard basis in in Ch 6, they will be used to denote a general orthogonal list.)

Proposition 5.2.19. If $e_1, \dots, e_m$ is orthonormal, then $||a_1 v_1 + \cdots + a_m v_m|| = \sqrt{|a_1|^2 + \cdots + |a_n|^2}$.

Proof. Expand $||a_1 e_1 + \cdots a_m e_m||^2 = \langle a_1 e_1 + \cdots + a_m e_m, a_1 e_1 + \cdots + a_m e_m \rangle = |a_1|^2 + \cdots + |a_m|^2$.

Example 5.2.20. $\langle a_1 e_1 + \cdots + a_m e_m, b_1 e_1 + \cdots + b_m e_m \rangle = a_1 \overline{b}_1 + \cdots + a_m \overline{b}_m$

Corollary 5.2.21. An orthonormal list is linearly independent.

Proof. Assume the list is $e_1, \dots, e_m$. If $\sum_j a_j e_j = 0$, then $\sum |a_j|^2 = 0$ so all $a_j = 0$.

Corollary 5.2.22. If $\dim V = n$, a list of n orthonormal vectors is a basis (orthonormal basis).

Theorem 5.2.23. If $e_1, \dots, e_m$ is an orthonormal basis, then $\forall v \in V$, $v = \sum_{i=1}^n \langle v, e_j \rangle e_j$.

Proof. We know there are $\lambda_1, \dots, \lambda_n$ such that $v = \sum_j \lambda_j e_j$. But $\langle v, e_k \rangle = \sum_j \lambda_j \langle e_j, e_k \rangle = \lambda_k$.

Gram-Schmidt Procedure: An algorithm with:

- input: basis $v_1, \dots, v_n$
- output: orthonormal basis: $e_1, \dots, e_n$ such that $e_j \in \text{span}\{v_1, \dots, v_j\}$.

Theorem 5.2.24 (Gram-Schmidt Procedure)**.** Given a basis $v_1, \dots, v_n \in V$, define

$$e_j = \frac{v_j - \sum_{k=1}^{j-1} \langle v_j, e_k \rangle e_k}{||v_j - \sum_{k=1}^{j-1} \langle v_j, e_k \rangle e_k||}$$

Proof. We need to check:

- e_j is well defined.
- $||e_j|| = 1$
- $e_j \perp e_k$, $k < j$
- $\text{span}(e_1, \dots, e_j) = \text{span}(v_1, \dots, v_j)$

Assume we are at step j:
For 1rst item, we need $||v_j - \sum_{k=1}^{j-1} \langle v_j, e_k \rangle e_k|| \neq 0$. True since $e_k \in \text{span}(v_1, \dots, v_{j-1})$, $k < j$. But $v_j \notin \text{span}(v_1, \dots, v_{j-1})$.
2cnd item is clear.
For 3rd item, compute $\langle v_j - \sum_{j=1}^n \langle v_j, e_k \rangle, e_l \rangle$ (for $l < k$) $= \langle v_j, e_l \rangle - \langle v_j, e_l \rangle = 0$.

For 4th item, note that $e_1, \dots, e_{j-1}$ already in $\operatorname{span}(v_1, \dots, v_{j-1})$, by def $e_j \in \operatorname{span}(v_1, \dots, v_j)$. Moreover $e_1, \dots, e_j$ are linearly independent so $\operatorname{span}(e_1, \dots, e_j) = \operatorname{span}(v_1, \dots, v_j)$.

5.3 March 28

5.3.1 6.B - Orthogonal Bases

Correction: For a vector space V, use V' to denote its dual space. For $T \in \operatorname{Hom}(V, W)$, use T' to denote its dual map.

Theorem 5.3.1 (Bessel's Inequality). In V, if the list $e_1, \dots, e_n$ is an orthonormal, then $\forall v \in V$,

$$||v||^2 \geq \sum_{j=1}^{m} |\langle v, e_j \rangle|^2.$$

The equality holds iff $v \in \operatorname{span}(e_1, \dots, e_n)$.

Proof. $0 \leq ||v - \sum_{j=1}^{m} \langle v, e_j \rangle e_j||^2 = ||v||^2 - \sum_{j=1}^{m} |\langle v, e_j \rangle|^2$.

Example 5.3.2. In $\mathcal{P}_2(\mathbb{C})$, define the inner product $\langle f, g \rangle = \int_{-1}^{1} f\overline{g}dx$. Applying G-s to $(1, x, x^2)$, we get $(\frac{\sqrt{2}}{2}, \frac{\sqrt{6}}{2}x, \frac{3\sqrt{10}}{4}(x^2 - \frac{1}{3})$.

Example 5.3.3. In $\mathbb{R}^2$, apply G-S to $(1, 2)$, $(3, 4)$, we get $\frac{\sqrt{5}}{5}(1, 2)$, $\frac{\sqrt{5}}{5}(2, -1)$.

Corollary 5.3.4. Every finite dimensional inner product space has an orthonormal basis.

Corollary 5.3.5. If V is finite dimensional, then every orthonormal list can be extended to an orthonormal basis.

Proof. Assume the list is $e_1, \dots, e_m$, extend it to a basis $e_1, \dots, e_m, v_{m+1}, \dots, v_n$. Apply G-S, the first m vectors don't change (can be shown inductively).

Corollary 5.3.6 (Schur's Theorem). Assume $\mathbb{F} = \mathbb{C}$, $\dim V < \infty$, every $T \in \operatorname{End}(V)$ has an upper triangular matrix with respect to some basis of V.

Proof. T is upper triangular with respect to $v_1, \dots, v_m$. Apply G-S to get orthonormal basis $e_1, \dots, e_m$. Now, $\forall j$m, $Te_j \in \operatorname{span}(Tv_j, Te_1, \dots, Te_{j-1}) \subset \operatorname{span}(Tv_1, \dots, Tv_j) \subset \operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_j)$. Thus, T us upper triangular with respect to $e_1, \dots, e_n$.

Theorem 5.3.7 (Reisz Representation Theorem, finite dimensional case). If V is finite dimensional, $\varphi \in V'$, then $\exists! u \in V$ such that $\langle v, u \rangle = \varphi(v)$, $\forall v \in V$.

Proof. Uniqueness: If u_1, u_2 satisfy $\varphi(v) = \langle v, u_1, \rangle = \langle v, u_2, \rangle$, $\forall v$. Then $\langle v, u_1 - u_2 \rangle = 0$ $\forall v$ so taking $v = u_1 - u_2$ implies $u_1 = u - 2$.
Existence: Take an orthonormal basis $e_1, \dots, e_n$. Take $u = \overline{\varphi(e_1)}e_1 + \overline{\varphi(e_2)}e_2 + \cdots + \overline{\varphi(e_n)}e_n$. Then $\forall v \in V$, assuming $v = \sum_{j=1}^n a_j e_j$, then $\langle v, u \rangle = a_1\varphi(e_1) + a_2\varphi(e_2) + \cdots + a_n\varphi(e_n) = \varphi(a_1e_1 + \cdots + a_ne_n) = \varphi(v)$.

Example 5.3.8. More intrinsic proof: Observe $\ker\varphi$ is an $n-1$ dimensional subspace of $\varphi \neq 0$.

Note: $e^{i\theta} = \cos\theta + i\sin\theta$, $\theta \in \mathbb{R}$.

Example 5.3.9. $\int_0^1 e^{2n\pi ix}e^{-2m\pi ix}dx = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$ where $m, n \in \mathbb{Z}$.

Let $V = \text{span}(1, e^{2\pi ix}, e^{-2\pi ix}, e^{4\pi ix}, e^{-4\pi ix} \subset C[0,1]$. Define $\langle \cdot, \cdot \rangle$ on V by $\langle f, g \rangle = \int_{-1/2}^{1/2} f\overline{g}dx$. Take $\varphi \in V'$ such that $\varphi(f) = \int_{-1/2}^{1/2} f \cdot x dx$. Find u such that $\varphi(f) = \langle f, u, \rangle$, $\forall g \in V$. $(u = \frac{1}{\pi}\sin(2\pi x) - \frac{1}{2\pi}\sin(4\pi x)$

Corollary 5.3.10 (QR Factorization)**.** Let $M \in \mathbb{F}^{n,m}$ have linearly independent columns, then $\exists!$ pair (Q, R) such that $Q \in \mathbb{F}^{n,m}$, $R \in \mathbb{F}^{m,m}$, $M = QR$, Q has orthonormal column vectors and R is upper triangular with positive diagonal entries.

Proof. Do G-S to columns of M for existence.

5.3.2 6.C - Orthogonal Complements and Minimization

Definition 5.3.11. For a subset $S \subset V$, define the orthogonal complement of S to be $S^\perp = \{v \in V : \langle v, s \langle = 0, \forall s \in S\}$.

Proposition 5.3.12.

(a) $S^\perp$ is a subspace of V.

(b) $\{0\}^\perp = V$

(c) $V^\perp = \{0\}$

(d) $S \cap S^\perp = \{0\}$

(e) $S_1 \subset S_2 \to S_2^\perp \subset S_1^\perp$

(f) $S^\perp = \text{span}(v : v \in S)^\perp$ - set of finite linear combinations of vectors from S.

Theorem 5.3.13. If $U \subset V$, U finite dimensional, then $V = U \oplus U^\perp$.

Proof. Let $e_1, \dots, e_n$ be an orthonormal basis of U, $\forall v \in V$, $v = \sum_{j=1}^n \langle v, e_j \rangle e_j + (v - \sum_{j=1}^n \langle v, e_j \rangle e_j)$. Then $\forall 1 \leq k \leq n$, $\langle v - \sum_{j=1}^n \langle v, e_j \rangle e_j, e_k \rangle = \langle v, e_j \rangle - \langle v, e_k \rangle \langle e_k, e_k \rangle = 0$. Hence $v_2 \in U^\perp$ and $v_1 \in U$ so $V = U + U^\perp$. Hence, $V = U \oplus U^\perp$.

Corollary 5.3.14. If $U \subset V$, $\dim V < \infty$, then $\dim U^\perp = \dim v - \dim U$.

Theorem 5.3.15. If $U \subset V$, $\dim U < \infty$, then $U = (U^\perp)^\perp$.

Proof. $\forall u \in U$, $\forall v \in U^\perp$, $\langle u, v \rangle = 0$. Hence $u \in (U^\perp)^\perp$ so $U \subset (U^\perp)^\perp$.
For $w \in (U^\perp)^\perp$, $w = w_1 + w_2$ for $w_1 \in U$, $w_2 \in U^\perp$. But $\langle w, w_2 \rangle = 0 = \langle w_1 + w_2, w_2 \rangle = ||w_2||^2$ so $w_2 = 0$ so $w \in U$. Thus, $U = (U^\perp)^\perp$.

Definition 5.3.16. If $U \subset V$, finite dimensional, define the orthogonal projection, P_U to be: for $v \in V$, write $v = v_1 + v_2$ with $v_1 \in U$, $v_2 \in U^\perp$ and define $P_U v = v_1$.

Theorem 5.3.17.

(a) $P_U \in \mathrm{End}(V)$

(b) $P_U^2 = P_U$

(c) $\mathrm{im} P_U = U$

(d) $\ker P_U = U^\perp$

Proof. By thm, $E(0, P_U) = U^\perp$, $E(1, P) = U$ (since $E(0, P_U) \oplus E(1, P_U) = v$, so P_U has no eigenvalues).

(e) $v - P_U v \in U^\perp$

(f) $||P_U v|| \leq v$ (since $v = P_u v + v_2$, Pythagorean Theorem)

(g) If $e_1, \ldots, e_n$ is an orthonormal basis of U, then $P_u v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$.

Theorem 5.3.18. If $U \subset V$, U finite dimensional,$v \in V$ then $\forall u \in U$, $||v - P_U v|| \leq ||v - u||$. Equality iff $u = P_U v$.

Chapter 6

Operators

6.1 April 4

6.1.1 Ch 7 - Linear Operators on Inner Product Spaces

Motivation: Which operators can be diagonalized using an orthonormal basis?
Ans: Spectral Theorem: Self-Adjoint operators for $\mathbb{F} = \mathbb{R}$, normal operators for $\mathbb{F} = \mathbb{C}$. Self adjoint/ normal operators defined using with simple expressions, show up naturally and are important in their own rights.
V, W : Finite dimensional inner product spaces throughout Ch 7

6.1.2 7.A - Self Adjoint / Normal Operators

Definition 6.1.1. For $T \in \text{Hom}(V, W)$m the adjoint of T is an operator $T^* : W \to V$ such that

$$\langle Tv, w\rangle = \langle v, T^*w\rangle, \, \forall v \in V, \, \forall w \in W \tag{1}$$

Well Defined: (1) defines T^* uniquely and $\exists T^*$ satisfying (1).

Proof. For $w \in W$, definte the linear functional $\varphi : v \mapsto \langle Tv, w\rangle$. There is a $T^*w \in W$ such that $\varphi(v) = \langle Tv, w\rangle = \langle v, T^*w\rangle$ by the Reisz Representation Theorem. Uniqueness follows. □

Theorem 6.1.2. T^* is a linear map (ie. $T^* \in \text{Hom}(W, V)$)

Proof. $\langle v, T^*(w_1+w_2)\rangle = \langle Tv, w_1+w_2\rangle = \langle Tv, w_1\rangle+\langle Tv, w_2\rangle = \langle v, T^*w_1\rangle+\langle v, T^*w_2\rangle = \langle v, T^*w_1+T^*w_2\rangle$. $\langle v, T^(\lambda w)\rangle = \langle Tv, \lambda w\rangle = \overline{\lambda}\langle Tv, w\rangle = \overline{\lambda}\langle v, T^*w\rangle = \langle v, \lambda T^*w\rangle$

Example 6.1.3. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the map such that $\mathcal{M}(T) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ with respect to the standard basis. Then,

$$\begin{aligned} \langle (x_1, x_2), T^*((y_1, y_2))\rangle &= \langle T((x_1, x_2)), (y_1, y_2))\rangle \\ &= \langle (\cos\theta x_1 - \sin\theta x_2, \sin\theta x_1 + \cos\theta x_2), (y_1, y_2)\rangle \\ &= \cos\theta x_1 y_1 - \sin\theta x_2 y_1 + \sin\theta x_1 y_2 + \cos\theta x_2 y_2 \\ &= \langle (x_1, x_2), (\cos\theta y_1 + \sin\theta y_2, -\sin\theta y_1 + \cos\theta y_2)\rangle \end{aligned}$$

so $\mathcal{M}(T^*)$ with respect to the standard basis is $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

Compare: If V_1, V_2 vector spaces, not equipped with an inner product, $T \in \text{Hom}(V_1, V_2)$ then $T' \in \text{Hom}(V_2', V_1')$.

Theorem 6.1.4 (Basic Properties of T^*)**.** $\forall \lambda \in \mathbb{F}$, $S, T \in \text{Hom}(V, W)$

(a) $(S+T)^* = S^* + T^*$

(b) $(\lambda T)^* = \overline{\lambda} T^*$

(c) $(T^*)^* = T$

(d) $(ST)^* = T^* S^*$

Proof. $\forall v \in V$, $\forall w \in W$,

1. $\langle v, (S+T)^* w\rangle = \langle (S+T)v, w\rangle = \langle Sv + Tv, w\rangle = \langle Sv, w\rangle + \langle Tv, w\rangle = \langle v, S^* w\rangle + \langle v, T^* w\rangle = \langle v, S^* w + T^* w\rangle$
2. $\langle v, (\lambda T)^* w\rangle = \langle \lambda T v, w\rangle = \lambda \langle Tv, w\rangle = \lambda \langle v, T^* w\rangle = \langle, v, \overline{\lambda} T w\rangle$
3. $\langle w, (T^*)^* v\rangle = \langle T^* w, v\rangle = \overline{\langle v, T^* w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle$
4. $\langle v, I^* w\rangle = \langle Iv, w\rangle = \langle v, w\rangle = \langle v, Iw\rangle$
   $\langle v, 0^* w\rangle = \langle 0v, w\rangle = 0 = \langle v, 0w\rangle$
5. Assume $v \in V$, $u \in U$,
   $\langle v, (ST)^* u\rangle = \langle STv, u\rangle = \langle Tv, S^* u\rangle = \langle v, T^* S^* y\rangle$

Proposition 6.1.5.

(a) $\ker T^* = (\text{im} T)^\perp$

(b) $\text{im} T^* = (\ker T)^\perp$

(c) $\ker T = (\text{im} T^*)^\perp$

(d) $\text{im} T = (\ker T^*)^\perp$

Proof (Proof of (a))**.** $(\text{im} T)^\perp = \{w : \langle Tv, w\rangle = 0, \forall v \in V\} = \{w : \langle v, T^* w\rangle = 0, \forall v \in V\} = \{w : T^* w = 0\} = \ker T^*$

What is $\mathcal{M}(T^*)$? (if $T \in \text{Hom}(V, W)$)
Fix an orthonormal basis $e_1, \ldots, e_n$ of V and $f_1, \ldots, f_m$ of W.

Then for $v = \sum_{j=1}^{n} a_j e_j$, $w = \sum_{k=1}^{m} b_k f_k$, assume $\mathcal{M}(T) = A$,

$$\begin{aligned}
\langle Tv, w\rangle &= \langle \sum_{j=1}^{n} a_j T e_j, \sum_{k=1}^{m} b_k f_K \rangle \\
&= \langle \sum_{k=1}^{m} (\sum_{j=1}^{n} a_j A_{k,j}) f_k, \sum_{k=1}^{m} b_k f_k \rangle \\
&= \sum_{k=1}^{m} \sum_{j=1}^{n} a_j A_{k,j} \overline{b_k} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{m} \langle A_{k,j}, \overline{b_k} \rangle a_j \\
&= \langle \sum_{j=1}^{k} a_j e_j, \sum_{j=1}^{n} (\sum_{k=1}^{m} \overline{A_{j,k}} b_k) e_j \rangle
\end{aligned}$$

By definition of T^*, $T^*(\sum_{k=1}^{m} b_k f_k) = \sum_{j=1}^{n} (\sum_{k=1}^{m} \overline{A_{j,k}} b_k) e_j$ so $\mathcal{M}(T^*)_{j,k} = \overline{A_{k,j}}$

Definition 6.1.6. For $A \in \mathbb{F}^{m,n}$, the conjugate transpose $\overline{A}^t \in \mathbb{F}^{n,m}$ is defined by $(\overline{A}^t)_{j,k} = \overline{A}_{k,j}$.

Example 6.1.7. $\overline{\begin{pmatrix} 1 & i \\ 2 & 3 \end{pmatrix}}^t = \begin{pmatrix} 1 & 2 \\ -i & 3 \end{pmatrix}$

Theorem 6.1.8. If $T \in \mathrm{Hom}(V, W)$, fix an orthonormal basis $e_1, \dots, e_n$ of V, and an orthonormal basis $f_1, \dots, f_m$ of W. Then $\mathcal{M}(T^*) = \overline{\mathcal{M}(T)}^t$.

Definition 6.1.9. $T \in \mathrm{End}(V)$ is self adjoint (or Hermetian) if $T^* = T$. $A \in \mathbb{F}^{m,n}$ is self adjoint (or Hermetian) if $\overline{A}^t = A$.

T self adjoint $\leftrightarrow$ $\mathcal{M}(T)$ is self adjoint with respect to some orthonormal basis $\leftrightarrow$ $\mathcal{M}(T)$ is self adjoint with respect to every orthonormal basis.
Compare: When $A^t = A$, we say A is symmetric. If $\mathbb{F} = \mathbb{R}$, A is self adjoint $\leftrightarrow$ A is symmetric.

Proposition 6.1.10. Every eigenvalue of a self adjoint $T \in \mathrm{End}(V)$ is real.

Proof. IF v is an eigenvector of T with eigenvalue λ,
$\lambda ||v||^2 = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, T^* v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda} ||v||^2$. Since $||V||^2 > 0$, $\lambda = \overline{\lambda}$.

Proposition 6.1.11. If $\mathbb{F} = \mathbb{C}$, $\langle Tv, v\rangle = 0$ $\forall v$ then $T = 0$.

Proof. Note that

$$\langle Tv, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4} - i\frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4}$$

Corollary 6.1.12. If $\mathbb{F} = \mathbb{C}$, $\langle Tv, Tv\rangle \in \mathbb{R}\ \forall v \leftrightarrow T$ is self adjoint.

Proof. Since $\langle Tv, Tv\rangle \in \mathbb{R}, \forall v \leftrightarrow \langle (T - T^*)v, v\rangle = 0\ \forall v \in V$.

Proposition 6.1.13. If $\mathbb{F} = \mathbb{R}$ and T is self adjoint, then if $\langle Tv, v\rangle = 0\ \forall v \in V$, $T = 0$.

Proof.

$$\langle Tv, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}$$

(true if $T^* = T$)

Example 6.1.14. Prove by considering the matrices.

Definition 6.1.15. $T \in \mathrm{End}(V)$ is normal if $TT^* = T^*T$, $A \in \mathbb{F}^{n,m}$ is normal if $A\overline{A}^t = \overline{A}^t A$.

Example 6.1.16. If $\mathcal{M}(T) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ with respect to the standard basis of $\mathbb{R}^2$, then T is normal.

Lemma 6.1.17. T is normal iff $||Tv|| = ||T^*v||\ \forall v \in V$

Proof.

$$\begin{aligned} T \text{ is normal } &\leftrightarrow T^*T - TT^* = 0 \\ &\leftrightarrow \langle (T^*T - TT^*)v, v\rangle = 0\,\forall v \in V \\ &\leftrightarrow ||Tv||^2 = ||T^*v||^2\,\forall v \in V \end{aligned}$$

Unusual! Take orthonormal bases $e_1, \ldots, e_m$, $f_1, \ldots, f_n$

$$\sum_{j=1}^{m} ||Te_j||^2 = \sum_{j=1}^{m}\sum_{k=1}^{n} ||\langle Te_j, f_k\rangle f_k||^2 = \sum_{j=1}^{m}\sum_{k=1}^{n} ||\langle e_j, T^*f_k\rangle f_k||^2 = \sum_{k=1}^{n} ||T^*f_k||^2$$

Sum independent of orthonormal basis.

6.2 April 6

6.2.1 7.A - Self Adjoint/ Normal Operators

$\sqrt{\sum_{j=1}^{n} ||Te_j||^2}$ $(e_1, \ldots, e_n)$ is called the Hilbert-Schmidt norm of T.

Theorem 6.2.1. If T is normal,

(a) T and T^* have the same eigenvectors.

(b) Eigenvectors of T corresponding to different eigenvalues are orthogonal.

Proof.

1. Note $T - \lambda I$ is also normal because $(T-\lambda I)(T-\lambda)^* = TT^* - \lambda T^* \overline{\lambda} T + |\lambda|^2 I = (T-\lambda I)^*(T-\lambda I)$. If $Tv = \lambda v$, then $||(T-\lambda I)v|| = 0$ so $||(T-\lambda I)^* v|| = 0$ so $(T-\lambda I)^* v = 0$ so $T^* v = \overline{\lambda} v$

2. If $Tv_1 = \lambda_1 v_1, Tv_2 = \lambda_2 v_2$, then $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = \lambda_1 \langle v_1, v_2\rangle - \lambda_2 \langle v_1, v_2\rangle = \langle Tv_1, v_2\rangle - \langle v_1, T^* v_2\rangle = 0$.

6.2.2 7.B- Spectral Theorems

Ways to Prove Them:

$$\begin{cases} \text{Prove } \mathbb{C} \text{ version and } \mathbb{R} \text{ version separately} \\ \text{Prove both versions in closely related way} \\ \text{Prove } \mathbb{C} \text{ version, then derive } \mathbb{R} \text{ version as a corrollary} \end{cases}$$

Theorem 6.2.2. If $\dim V \geq 1$, $T \in \text{End}(V)$, T is self adjoint, then T has an eigenvalue.

Lemma 6.2.3. If $T \in \text{End}(V)$ is self adjoint, then if $b^2 < 4c$, $T^2 + bT + c$ is invertible.

Proof. $\forall v \neq 0$, $\langle (T^2 + bT + c)v, v\rangle = \langle (T^2 + \frac{b}{2}I)^2 + (c - \frac{b}{4})^2 I\rangle = ||(T + \frac{b}{2}I)^2||_(^2 c - \frac{b}{4})^2 ||v||^2 > 0$. Hence $(T^2 + bT + c)v \neq 0$ so $T^2 + bT + c$ is invertible.

Proof (Proof Theorem)**.** WLOG, $\mathbb{F} = \mathbb{R}$, let $n = \dim V$. Take $v \neq 0$, $v, Tv, \ldots, T^n v$ is linearly dependent. $\exists$ nonzero $f \in \mathcal{P}(\mathbb{R})$ such that $f(T)v = 0$. Then since $(T^2 + b_1 T + c_1 I)\cdots(T^2 + b_M T + c_M I)(T - \lambda_1 I)\cdots + (T - \lambda_m I)v = 0$, $b_i, c_i \in \mathbb{R}$, $b_j^2 < 4c_j$ (by factorization of $\mathbb{R}$) each $T^2 + b_j T + c_j I$ is invertible so $(T - \lambda_1 I)\cdots(T - \lambda_n I)v = 0$. Hence some $T - \lambda_k I$ is not invertible.

Theorem 6.2.4. If $T \in \text{End}(V)$ is normal (sef adjoint), $U \subset V$ invariant under T, then

(a) $U^\perp$ is invariant under T

(b) $T|_U$ is normal (self adjoint)

(c) $T|_{U^\perp}$ is normal (self adjoint)

Proof.

(a) First, find $(T|_U)^*$. $\forall u_1, u_2 \in U$, $\langle T|_U u_1, u_2\rangle = \langle Tu_1, u_2\rangle = \langle u_1, T^* u_2\rangle = \langle u_1, P_U T^* u_2\rangle$. Hence

$(T|_U)^* = P_U T^*$.
Now, let $e_1, \dots, e_n$ be an orthonormal basis of U, consider the H-S norm of $T|_U$. $\sum_{j=1}^n ||T^* e_j||^2 = \sum_{j=1}^n ||Te_j||^2 = \sum_{j=1}^n ||T|_U e_j||^2 \sum_{j=1}^n ||P_U T^* e_j||^2$. This implies $T^* e_j$ is in U so U is invariant under T^*. For every $v \in U^\perp$, $u \in U$, $\langle Tv, u\rangle = \langle v, T^* v\rangle = 0$. So $U^\perp$ is invariant under T.

(b) $\forall u \in U$. $(P_U T^*)(T|_U)u = P_U T^* T u = P_U T T^* u = TT^* u$ (since U is invariant under T, T^*) $= TP_U T^* u = T|_U (P_U T^*) u$.

(c) since $U^\perp$ is invariant apply (b)

Theorem 6.2.5 (Spectral Theorem)**.** $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, TFAE:

(a) T is self adjoint (if $\mathbb{F} = \mathbb{R}$)
  T is normal (if $\mathbb{F} = \mathbb{C}$)

(b) V has an orthonormal basis of eigenvectors of T.

(c) T has a diagonal matrix over an orthonormal basis of of V.

Proof. $b \to c$) $\mathcal{M}(T)$ with orthonomral basis of eigenvectors is diagonal.
$c \to a$) look at $\mathcal{M}(T)$ with respect to some orthonormal basis.
If $\mathbb{F} = \mathbb{R}$, $\mathcal{M}(T) = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_i \in \mathbb{R}$. Then $\mathcal{M}(T^*) = \text{diag}(\lambda_1, \dots, \lambda_n) = \mathcal{M}(T)$ so $T^* = T$.
If $\mathbb{F} = \mathbb{C}$, $\mathcal{M}(T) = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_j \in \mathbb{C}$, $\mathcal{M}(T^*) = \text{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$. $\mathcal{M}(T)\mathcal{M}(T^*) = \text{diag}(|\lambda_1|^2, \dots, |\lambda_n|^2) = \mathcal{M}(T^*)\mathcal{M}(T)$ so $TT^* = T^*T$.
$a \to b$) Induction on $\dim V = n$. Easy for $\dim V = 0$ or 1.
Now, suppose $\dim V > 1$, the theorem holds for all W with $\dim W < \dim V$. T has an eigenvalue. Let v be a corresponding eigenvector, $||v|| = 1$. $(\text{span}(T))^\perp$ is invariant under T and $T|_{(\text{span}(v))^\perp}$ is $\begin{cases} \text{self adjoint}, & \mathbb{F} = \mathbb{R} \\ \text{normal}, & \mathbb{F} = \mathbb{C} \end{cases}$, By IH, the restriction of T on $(\text{span}(v))^\perp$ is diagonalizable by an orthonormal basis $v_1, \dots, v_{n-1}$. Now, $v, v_1, \dots, v_{n-1}$ is an orthonormal basis of eigenvectors of T.

6.2.3 7.C - Positive Operators and Isometries

Important Normal Operators: $\begin{cases} \text{self adjoint operators} \\ \text{isometries} \end{cases}$ under orthonormal basis $\begin{cases} A^t = A \\ A^t A = AA^t = I, \mathbb{F} = \mathbb{R} \quad \overline{A}^t A - A\overline{A}^t = I, \end{cases}$

Definition 6.2.6. $T \in \text{End}(V)$ is positive if T is self adjoint and $\langle Tv, v\rangle \geq 0$, $\forall v \in V$.

Example 6.2.7. Positive Operators:

(alph*) Orthogonal Projections

(alph*) $T^2 + bT + cI$, $b, c \in \mathbb{R}$, $b^2 < 4c$, T is self adjoint.

Definition 6.2.8. If $R \in \text{End}(V)$, $R^2 = T$, R is called a square root of T.

Theorem 6.2.9. $T \in \text{End}(V)$. TFAE:

(a) T is positive.

(b) T is self adjoint and all eigenvalues of $T \geq 0$.

(c) T has a positive square root.

(d) T has a self adjoint square root.

(e) $\exists R \in \text{End}(V)$ such that $T = R^*R$

Proof. $a \to b)$ T is self adjoint by assumption. If λ is an eigenvalue and v is a corresponding eigenvector, then $0 \leq \langle Tv, v\rangle = \lambda\langle v, v\rangle$ so $\lambda \geq 0$.
$b \to c)$ $\exists$ an orthonormal basis $e_1, \ldots, e_n$ such that $Te_j = \lambda_j e_j$ and $\lambda_j \geq 0$. Define R such that $Re_j = \sqrt{\lambda_j} e_j$. R is positive and $R^2 = T$.
$c \to d)$ Take the positive square root. It is self adjoint.
$d \to e)$ If $R^2 = T$, R is self adjoint, then $T = R^*R$.
$e \to a)$ First, $T^* = (R^*R)^* = R^*R = T$. Moreover, $\langle Tv, v\rangle = \langle R^*Rv, v\rangle = \langle Rv, Rv\rangle = ||Rv||^2 \geq 0$.

Theorem 6.2.10. If $T \in \text{End}(V)$ is positive, then it has a unique positive square root.

Proof. Existence: Previous theorem $(a \to c)$
Uniqueness: If R is positive, $T = R^2$, WTS $\forall\lambda \geq 0$ eigenvalues of T and $v \neq 0$ in $E(\lambda, T)$, $Rv = \sqrt{\lambda}v$. This implies uniqueness.
Suppose R is diagonalized with orthonormal basis $e_1, \ldots, e_n$ and $Re_j = \sqrt{\lambda_j} e_j$, $\lambda_j \geq 0$ and suppose $v = \sum_{i=1}^n a_j e_j$.

6.3 April 11

6.3.1 7.C - Positive Operators and Isometries

Proof (Proof of Thm 6.2.10 cont)**.** Now, $\sum_{j=1}^n \lambda_j a_j e_j = Tv = R^2 v = \sum \lambda_j a_j e_j$, hence $\sum_{j=1}^n (\lambda - \lambda_j) a_j e_j = 0$. Comparing coefficients: $a_j = 0$ if $\lambda_j \neq \lambda$ hence $Rv = \sqrt{\lambda} v$.

6.3.2 Isometries

Definition 6.3.1. $S \in \text{End}(V)$ is called an isometry if $||Sv|| = ||v||$, $\forall v \in V$.

Example 6.3.2. $S \in \text{End}(\mathbb{R}^2)$ iff its matrix under the standard basis $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Theorem 6.3.3. For $S \in \text{End}(V)$, TFAE:

(a) S is an isometry

(b) $\langle Su, Sv\rangle = \langle u, v\rangle \ \forall u, v \in V$

(c) $\forall e_1, \dots, e_n$ orthonormal list, $Se_1, \dots, Se_n$ is an orthonormal list

(d) $\exists$ orthonormal basis $e_1, \dots, e_n$ such that $Se_1, \dots, Se_n$ is an orthonormal basis.

(e) $S * S = I$

(f) $SS^* = I$

(g) S^* is an isometry

(h) S is an invertible with $S^{-1} = S^*$

Proof. $a \to c)$ $||Se_1|| = ||e_1|| = 1$, for $i \neq j$,

$$t^2 + 1 = ||e_i + te_j||^2 = ||S(e_i + te_j)||^2 = ||Se_i + Se_j||^2 = t^2 + 1 + 2\mathrm{Re}(\bar{t}\langle Se_i, Se_j\rangle)$$

for all t so $\langle Se_i, Se_j\rangle = 0$.
$c \to d)$ Any orthonormal basis suffices s
$d \to b)$ If $u = \sum a_j e_j$, $v = \sum b_j e_j$ with $e_1, \dots, e_n$ an orthonormal basis, then $S_u = \sum a_j Se_j$, $Sv = \sum b_j Se_j$ so $\langle u, v\rangle = \sum_{j=1}^n a_j \overline{b_j} = \langle Su, Sv\rangle$
$b \to e)$ $\langle S^*Su, v\rangle = \langle u, v\rangle \ \forall u, v \in V$ so $S^*S = I$
$e \to f)$ $S^*S = I$ so S is invertible and $SS^*S = S$ so multiplying by S^{-1} on the right, we get $SS^* = I$
$f \to g)$ $||S^*v||^2 = \langle SS^*v, v\rangle = \langle v, v\rangle = ||v||^2$
$g \to h)$ By previous reasoning "$(a) \to (e) \to (f)$", when $\tilde{S}$ is an isometry, $\tilde{S}$ and $\tilde{S}^*$ are invertible and $\tilde{S} = (\tilde{S}^*)^{-1}$. Take $\tilde{S} = S^*$ satisfies (h).
$h \to a)$ First, note $S^*S = I$, then $||Sv||^2 = \langle S^*Sv, v\rangle = \langle v, v\rangle = ||v||^2$.

Theorem 6.3.4. If $\mathbb{F} = \mathbb{C}$, $S \in \mathrm{End}(V)$, then S is an isometry $\leftrightarrow \exists$ an orthonormal basis of eigenvectors of S with absolute value 1.

Proof. $\to)$ S is normal. By the spectral theorem, S is diagonalized by an orthonormal basis. Since $S^*S = I$, all diagonal terms must have absolute value 1.
$\leftarrow)$ Assume the orthonormal basis is $e_1, \dots, e_n$ under which $\mathcal{M}(S) = \mathrm{diag}(a_1, \dots, a_n)$, $|a_j| = 1 \ \forall j$. Hence $\mathcal{M}(S^*S) = \mathrm{diag}(|a_1|^2, \dots, |a_n|^2) = \mathcal{M}(I)$.

Example 6.3.5. S^* is an isometry iff:
$A = \mathcal{M}(S)$ has an orthonormal basis satisfying $A^*A = AA^* = I$
Such an A is called an $\begin{cases}\text{orthogonal} \\ \text{unitary}\end{cases}$ matrix if $\mathbb{F} = \begin{cases}\mathbb{R} \\ \mathbb{C}\end{cases}$.

6.3.3 Polar Decomposition and Singular Value Decomposition

Definition 6.3.6. For positive T, let $\sqrt{T}$ be the unique positive square root of T.

Theorem 6.3.7 (Polar Decomposition). If $T \in \text{End}(V)$, $\exists$ isometry $S \in \text{End}(V)$ such that $T = S\sqrt{T^*T}$. (if T invertible, then $S = T(\sqrt{T^*T})^{-1}$: $T(\sqrt{T^*T})^{-1}(\sqrt{T^*T})^{-1}T^* = T(T * T)^{-1}T^* = TT^{-1}(T^*)^{-1}T^* = I$

Proof. Define $S_1 : \text{im}(\sqrt{T^*T}) \to \text{im}T$ by $\sqrt{T * T}v \mapsto Tv$
S_1 well defined: if $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$, then

$$0 = \langle \sqrt{T^*T}(v_1 - v_2), \sqrt{T^*T}(v_1 - v_2) \rangle = \langle T^*T * (v_1 - v_2), v_1 - v_2 \rangle = ||T(v_1 - v_2)||^2$$

isometry: $||\sqrt{T^*T}v||^2 = \langle \sqrt{T^*T}v, \sqrt{T^*T}v \rangle = \langle T^*Tv, v \rangle = \langle v, v \rangle$
So S_1 is injective, $\text{im}S_1 = \text{im}T$.
Hence if $V_1 = \text{im}(\sqrt{T^*T})$ and $V_2 = \text{im}(T)$ where $\dim V_1 = \dim V_2$ so $\dim V_1^\perp = \dim V_2^\perp$. It's possible to define S_2 to be an isometry between $V_1^\perp$ and $V_2^\perp$ by taking an orthonormal basis of each and mapping corresponding basis vectors to each other.
For $v \in V$, $v = u + w$ for $u \in V_1$, $w \in V_1^\perp$, define $Sv = S_1u + S_2w$. Then S is an isometry and $\forall v \in V$, $S\sqrt{T^*T}v = S_1\sqrt{T^*T}v = Tv$.

Note: S need not commute with $S\sqrt{T^*T}$ or have any relation.

6.3.4 Singular Value Decomposition

Definition 6.3.8. For $T \in \text{End}(V)$, the singular values of T are the eigenvalues of $\sqrt{T * T}$, where each eigenvalue λ is repeated $\dim E(\lambda, \sqrt{T * T})$ times.

Theorem 6.3.9 (SVD). For $T \in \text{End}(V)$ with singular values $s_1, \dots, s_n$ $\exists$ orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_n$ such that

$$Tv = s_1\langle v, e_1 \rangle f_1 + s_2\langle v, e_2 \rangle f_2 + \dots + s_n\langle v, e_n \rangle f_n$$

Proof. Take $e_1, \dots, e_n$ that diagonalizes $T \in \text{End}(V)$ as $\text{diag}(s_1, \dots, s_n)$. Form the polar decomposition $T = S\sqrt{T^*T}$. Assume $Se_j = f_j$, $f_1, \dots, f_n$ orthonormal basis. Now, $T(\sum a_je_j) = S\sqrt{T^*T}(\sum a_je_j) = S(\sum s_ja_je_j) = \sum s_ja_jf_j$

Corollary 6.3.10. $\forall A \in \mathbb{F}^n, m$, $\exists$ unitary/orthogonal U_1, U_2 and $\text{diag}(s_1, \dots, s_n) \in \mathbb{F}^{n,n}$ such that $s_1, \dots, s_n > 0$ and $A = U_1\text{diag}(s_1, \dots, s_n)U_2$.

Example 6.3.11. For $A = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{2,2}$, $A^tA = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$
$(1, \pm 1)$ eigenvectors of A^tA $\quad A = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$

Proposition 6.3.12. Singular values of T are the square roots of the eigenvalues of T^*T with each $\sqrt{\lambda}$ repeated $\dim E(\lambda, T^*T)$ times.

6.3.5 Ch 8: Operators on Complex Vector Spaces

V: finite dimensional vector space over $\mathbb{F}$ (no inner product) throughout Ch 8. Main Theorem: Jordan Normal Form- for $T \in \text{End}(V)$, $(V/\mathbb{C})$ has the Jordan Form under some basis:

$$\begin{pmatrix} \begin{pmatrix} \lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{pmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{pmatrix} \lambda_p & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & p \\ & & & \lambda_p \end{pmatrix} \end{pmatrix}$$

6.4 April 13

6.4.1 Jordan Form

To prove Jordon Form Thm - 2 Steps:

1. Bézout's Lemma $\rightarrow V$ is a direct sum of generalized eigenspaces
2. Structure of nilpotent operators

Lemma 6.4.1. For $T \in \text{End}(V)$, $\exists p \neq 0 \in \mathcal{P}(\mathbb{F})$ such that $p(T) = 0$

Proof. $I, T, T^2, \ldots, T^{(\dim V)^2}$ is linearly dependent.

Lemma 6.4.2 (Bézout's Lemma, special case)**.** For nonzero polynomials $p_1, \ldots, p_n \in \mathcal{P}(\mathbb{C})$, we have either:

(a) $\exists \lambda \in \mathbb{C}$ is a common root of $p_1, \ldots, p_m$ or

(b) $\exists q_1, \ldots, q_m \in \mathcal{P}(\mathbb{C})$ such that $1 = p_1q_1 + p_2q_2 + \cdots + p_mq_m$

Proof. Consider the set $S = \{g_1p_1 + \cdots + g_np_n : g_1, \ldots, g_m \in \mathcal{P}(\mathbb{C})\}$. S is closed under addition and multiplication by $\forall g \in \mathcal{P}(\mathbb{C})$.
Claim: $\exists$ a nonzero $p \in \mathcal{P}(\mathbb{C})$ such that $S = \{g \cdot p : g \in \mathcal{P}(\mathbb{C})\}$ (holds for any $\mathbb{F}$)
Proof of claim: S contains polynomials of deg ≥ 0 such as p_1. Choose $0 \neq p \in S$ such that p has the smallest possible degree. Suffices to show any $\tilde{p} \in S$ is divisible by p. Since we can write $\tilde{p} = sp + r$, $\deg r < \deg p$, $r \in S$ so by the minimality of the degree of p, $r = 0$ so p divides $\tilde{p}$.
If p in the claim is a constant, we have (b). Otherwise, by FTA, p has a root so (a) follows.

Definition 6.4.3. If $T \in \text{End}(V)$, $v \in V$ is called a generalized eigenvector corresponding to λ if $v \neq 0$ and there is a positive integer j such that $(T - \lambda I)^j v = 0$.

Example 6.4.4. If there is such a v, then λ has to be an eigenvalue.

Proof. Suppose $(T - \lambda I)^j v = 0$, $v \neq 0$. Choose a minimal k such that $(T - \lambda I)^k v = 0$. Then $(T - \lambda I)^{k-1} v \neq 0$ by $(T - \lambda I)((T - \lambda I)^{k-1} v)) = 0$ so $T - \lambda I$ is not injective, hence λ is an eigenvalue. □

Definition 6.4.5. For eigenvalue λ, all general eigenvectors of T corresponding to λ together with 0 form a general eigenspace of T corresponding to λ, denoted $G(\lambda, T)$.

Example 6.4.6. Prove that $G(\lambda, T)$ is a subspace.

Convention: if λ is not an eigenvalue, $G(\lambda, T) = \{0\}$.
Assuming $\mathbb{F} = \mathbb{C}$, by first lemma $\exists f \neq 0 \in \mathcal{P}(\mathbb{C})$ such that $f(T) = 0$.
Assuming f monic (ie. f has highest degree coefficeint 1), by FTA $f(z) = (z - \lambda_1)^{j_1} \cdots (z - \lambda_m)^{j_m}$

Proposition 6.4.7.

(a) $\ker(T - \lambda_k)^{j_k} = G(\lambda_k, T)$ and

(b) $V = G(\lambda_1, T) \oplus + \cdots + G(\lambda_m, T)$

Proof.

(a) $\ker(T - \lambda_k I)^{j_k} \subset G(\lambda_k, T)$ by def
For $v \neq 0$ in $G(\lambda_k, T)$, $\exists \tilde{j_k}$ such that $(T - \lambda_k I)^{\tilde{j_k}} v = 0$
If $\tilde{j_k} \leq j_k$, then $v \in \ker(T - \lambda_k)^{j_k}$
If $\tilde{j_k} > j_k$, set $g_k(z) = \frac{f(z)}{(z-\lambda_j)^{j_k}} = (z_1 - \lambda_1)^{j_1} \cdots \widehat{(z - \lambda_k)}^{j_k} \cdots (z - \lambda_m)^{j_m}$
Now, $g_k(T)(T - \lambda_k I)^{j_k} v = 0 \quad (1)$, $(T - \lambda_k)^{\tilde{j_k} - j_k}(T - \lambda_k)^{j_k} v = 0 \quad (2)$
By Bézout, $\exists q_1, q_2 \in \mathcal{P}(\mathbb{C})$ such that $q_1 g_k + q_2 (z - \lambda)^{\tilde{j_k} - j_k} = 1$. Then,
$q_1(T)\,(1)\ + q_2(T)\,(2)$ gives: $I(T - \lambda_k)^{j_k} v = 0$ so $v \in \ker(T - \lambda_k I)^{j_k}$.

(b) As before, let $g_k(z) \frac{f(z)}{(z-\lambda_k)^{j_z}}$
Direct Sum: If $v_1 + \cdots + v_m = 0 \quad (3)$ where each $v_k \in (G\lambda_k, T)$ then $(T - \lambda_k I)^{j_k} v_k = 0$, then $g_k(T)\tilde{v_k} - 0$, $\tilde{k} \neq k$. Applying $g_k(T)$ to (3), we get $g_k(T) v_k = 0 \quad (4)$, also $(T - \lambda_k T)^{j_k} v = 0 \quad (5)$.
By Bézouts, $\exists q_{3,k}, q_{4,k} \in \mathcal{P}(\mathbb{C})$ such that $q_{3,k}(z) + g_k(z) + q_{4,k}(z)(z - \lambda_k)^{j_k} = 1$. Applying $q_{3,k}$ to (4) and $q_{4,k}$ to (5), we see that $I v_k = 0$, $\forall k$.
Adding up to V: By Bézout, $\exists h_1, \ldots, h_m \in \mathcal{P}(\mathbb{C})$ such that $1 = \sum_{j=1}^m h_k g_k$, $\forall v \in V$, $v = \sum_{k=1}^m g - k(T) h_k(T) v$. Now, $g_k(T) w$ is such that $(T - \lambda_k)^{j_k} g_k(T) w = f(T) w = 0$ so $g_k(T) w \in G(\lambda_k, T)$. Hence each $g_k(T) h_k(T)$ on RHS is in $G(\lambda_k, T)$.

Example 6.4.8. Each $G(\lambda_k, T)$ is invariant under T.

Proof. Follows since $G(\lambda_k, T)$ is the kernel of some polynomial. □

Note: $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent (ie. some power of it is 0). Hence we need to study the structure of nilpotent operators.

Theorem 6.4.9 (Study of Nilpotent Operators)**.** Let $N \in \mathrm{End}(V)$ be nilpotent. Then $\exists v_1, \ldots, v_n \in V$, $m_1, \ldots, m_n \in \mathbb{N}$ such that

1. $N^{m_1}v_1, N^{m_1-1}v_1, \ldots, v_1, N^{m_2}v_2, N^{m_2-1}v_2, \ldots, v_2, \ldots, N^{m_n}v_n, N^{m_n-1}v_n, \ldots, v_n$ is a basis of V.
2. $N^{m_1+1}v_1 = N^{m_2+1}v_2 = \cdots = N^{m_n+1}v_n = 0$.

Example 6.4.10. Using decomposition into generalized eigenspaces and structure of nilpotent operators, we get the Jordan Form:
$\forall T, \exists \lambda_1, \ldots, \lambda_p \in \mathbb{C}$ and $m_1, \ldots, m_p > 0$ and a basis such that $\mathcal{M}(T)$ under the basis is:

$$\begin{pmatrix} \underbrace{\begin{pmatrix} \lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{pmatrix}}_{m_1} \left.\vphantom{\begin{pmatrix} \\ \\ \\ \\ \end{pmatrix}}\right\} m_1 & & 0 \\ & \ddots & \\ 0 & & \underbrace{\begin{pmatrix} \lambda_p & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_p \end{pmatrix}}_{m_p} \left.\vphantom{\begin{pmatrix} \\ \\ \\ \\ \end{pmatrix}}\right\} m_p \end{pmatrix}$$

Proof (Proof of Thm 6.4.9)**.** Induct on $\dim V$. True for $\dim V = 0$.
Assuming, we know the claim is true for $\dim W < \dim V$.
N nilpotent $\to$ N is not invertible $\to$ $N \subsetneq V$.
$\mathrm{im}N$ invariant under N, $N|_{\mathrm{im}N}$ is nilpotent too. By IH, $N|_{\mathrm{im}N}$, $\exists u_1, \ldots, u_n \in \mathrm{im}N$, $l_1, \ldots, l_r \in \mathbb{N}$ such that $N^{p_1}u_1, \ldots, u_1, N^{p_2}u_2, \ldots, u_2, \ldots, N^{p_r}, \ldots, u_r$ is a basis of $\mathrm{im}N$ and $N^{l_i+1}u_i = 0$, $\forall i$. Since each $u_i \in \mathrm{im}N$, we have $v_1, \ldots, v_r$ such that $u_i = Nv_i$, $\forall i$. Define $m_i = l_i = 1$.

6.5 April 18

6.5.1 Jordan Form

Proof (Proof of Thm 6.4.9 (cont.))**.**
Claim 1: $N^{m_1}v_1, N^{m_1-1}v_1, \ldots, v_1, \ldots, N^{m_r}v_r, \ldots, v_r$ is linearly independent.

Proof. If $\exists$ a linear combination of these vectors that equals 0, applying N to it, we get a linear combination of $N^{l_1}u_1, \ldots, u_1, N^{l_2}u_2, \ldots, u_2, \ldots$, and some 0's. Hence the coefficeints of $N^{m_1-1}v_1, \ldots, v_1$, $N^{m_2-1}v_2, \ldots, v_2, \ldots$ are all 0. Look at the remaining $a_1N^{m_1}v_1 + a_2N^{m_2}v_2 + \cdots + a_rN^{m_r}v_r = 0$. Since it is contained within $\mathrm{im}N$, $a_1 = a_2 = \cdots = 0$.

Let $U = \mathrm{span}(N^{m_1}v_1, \ldots, v_1, N^{m_2}v_2, \ldots, v_2, \ldots)$. Extend $N^{m_1}v_1, \ldots, v_1, N^{m_2}v_2, \ldots, v_2, \ldots$ to a basis

$N^{m_1}v_1, N^{m_1-1}v_1, \ldots, v_1, \ldots, N^{m_r}v_r, \ldots, v_r, w_1, \ldots, w_s$.
Claim 2: $\forall u \in V$, $\exists x \in U$ such that $Nw = Nx$

Proof. Since $\mathrm{im}N|_U = \mathrm{im}N$

We choose x_i as in claim 2 for each w_i/ Let $v_{r+i} = w_i - x_i$. Then $Nv_{r+i} = 0$. Hence $N^{m_1}v_1, N^{m_1-1}v_1, \ldots, v_1, \ldots, N^{m_r}v_r, \ldots, v_r, v_{r+1}, \ldots, v_n$ is a basis satisfying the desired conclusion.

Remark 6.5.1. The matric of N under the basis in the previous thm is

$$\begin{pmatrix} \underbrace{\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}}_{m_1} \left.\vphantom{\begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}}\right\} m_1 & & 0 \\ & \ddots & \\ 0 & & \underbrace{\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}}_{m_p} \left.\vphantom{\begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}}\right\} m_p \end{pmatrix}$$

6.5.2 Genralized Eigenvectors

Theorem 6.5.2.

- For $T \in \mathrm{End}(V)$, $0 \subset \ker T^0 \subset \ker T^1 \subset \ker T^2 \subset \cdots$
- If $\ker T^m = \ker T^{m+1}$, then $\ker T^m = \ker T^{m+1} = \ker T^{m+2} = \cdots$

Proof (Proof of Second Claim)**.** Suffices to prove $\ker T^{m+1} = \ker T^{m+2}$. $\forall v \in \ker T^{m+2}$, $T^{m+2}v = 0$ so $T^{m+1}(Tv) = 0$ so $T^m(Tv) = 0$ so $T^{m+1}v = 0$.

Corollary 6.5.3. $\ker T^{\dim V} = \ker T^{\dim V+1} = \cdots$

Proof. Otherwise, $0 \subsetneq \ker T^0 \subsetneq \ker T^1 \subsetneq \cdots \subsetneq \ker T^{\dim V} \subsetneq \ker T^{\dim V+1}$, a contradiction since the dimension increases by at least 1 each inclusion.

Proposition 6.5.4. $V = \ker T^{\dim V} \oplus \mathrm{im}T^{\dim V}$

Proof. Direct Sum: If $v \in \ker T^{\dim V} \cap \mathrm{im}T^{\dim V}$, $\exists u$ such that $v = T^{\dim V}u$. Also, $T^{\dim V} = T^{\dim V+1}u = 0$ so $T^{\dim V}u = 0 = v$.

Direct sum $= V$ follows by couting dimension.

Proposition 6.5.5. If $T \in \mathrm{End}(V)$, $G(\lambda, T) = \ker(T - \lambda I)^{\dim V}$.

Proof. $\forall v \in G(\lambda, T)$, $\exists m$ such that $(T - \lambda I)^m v = 0$, by above propositions $(T - \lambda)^{\dim V} v = 0$.

Proposition 6.5.6. if $\mathbb{F} = \mathbb{C}$, $V = \bigoplus_{G(\lambda,T) \neq \{0\}} G(\lambda, T)$

Proposition 6.5.7. If N is nilpotent, then $N^{\dim V} = 0$.

6.5.3 8.B - Decompositions of an Operator

Proposition 6.5.8. If $T \in \mathrm{End}(V)$, $p \in \mathcal{P}(\mathbb{F})$, then $\ker p(T)$ and $\mathrm{im} p(T)$ are invariant under T.

Proof. If $p(T)v = 0$, then $Tp(T)v = 0$ so $p(T)(Tv) = 0$.

Theorem 6.5.9.

(a) Each $G(\lambda_j, T)$ is invariant under T.

Proof. Follows since $G(\lambda_j, T) = \ker(T - \lambda_j I)^{\dim V}$

(b) Each $(T - \lambda_j I)|_{G(\lambda_j, T)}$ is nilpotent.

Corollary 6.5.10. If $\mathbb{F} = \mathbb{C}$, $T \in \mathrm{End}(V)$, then $\exists$ a basis of V of generalized eigenvectors of T.

Definition 6.5.11. For $T \in \mathrm{End}(V)$, the multiplicity of each eigenvalue λ of T is $\dim G(\lambda, T) = \dim \ker(T - \lambda I)^{\dim V}$.

Example 6.5.12. If $\mathcal{M}(T) = \begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}$, then the multiplicity of the eigenvalue 6 is 2 and of the eigenvalue 7 is 1.

Corollary 6.5.13. If $\mathbb{F} = \mathbb{C}$, the multiplicities of all eigenvalues of T add up to $\dim V$.

Definition 6.5.14. A block diagonal matrix is a matrix of the from

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & A_m & \end{pmatrix}$$

where $A_1, \ldots, A_m$ are square matrices lying along the diagonal and all othe entries equal 0.

Jordan Block: $\begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_1 \end{pmatrix}$

Example 6.5.15. Use this notion to describe the Jordan Form.

Theorem 6.5.16. If $N \in \mathrm{End}(V)$ is nilpotent, $I + N$ has a square root.

Proof. A formal power series is $\sum_{n=0}^{\infty} a_n x^n$, $a_n \in \mathbb{F}$. We define the product

$$(\sum a_n x_n)(\sum b_n x_n) = \sum_{n \geq 0} (\sum_{0 \leq k \leq n} a_k b_{n-k}) x^n$$

Claim:

Proof. Take $a_0 = 1$ and define the coefficeints inductively.

Then, we can check $(\sum a_n N^n)^2 = 1 + n$ where $\sum a_n N^n$ is a finite sum since N is nilpotent.

Theorem 6.5.17. When $\mathbb{F} = \mathbb{C}$, every invertible $T \in \mathrm{End}(V)$ has a square root.

Proof. Write $V = \bigoplus_{i=1}^m G_j$ where $G_j = G(\lambda_j, T)$, λ_j distinct, nonzero.
$\frac{1}{\lambda_j} T|_{G_j} = I + N_j$ so $\exists S_j \in \mathrm{End}(G_j)$ such that $S_j^2 = \frac{1}{\lambda_j} T|_{G_j}$. Take $\mu_j \in \mathbb{C}$ such that $\mu_j^2 = \lambda_j$ and let $R_j = \mu_j S_j$. Then $R_j^2 = T|_{G_j}$. Define $R(\sum_{i=1}^m a_j v_j) = \sum_{j=1}^m a_j R_j v_j$ where $v_j \in G_j$. Then $R^2 = T$.

6.5.4 8.C - Minimal Polynomials

Definition 6.5.18. A monic polynomial is a polynomial with highest degree coefficeint equal to 1.

Example 6.5.19. $x + 1, x^2 + 1, x^{10} + 6x + 5$ are monic.

Proposition 6.5.20. For $T \in \mathrm{End}(V)$, $\exists$ a unique monic polynomial p of smallest degree such that $p(T) = 0$. It is called the minimal polynomial of T.

6.6 April 20

6.6.1 8.C - Minimal Polynomials

Proof. Consider $N(T) = \{q \in \mathcal{P}(\mathbb{F}) : q(T) = 0\}$. $N(T) \neq \{0\}$ and is closed under addition and multiplication (by elements of $\mathcal{P}(\mathbb{F})$). Take a nonzero $p \in N(T)$ of smallest degree. WLOG p is monic. For every $q \in N(T)$, $q = sp + r$ with $\deg r < \deg p$, but $q(T) = 0$, $p(T) = 0$, $(sp)(T) = 0$, then $r(T) = 0$. Since $\deg r < \deg p$, $r = 0$. Hence $q = sp$, If q is also monic, then either $s = 1$ so $q = p$ or $\deg q > \deg p$.

Corollary 6.6.1. If $q(T) = 0$, then q is a multiple of the minimal polynomial of T.

Example 6.6.2. For the matrix $\begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{n-2} \\ & & 1 & -a_{n-1} \end{pmatrix}$ in $\mathbb{F}^{n,n}$, the minimal polynomial of the corresponding operator is $z^n + a_{n-1}z^{n-1} + a_{n-2}^{n-2} + \cdots + a + 0$
First, note that $Tv_1 = v_2, Tv_2 = v_3, \ldots, Tv_n = -a_0v_1 - \cdots - a_{n-1}v_n$. Hence $(T^n + a_{n-1}T^{n-1} + \cdots + a_0I)v_1 = 0$.
Now, $p(T)v_2 = p(T)Tv_1 = Tp(T)v_1 = 0, \ldots$.

Example 6.6.3. If $\mathcal{M}(T) = \begin{pmatrix} 0 & 1 & \\ & 0 & \\ & & 2 \end{pmatrix}$, then T has minimal polynomial $z^2(z-2)$

If $\mathcal{M}(T) = \begin{pmatrix} 0 & 1 & \\ & 1 & \\ & & 2 \end{pmatrix}$, then T has minimal polynomial $z(z-1)(z-2)$

Theorem 6.6.4. For $T \in \text{End}(V)$, the zeros of the minimal polynomial are precisely the eigenvalues of T. (Doesn't say anything about multiplicity).

Proof. Let the minimal polynomial be $p(z)$. If $p(\lambda) = 0$, then $p(T) = (T - \lambda I)q(T)$. If $(T - \lambda I)$ is invertible, then $q(T) = 0$, contradicting the minimality of the degree of p. Thus, $(T - \lambda I)$ is not invertible.
If $T - \lambda I$ is not invertible, then $\exists v \neq 0$ such that $(T - \lambda I)v = 0$, and $p(T)v = 0$. Let $p(z) = (z - \lambda I)q(z) + r$, r constant. Then applying z, we get $0 = 0 + r$ so $r = 0$. Thus $p(\lambda) = 0$.

Example 6.6.5. if $\mathcal{M}(T)$ is some Jordan Form what is the minimal polynomial of it?
eg. $\begin{pmatrix} 0 & 1 & & & \\ & 0 & & & \\ & & 0 & 1 & \\ & & & 0 & \\ & & & & 1 \end{pmatrix}$, for this example its is $z^2(z-1)$

Definition 6.6.6. A Jordan basis for $T \in \text{End}(V)$ is some $v_1, \ldots, v_n$ such that the matrix of T under it is in Jordan Form.

Example 6.6.7. If $\mathcal{M}(T) = \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}$ under a basis, then $\mathcal{M}(T)$ can't be $\begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{pmatrix}$ under another basis. (Consider T^2)

Remark 6.6.8. If $\mathcal{M}(T)$ is a jordan form, then we can read off the eigenvalues and their multiplicities immediately.

Ch 10 - Trace and Determinant

V: finite dimensional vector space, $\dim V > 0$ throughout this chapter.

6.6.2 10.A - Trace

Recall:

- I means the identity operator or the identity matrix
- We denote the inverse of a matrix A, A^{-1}
- Change of basis formula: If $A = \mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$, then

$$\mathcal{M}(T, (u_1, \ldots, u_n)) = A^{-1}\mathcal{M}(T, (v_1, \ldots, v_n)A)$$

Definition 6.6.9. FOr $A \in \mathbb{F}^{n,n}$ $(n \geq 1)$, then the trace of A is

$$\operatorname{tr}(A) = A_{1,1} + A_{2,2} + \cdots + A_{n,n}$$

Proposition 6.6.10. For $A \in \mathbb{F}^{n,m}$, $B \in \mathbb{F}^{m,n}$, $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

Proof. $\operatorname{tr}(AB) = \sum_{j=1}^{n}(AB)_{j,j} = \sum_{j=1}^{n}\sum_{k=1}^{m} A_{j,k}B_{k,j} = \sum_{k=1}^{m}\sum_{j=1}^{n} A_{j,k}B_{k,j} = \sum_{k=1}^{m}(BA)_{k,k} = \operatorname{tr}(BA)$.

Corollary 6.6.11. For $A \in \mathbb{F}^{n,n}$, $S \in \mathbb{F}^{n,n}$ invertible, $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}(A)$

Proof. $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}(ASS^{-1}) = \operatorname{tr}(A)$

Definition 6.6.12. For $T \in \operatorname{End}(V)$, define the trace of T, $\operatorname{tr}(T)$ to be $\operatorname{tr}(\mathcal{M}(T))$ (under arbitrary basis)

Well defined since by the previous corollary and the change of basis formula, it is independent of choice of basis.

Proposition 6.6.13. if $\mathbb{F} = \mathbb{C}$, then $\operatorname{tr}(T) = \sum$ eigenvalues of T (counting multiplicity)

Proof. Find a jordan basis, $\mathcal{M}(T)$ has diagonal elements being the eigenvalues of T (with multiplicity)

Proposition 6.6.14. For $T \in \mathrm{End}(V)$,

(1) $\mathrm{tr}(T^{-1} = \mathrm{tr}(T))$

(2) $\mathrm{tr}(T_1 + T_2) = \mathrm{tr}(T_1) + \mathrm{tr}(T_2)$

(3) $\mathrm{tr}(kT) = k \cdot \mathrm{tr}(T)$, $k \in \mathbb{F}$

For any $A, A_1, A_2 \in \mathbb{F}^{n,n}$

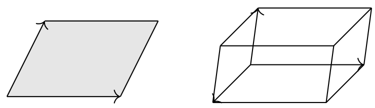
1. $\mathrm{tr}A = \mathrm{tr}A^t$
2. $\mathrm{tr}(A_1 + A_2) = \mathrm{tr}(A_1) + \mathrm{tr}(A_2)$
3. $\mathrm{tr}(kA) = k \cdot \mathrm{tr}(A)$, $k \in \mathbb{F}$

Remark 6.6.15. If V is an innter product, $\langle T_1, T_2 \rangle_{\mathrm{HS}} = \mathrm{tr}(T_1 T_2^*)$ is an innter product which gives the H-S norm.

Example 6.6.16. $\mathrm{tr}(T) \geq 0$ if T is positive.

6.6.3 10.B - Determinants

Motivation: Signed area or volume of a parralelipiped.



Each operator in $\mathrm{End}(\mathbb{R}^n)$ change the signed volume of the parralelipiped proportionally.

Definition 6.6.17. det: $\mathbb{F}^{n,n} \to \mathbb{F}$ is a function satisfying

(i) Multilinearity:

$$\det\big((v_1)\,(v_2)\,\cdots\,(a_1 u + a_2 w)\,\cdots\,(v_n)\big) = a_1 \det\big((v_1)\,(v_2)\,\cdots\,(u)\,\cdots\,(v_n)\big) + a_2 \det\big((v_1)\,(v_2)\,\cdots\,(2)\,\cdots\,(v_n)\big)$$

(ii) Alternating:

$$\det\big(\cdots(v)\cdots(v)\cdot\big)$$

(iii) $\det(I) = 1$

Proposition 6.6.18. det is unique and well defined.

Proof. If det satisfies $(i), (ii), (iii)$ then by multilinearity, a general determinant of A is an explicit linear combination of $\det(e_{k_1}, e_{k_2}, \ldots, e_{k_n})$ where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{F}^n$. If $\exists$ repetitions $k_1, \ldots, k_n$ then $\det = 0$. Otherwise $k_1, \ldots, k_n$ is a permutation of $1, \ldots, n$. By (i) + (ii),

$$\det(\cdots(v+w)\cdots(v+w)\cdots) - \det(\cdots(v)\cdots(v)\cdots) - \det(\cdots(w)\cdots(w)\cdots) = 0$$

so

$$\begin{aligned}
&\det(\cdots(v+w)\cdots(v+w)\cdots) - \det(\cdots(v)\cdots(v)\cdots) - \det(\cdots(w)\cdots(w)\cdots)\\
&= \det(\cdots(v)\cdots(v+w)\cdots) - \det(\cdots(v)\cdots(v)\cdots) + \det(\cdots(w)\cdots(v+w)\cdots) - \det(\cdots(w)\cdots(w)\cdots)\\
&= \det(\cdots(v)\cdots(w)\cdots) + \det(\cdots(w)\cdots(v)\cdots)\\
&= 0
\end{aligned}$$

so

$$\det(\cdots(v)\cdots(w)\cdots) = -\det(\cdots(w)\cdots(v)\cdots)$$

6.7 April 25

6.7.1 10.B - Determinants

Definition 6.7.1. Call $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ a permutation if it is a bijection. σ can also be denoted by a list $(m_1 \ldots m_n)$ wwwehre each number in $1, \ldots, n$ shows up once. Let $\mathrm{perm}(n)$ (or S_n) denote the set of all permuations on $\{1, \ldots, n\}$.

Definition 6.7.2. For $\sigma \in \mathrm{perm}(n)$, the sign (or signature) of σ, $\mathrm{sign}(\sigma)$ is defined to be

$$(-1)^{|\{i,j:1\leq i<j\leq n,\, \sigma^{-1}(i)>\sigma^{-1}(j)\}|} = (-1)^{|\{i,j:1\leq i<j\leq n,\, \sigma(i)<\sigma(j)\}|}$$

Example 6.7.3. When $n = 3$, if $\sigma = (1\,2\,3)$, then $\mathrm{sign}(\sigma) = -1$.
If $\sigma = (2\,1\,3)$, then $\mathrm{sign}(\sigma) = 1$.

Lemma 6.7.4. Swapping 2 entries of σ results in a change of sign.

Proof. Assuming we are swapping m_j and m_k in $(\cdots m_j \cdots m_k \cdots)$. We say (a, b) is an inversion if $a > b$. For every number m_l such that $j < l < k$, (m_j, m_l) is an iversion iff (m_l, m_j) is not and (m_l, m_k) is an iversion iff (m_k, m_l) is not. Pairing (m_j, m_l) and (m_l, m_k) for each l, we see these terms contribute no change in parity. Now, exactly one of (m_j, m_k) and (m_k, m_j) is an inversion so the parity changes because of this. (By above there are no other changes in parity)

Proof (Proof of Existence of Determinant)**.** (Inspired By Uniqueness)
Claim: $\det A$ can be defined by

$$\det A = \sum_{\sigma\in\mathrm{perm}(n)} (\mathrm{sign}(\sigma)) \cdot A_{\sigma(1),1}A_{\sigma(2),2}\cdots A_{\sigma(n),n}$$

eg. $A \in \mathbb{F}^{3,3}$: $\begin{pmatrix} A_{1,1} & A_{1,2} & \boxed{A_{1,3}} \\ A_{2,1} & \boxed{A_{2,2}} & A_{2,3} \\ \boxed{A_{3,1}} & A_{3,2} & A_{3,3} \end{pmatrix}$, one of the terms in the det is $(-1)A_{3,1}A_{2,2}A_{1,3}$.
The expression satisfies (i) and (iii).

For (ii), WLOG assume $n \geq 2$, if jth and $k$the columns are identical, then we can pair all σ's in perm(n) into (σ_1, σ_2) such that σ_1 and σ_2 are in the same pair if the only differ ni the jth and kth entries. The signs in each pair are different (by lemma) and for each σ,

$$A_{\sigma(1),1} \cdots A_{\sigma(n),n} = A_{\sigma(1),1} \cdots A_{\sigma(k),j} \cdots A_{\sigma(j),k} \cdots A_{\sigma(n),n}$$

So the contribution to the determinant by each pair together is 0. Thus the determinant is 0.

Corollary 6.7.5. If $f((v_1), \dots, (v_n)))$ satisfies (i) and (ii), then $\exists c \in \mathbb{F}$ such that $f((v_1), \dots, (v_n)) = c \cdot \det((v_1), \dots, (v)n)$. More specifically, $c = f(I)$.

- The set of all alternating multinear forms has dimension 1.

Algorithm for Computing the Det: Column Reduction

- Subtracting $\lambda\cdot$ column j from column k, $(k \neq j)$, doesn't change determinant.

eg.

$$\begin{aligned}\det\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\end{pmatrix} &= \begin{pmatrix}1 & 1 & 1\\ 2/3 & 2/3 & 0\\ 1 & 4 & 9\end{pmatrix}\\ &= \begin{pmatrix}8/9 & 5/9 & 0\\ 2/3 & 2/3 & 0\\ 1 & 4 & 9\end{pmatrix}\\ &= \begin{pmatrix}1/3 & 0 & 0\\ 2/3 & 2/3 & 0\\ 1 & 4 & 9\end{pmatrix}\\ &= \begin{pmatrix}1/3 & 0 & 0\\ 0 & 2/3 & 0\\ 0 & 0 & 9\end{pmatrix} = 2\end{aligned}$$

Example 6.7.6. $\det A = \det A^t$

Example 6.7.7. $\det\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix} = -\det\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} = -\det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} - -1$

Example 6.7.8 (Hw Problem)**.** $\det\begin{pmatrix}A & C\\ 0 & B\end{pmatrix} = (\det A)\cdot(\det B)$ where A, B square matrices.

Corollary 6.7.9.

$$\det\begin{pmatrix}A_{1,1} & & & *\\ & A_{2,2} & & \\ & & \ddots & \\ 0 & & & A_{n,n}\end{pmatrix} = A_{1,1}A_{2,2}\cdots A_{n,n}$$

Proposition 6.7.10. Let $A, B \in \mathbb{F}^{n,n}$,

- If B is obtained by swapping two rows or columns of A, then $\det B = -\det A$.
- If columns (or rows) of A are linearly dependent, then $\det(A) = 0$.

Key Properties of det:

Theorem 6.7.11. Let $A, B \in \mathbb{F}^{n,n}$

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

(2) A is invertible $\iff$ $\det A \neq 0$

Proof.

(a) For column vectors $(v_1), \ldots, (v_n)$, consider $f(v_1, \ldots, v_n) = \det(Av_1, \ldots, Av_n)$. f is multinear, alternating, so by previous corollary $f(v_1, \ldots, v_n) = c \cdot \det(v_1, \ldots, v_n)$. Take $v_1, \ldots, v_n$ to be the std basis then,

$$\begin{aligned} f(e_1, \ldots, e_n) &= \det(Ae_1, \ldots, Ae_n) \\ &= \det(A_{\cdot,1}, \ldots, A_{\cdot,n}) \\ &= \det A \end{aligned}$$

so $c = \det A$. Taking $v_1, \ldots, v_n$ to be the columns of B, it follows that if $\det(A)\det(B) = \det(AB)$.

(b) $\rightarrow$) by (a), $\det(A)\det(A^{-1}) = 1$. Hence $\det(A) \neq 0$
$\leftarrow$) If $\det(A) \neq 0$, view A as a linear map in $\text{End}(\mathbb{F}^n)$. A maps standard basis to columns of A which are linearly independent. Hence, A is invertible as a map, and as a matrix.

Definition 6.7.12. For $T \in \text{End}(V)$, define $\det(T) = \det(\mathcal{M}(T))$ under any basis.

It is independent of basis since $\det(S^{-1}AS) = \det(S^1)\det(A)\det(S) = \det(A)$.

Corollary 6.7.13. $T \in \text{End}(V)$ is invertible iff $\det(T) \neq 0$.

Example 6.7.14. $\det T = \det T' = \overline{\det T^*}$

Proposition 6.7.15. If $\mathbb{F} = \mathbb{C}$, $T \in \text{End}(V)$, then $\det T$ is the product of all eigenvalues counting multiplicity.

Proof. Take a Jordan form.

Theorem 6.7.16. If V is an inner product space, $T \in \text{End}(V)$ is an isometry, then $|det(T)| = 1$.

Proof. $(\det T)(\det T^*) = 1 \leftrightarrow |\det T|^2 = 1$.

Theorem 6.7.17. If V is an inner product space, then $|detT| = det(\sqrt{T^*T})$.

Proof. $|detT|^2 = \det T \cdot \det T^* = \det T^*T = \det(\sqrt{T^*T})^2$. Moreover, $\det(\sqrt{T^*T}) \geq 0$.

Definition 6.7.18.

(i) For $A \in \mathbb{F}^{n,n}$, define the characteristic polynomial to be $\det(zI - A)$, a polynomial in z.

(ii) For $T \in \operatorname{End}(V)$, define the characteristic polynomial of T to be $\det(zI - \mathcal{M}(I))$ with respect to any basis.

6.8 April 27

6.8.1 10.B - Determinants

characteristic polynomial well defined, independent of basis since $\det(zI - S^{-1}AS) = \det(S^{-1})\det(zI - A\det(S)$.

Example 6.8.1. Characteristic polynomial of T is monic, with $\deg = \dim V$ since the only permutation contributing z^n is the identity permutation.

Theorem 6.8.2. Assuming $\mathbb{F} = \mathbb{C}$, for $T \in \operatorname{End}(V)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$ with multiplicities $d_1, \ldots, d_m$, respectively. The characteristic polynomial of T is $(z - \lambda_1)d_1 \cdots (z - \lambda_m)d_m$

Proof. Make $\mathcal{M}(T)$ is the Jordan form:

$$\begin{pmatrix} \underbrace{\begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{matrix}}_{d_1} & & & & & & * \\ & \lambda_2 & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & & \lambda_m & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_m \end{pmatrix}$$

Then,

$$zI - \mathcal{M}(T) = \begin{pmatrix} \underbrace{\begin{matrix} z-\lambda_1 & & \\ & \ddots & \\ & & z-\lambda_1 \end{matrix}}_{d_1} & & & & & * \\ & z-\lambda_2 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & z-\lambda_m & \\ 0 & & & & & \ddots \\ & & & & & & z-\lambda_m \end{pmatrix}$$

The theorem follows by taking the determinant(product of diagonal entries)

Proposition 6.8.3. For $T \in \text{End}(V)$, the characteristic polynomial of T is $z^n - \text{tr}(T)z^{n-1} + \cdots + (-1)^n \det(T)$ where $n = \dim V$.

Proof. Let $A = \mathcal{M}(T)$ under some basis. Then,

$$ZI - A = \begin{pmatrix} z - A_{1,1} & -A_{1,2} & \cdots & -A_{1,n} \\ -A_{2,1} & z - A_{2,2} & \cdots & -A_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ -A_{n,1} & -A_{n,2} & \cdots & z - A_{n,n} \end{pmatrix}$$

The contribution to the constant term is $\det(-A) = (-1)^n \det(A)$.
The contribution to the coefficient of z^{n-1} is $-A_{1,1} - A_{2,2} - \cdots - A_{n,n} = -\text{tr}(A)$

Theorem 6.8.4 (Cayley-Hamilton)**.** Suppose $T \in \text{End}(V)$ with characteristic polynomial q, then $q(T) = 0$.

Proof. Take any $v \neq 0$. It suffices to prove $q(T)v = 0$. Let $m \geq 1$ be the least integer such that $v, Tv, \ldots, T^{m-1}v$ is linearly independent. Then $T^m v + a_{m-1}T^{m-1} + \cdots + a_0 v = 0$. $U = \text{span}(v, Tv, \ldots, T^{m-1}v)$

is invariant under T. Extend $v, Tv, \ldots, T^{m-1}v$ to a basis of V. $\mathcal{M}(T)$ looks like:

$$\left(\begin{array}{ccccc|c} 0 & & & & -a_0 & \\ 1 & 0 & & & -a_1 & \\ & 1 & \ddots & & -a_2 & * \\ & & \ddots & & \vdots & \\ & & & 0 & -a_{n-2} & \\ & & & 1 & -a_{n-1} & \\ \hline & & & 0 & & * \end{array}\right)$$

Hence, $\mathrm{char}(T|_U)$ divides $\mathrm{char}(T)$, but $\mathrm{char}(T|_U) = z^m + a_{m-1}z^{m-1} + \cdots + a_0 := q_1(z)$. Now, $q_1(T)v = 0$ and q_1 divides q so $q(T)v = 0$.

Corollary 6.8.5. The minimal polynomial of T divides the characteristic polynomial of T.

6.8.2 Multilinear Maps and Tensor Products

Let U, V, W be finite dimensiononal vector spaces

Definition 6.8.6. A bilinear map is a map $f : V \times W \to U$ such that $f(\cdot, w)$ and $f(v, \cdot)$ are both linear. Similarly, define a k-linear form/functional (when $U = \mathbb{F}$)

Want: Define an $\mathbb{F}$- vector space $V \otimes W$ such that the bilinear forms on $V \times W \leftrightarrow$ the linear functions of $V \otimes W$.

Proposition 6.8.7. For V, W, $\exists(\mathcal{T}, g)$ such that $\mathcal{T}$ is an $\mathbb{F}$ vector space, $g : V \times W \to \mathcal{T}$ is a bilinear map satisfying the universal property: $\forall\ U/\mathbb{F}$ and bilinear maps $f : V \times W \to U$, $\exists!$ linear map $\tilde{f} : \mathcal{T} \to U$ such that $f = \tilde{f} \circ g$.

$$\begin{tikzcd} & \mathcal{T} \arrow[dr, dashed, "\tilde{f}"] & \\ V \times W \arrow[ur, "g"] \arrow[rr, "f"] & & U \end{tikzcd}$$

Moreover, if $(\mathcal{T}_1, g)_1$ and $(\mathcal{T}_2, g_2)$ satisfy the above property there $\exists!$ (canonical) isomorphism $j : \mathcal{T}_1 \to \mathcal{T}_2$ such that $j \circ g_1 = g_2$.

Proof (Existence Proof 1)**.** Take a basis $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$. Define an abstract vector space with basis $v_1 \otimes w_1, v_2 \otimes w_2, \ldots, v_n \otimes w_m$ and

$$g(\sum a_j v_k, \sum b_k w_k) = \sum a_j b_k v_k \otimes w_k$$

Proof (Existence Proof 2)**.** Consider $\mathbb{F}^{V \times W} = \bigoplus_{\substack{v \in V \\ w \in W}} \mathbb{F}_{v,w}$. Here, we let $\bigoplus$ denote the abstract direct sum where the elements can be thought of as formal sums of elements form eaceh $F_{u,w}$ where there are

only finitely many nonzero terms.
Let $X \subset F^{V\times W}$ be the subspace spanned by all vectors of the for: $(v+v',w)-(v,w)-(v',w)$, $(v,w+w')-(v,w)-(v,w')$, $(av,w)-a(v,w)$, $(v,aw)-a(v,w)$. Then, we can take $\mathbb{F}^{V\times W}/X$ and $g(v,w)$ ="formal sum with 1 on $\mathbb{F}_{v,w}$, 0 everywhere else."

Proof (Uniqueness)**.** Suppose we have $(\mathcal{T}_1, g_1)$ and (T_2, g_2) satisfying the universal property. First, we observe that by the uniqueness of $\tilde{f}$ we must have $\mathcal{T}_i = \text{span}(\text{im} g_i)$. (Otherwise each $\tilde{f}$ could have multiple valid extensions to $\mathcal{T}_i$). Next since g_1 and g_2 are bilinear maps from $V\times W$, by the universal property, there are unique linear maps f_1 and f_2 such that $g_2 = f_1\circ g_1$ and $g_1 = f_2\circ g_2$. This implies that $g_1 = f_2\circ f_1 g_1$ and $g_2 = f_1\circ f_2\circ g_2$ so $f_2\circ f_1$ is the identity of $\text{im} g_1$ and hence is the identity on $\mathcal{T}_1$. Similarly, $f_1\circ f_2$ is the identity on $\mathcal{T}_2$ so $f_2 = f_1^{-1}$. Thus, f_1 is the desired (unique) isomorphism from $\mathcal{T}_1$ to cT_2.

Notation: Use $v\otimes w$ to denote $g(v,w)$.

Example 6.8.8. $(v_1+v_2)\otimes(w_1+w_2) = v_1\otimes w_1 + v_2\otimes w_1 + v_1\otimes w_2 + v_2\otimes w_2$.

Note: not everything in $V\otimes W$ is of the form $v\otimes w$.

Proposition 6.8.9. A k-linear form on $V\times\cdots\times V \leftrightarrow$ a linear functional on $\underbrace{V\otimes\cdots\otimes V}_{k\text{ times}} := V^{\otimes k}$.

Proof. Consider $\psi : f\mapsto\varphi$ such that $f(v_1,\dots,v_n) = \varphi(v_1\otimes\cdots\otimes v_n)$.

Remark 6.8.10. If $v_1,\dots,v_n$ forms a basis of V, then $\underbrace{v_i\otimes v_j\otimes v_k}_{n^3\text{ terms}}$ forms a basis of $V^{\otimes 3}$

Proposition 6.8.11. $\text{Hom}(V,W)\cong V^*\otimes W$ by $\phi : \varphi\otimes w\mapsto T$ such that $T(v) = \varphi(v)w$.
Define the evaluation map (linear)- ev: $V^*\otimes V\to\mathbb{F}$ by $\varphi\otimes v\mapsto\varphi(v)$. Then $\forall T\in\text{End}(V)$,

$$\text{tr}(T) = \text{ev}(\phi^{-1}(T))$$

Proof. For an arbitrary basis $v_1,\dots,v_n$ of V, suppose $Tv_i = \sum_{j=1}^n a_{i,j}v_j$. Observe that $\text{tr}(T) = \sum_{j=1}^n a_{j,j}$. Now, let $\varphi_1,\dots,\varphi_n$ be a dual basis of $v_1,\dots,v_n$. Observe that

$$\phi^{-1}(T) = \sum_{j=1}^n \varphi_j\otimes Tv_j$$

So computing $ev(\phi^{-1}(T))$, we see

$$\text{ev}(\phi^{-1}(T)) = \text{ev}\Big(\sum_{j=1}^n\varphi_j\otimes Tv_j\Big) = \sum_{j=1}^n\varphi_j(Tv_j) = \sum_{j=1}^n a_{j,j} = \text{tr}(T)$$

Definition 6.8.12. A k-linear form $f : \underbrace{V \times \cdots \times V}_{k \text{ times}} \to \mathbb{F}$ is called alternating if

$$f(v_1, \ldots, v, \ldots, v, \ldots, v_n) = 0.$$

One can define the wedge product $\bigwedge^k V$ such that linear functionals of $\bigwedge^k V \leftrightarrow$ alternating k-linear maps of V. Construction: In $V^{\otimes k}$, consider the subspace $Y = \operatorname{span}\{v_1 \otimes \cdots \otimes v_k \mid \text{two } v_j\text{'s are equal}\}$. $\bigwedge^k V = V^{\otimes k}/Y$. We use $v_1 \wedge \cdots \wedge v_n$ to denote $v_1 \otimes \cdots \otimes v_n + Y$.

Proposition 6.8.13. If $\dim V = n$, then $\bigwedge^n V$ is 1-dimensional. If $v_1, \ldots, v_n$ is a basis of V then $\bigwedge^k V = \operatorname{span}(v_1 \wedge \cdots \wedge v_n + Y)$

Given $T \in \operatorname{Hom}(V_1, W_1)$, $S \in \operatorname{Hom}(V_2, W_2)$ define $T \otimes S \in \operatorname{Hom}(V_1 \otimes W_1, V_2 \otimes W_2)$ by $v_1 \otimes v_2 \mapsto Tv_1 \otimes Sv_2$. $\bigotimes^k T$ induces $\bigwedge^k T \in \operatorname{End}(\bigwedge^k T)$ as $v_1 \wedge \cdots \wedge v_n \mapsto Tv_1 \wedge \cdots \wedge Tv_n$.

Proposition 6.8.14. For $T \in \operatorname{End}(V)$, $\det(T)$ is the scalar such that $\bigwedge^n T = \det(T) \cdot I$ where $n = \dim V$.

Proof. Since $\bigwedge^k V$ is 1-dimensional and for basis $v_1, \ldots, v_n$, $\bigwedge^k V = \operatorname{span}(v_1 \wedge \cdots \wedge v_n)$, it suffices to calculate $\bigwedge^k T(v_1 \wedge \cdots \wedge v_n)$. Suppose $Tv_i = \sum_{j=1}^n a_{i,j} v_j$, then

$$\bigwedge^n T(v_1 \wedge \cdots \wedge v_n) = \sum_{j=1}^n a_{1,j} v_j \wedge \cdots \wedge \sum_{j=1}^n a_{n,j} v_j$$

Expanding this, we get all possible terms of the form $a_{1,k_1} \wedge \cdots \wedge a_{1,k_n}$ where $k_i \in \{1, \ldots, n\}$. Further, since by construction of the wedge product, any term such that $k_j = k_i$ for $i \neq j$ is zero. So the nonzero terms in the sum are precisely $a_{1,\sigma(1)} v_{\sigma(1)} \wedge \cdots \wedge a_{n,\sigma(n)} v_{\sigma(n)}$ for $\sigma \in \operatorname{perm}(n)$. Observing that

$$\begin{aligned} a_{1,\sigma(1)} v_{\sigma(1)} \wedge \cdots \wedge a_{n,\sigma(n)} v_{\sigma(n)} &= a_{1,\sigma(n)} \cdots a_{n,\sigma(n)} (v_{1,\sigma(1)} \wedge \cdots \wedge v_{n,\sigma(n)}) \\ &= \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} (v_1 \wedge \cdots \wedge v_n) \end{aligned}$$

we see that the sum of all such terms is $\det(T)$ so $\bigwedge^n T = (\det T) \cdot I$, as desired.