

Math 115AH: Linear Algebra (Honors)

2023-24

Fall 23

Instructor: Robert E. Greene (+ Joshua Enwright)

Textbooks: Peter Petersen - Linear Algebra

P. R. Halmos - Finite-Dimensional Vector Spaces

(*) S. Axler - Linear Algebra Done Right.

Topics: Vector spaces, linear transformations & matrices, duality, inner product spaces, eigenvalues & eigenvectors, multilinear forms & determinants, matrix polynomials & decompositions

Table of Contents

1) Vector Spaces - 2

i) Fields - 2

ii) Vector Spaces - 3

iii) Dimension & Linear Independence - 6

2) Linear Transformations - 8

i) Matrices of Linear Maps - 11

ii) Duality - 12

iii) Inner Product & Orthogonality - 14

iv) Eigenvalues, Eigenvectors, & Diagonalization - 18

v) Multilinear Forms - 24

3) Advanced Topics - 28

i) Matrix Polynomials - 28

ii) Eigenspaces & Jordan Form - 30

(*) Matrix Exponentials - 36

iii) Complex Inner Products & Hermitian Adjoints - 43

Fields

9/20/23

Lecture 1

Def: Fields

A field is a set F associated with two binary operations - "addition" ($+ : F \times F \rightarrow F$) and "multiplication" ($\cdot : F \times F \rightarrow F$) satisfying certain properties (called "field axioms").

Field Axioms

Let $a, b, c \in F$. Then:

(1) Associativity: $a + (b + c) = (a + b) + c$; $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

(2) Commutativity: $a + b = b + a$; $a \cdot b = b \cdot a$

(3) Identities: There exist distinct elements $0, 1 \in F$ ($0 \neq 1$), such that: $a + 0 = a$; $a \cdot 1 = a$

(4) Inverses: For every $a \in F$ there exist elements $(-a)$, $a^{-1} \in F$ such that: $a + (-a) = 0$; $a \cdot a^{-1} = 1$

(5) Distributivity: $a \cdot (b + c) = a \cdot b + a \cdot c$

Notes on Fields

- A field F with operations $+$, \cdot may also be denoted $(F, +, \cdot)$.

- A field F represents a form of "number system" (where the numbers are the elements of F).

- (*) Not all "number systems" satisfy the field axioms [are fields]; a number system may have noncommutative multiplication, e.g.

- The elements 0 and 1 in a field are called the additive and multiplicative identity, respectively

- The " 0 " and " 1 " referenced in the additive ($(-a)$) and multiplicative (a^{-1}) inverse axioms refer to the additive & multiplicative identities; thus, identities must be defined before inverses

- Example fields: $\mathbb{R}, \mathbb{Q}, \mathbb{Z}, \mathbb{C}$

- (*) Not all fields are "regular": \mathbb{Z}_p (" \mathbb{Z} mod p ") is a field in which there only exist p elements ($\mathbb{Z}_p = \{0, 1, \dots, p-1\}$) and $(p-1) + 1 = 0$

Vector Spaces

9/29/23

Lecture
(cont.)

Def: Vector Spaces

A vector space is a set V over a field F (where the elements of V are called "vectors") associated with two binary operations - [vector] addition ($+ : V \times V \rightarrow V$) and scalar multiplication ($\cdot : F \times V \rightarrow V$) - satisfying certain axioms.

Vector Space Axioms

Let $u, v, w \in V; a, b \in F$. Then:

(1) Closure: for every $u, v \in V$ and $a \in F$, the sum $u+v$ and product av are themselves elements of V (i.e. $u+v, av \in V$).

(2) Associativity: $u+(v+w)=(u+v)+w$; $a(bv)=(ab)v$

(3) Commutativity: $u+v=v+u$

(4) Additive identity: There exists an element $0 \in V$ such that $v+0=v \forall v \in V$

(5) Multiplicative identity: $1v=v$ (where 1 denotes the multiplicative identity in F).

(6) Additive inverse: For every $v \in V$ there exists an element $(-v) \in V$ such that $v+(-v)=0$

(7) Distributivity: (i) Over vector addition: $a(u+v)=au+av$

(ii) Over scalar addition: $(a+b)v=av+bv$

Notes on Vector Spaces

- The closure requirement applies to fields as well, but more often comes up in the context of subsets of vector spaces
- Vector spaces are "over" fields mainly to provide scalars for scalar multiplication; as such, any vector space can typically be generalized across different fields
- Dot and cross product (used in \mathbb{R}^n) are not inherent properties of all vector spaces (e.g. ill-defined in \mathbb{C})
- Any field is a vector space over itself (and its subfields)

Vector Spaces (cont.)

10/2/23

Lecture 2

Linear Algebra: initially the study of finite unknowns in finite dimensions; later branched out to infinite dimensions

Notes on Vector Spaces (cont.)

- Given a set S and a field F , the function space $F(S, F)$ [the set of functions $f: S \rightarrow F$] has a vector space structure with addition $((f_1 + f_2)(v) = f_1(v) + f_2(v))$ and scalar multiplication $((\alpha f)(v) = \alpha f(v))$
- A coordinate (x_1, x_2, \dots, x_n) in \mathbb{R}^n can be seen as a function $f: \{1, 2, \dots, n\} \rightarrow \mathbb{R}$ (mapping an "index" in the coordinate to the coordinate's value [$v \in \mathbb{R}$] at that index). Thus, the set of all coordinates $(\{1, \dots, n\}, \mathbb{R})$ has a vector space structure over \mathbb{R} .
- Similarly, the space of all sequences on field F $(\{1, 2, \dots\}, F)$ (mapping $f: \mathbb{N} \rightarrow a_n$) is a vector space over F (where addition: $(s_n) + (t_n) = (s_n + t_n)$).
(*) ℓ^2 ("little ell two"), the set of infinite sequences $\{(x_1, x_2, \dots) : \sum_{n=1}^{\infty} x_n^2 < \infty\}$ [square summable sequences], can be shown to be closed & a vector space
- A field is always a vector space over any of its subfields, but subfields do not necessarily form vector spaces over the original field (ex: \mathbb{Q} is not a V.S. over \mathbb{R} ($q \in \mathbb{Q} \rightarrow \pi q \notin \mathbb{Q}$, e.g.))
- "Message": geometry (i.e. structure) still exists even in infinite dimensions

(*) Field Extension

Given a field F , we can extend the field to include additional elements not originally in the field (and also extending addition & multiplication accordingly).

$$\text{Ex: } \mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} : a, b \in \mathbb{Q}\} \quad ["\mathbb{Q} \text{ extended by } \sqrt{2}]$$

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) := \{a + b\sqrt{3} : a, b \in \mathbb{Q}(\sqrt{2})\}$$

$$\mathbb{R}(i) := \{a + bi : a, b \in \mathbb{R}\} = \mathbb{C} \quad ((*) \text{ Then we call } \mathbb{R} \text{ a } \underline{\text{subfield}} \text{ of } \mathbb{C}.)$$

Notes on Fields

10/3/23

Disc. 2

(*) On Fields

Notation: Although we generally accept the meaning of " $a+b+c+\dots$ " and " $a \cdot b \cdot c \cdot \dots$ " to be the repeated addition/multiplication of $a, b, c, \dots \in F$, this is not immediately true given the definition of a field (since addition and multiplication are binary operations - take only two operands). More technically correct expressions would be $+((a, +(b, +(c, \dots))))$ and $\cdot((a, \cdot(b, \cdot(c, \dots))))$ for binary operations $+(a, b)$ and $\cdot(a, b)$.

(*) Proving Field Properties

Many generally-assumed properties of fields are not immediately obvious from the axioms alone.

Ex: Prove $a \cdot 0 = 0$: $a \cdot 0 = a \cdot (0+0) = 0 \cdot a + 0 \cdot a$. Choose b such that $(a+0)+b=0$.

Then $0 = (a+0)+b = (a \cdot 0 + a \cdot 0) + b = a \cdot 0 + 0 = a \cdot 0$. \square

Ex: The uniqueness of 0 and 1 (i.e. that \nexists any elements $0' \neq 0$ or $1' \neq 1$ that possess the properties of additive/multiplicative identity) also requires the other axioms to prove.

(*) Def: Rings

A ring is a set R associated with two binary operations - addition ($+: R \times R \rightarrow R$) and multiplication ($\cdot: R \times R \rightarrow R$) - satisfying certain axioms (called "ring axioms").

Ring Axioms

The list of ring axioms is equivalent to the list of field axioms, with two exceptions: the ring axioms notably do not require commutative multiplication or the existence of the multiplicative identity [1].

(*) A ring can thus be thought of as an "almost-field", or as a generalization of fields to number systems with potentially noncommutative multiplication or lacking a multiplicative identity.

Dimension and Linear Independence

10/4/23

Lecture 3 Dimension of a Vector Space

Def: A vector space V is called finitely generated if \exists a finite set of vectors $v_1, \dots, v_k \in V$ such that every vector $v \in V$ can be expressed in the form $\sum_{j=1}^k \alpha_j v_j$ ($\alpha_1, \dots, \alpha_k \in F$).

(*) Def: Such a term $v = \sum_{j=1}^k \alpha_j v_j$ is called a linear combination of v_1, \dots, v_k .

Def: Dimension

If a vector space V is finitely generated, then we say V is finite dimensional and define the dimension of V , $\dim(V)$ to be the size of the smallest set of vectors that generates V .

Def: Linear Independence

Given a set of vectors $v_1, \dots, v_k \in V$, the following statements are equivalent:

- (1) The vectors v_1, \dots, v_k are linearly independent.
- (2) Every [finite] linear combination $\sum_{j=1}^k \alpha_j v_j$ is equal to 0 only when $\alpha_j = 0 \forall j \in \{1, \dots, k\}$.
- (3) Every nontrivial linear combination $\sum_{j=1}^k \alpha_j v_j$ (i.e. where $\exists \alpha_j \neq 0$) is nonzero.
- (4) No vector v_i in the set is a linear combination of the other vectors $v_{j \neq i}$.

Notes on Linear Independence

- If a set of vectors v_1, \dots, v_k is not [linearly] independent, then it is called dependent.
- If a set of vectors v_1, \dots, v_k is linearly independent, that implies $0 \notin \{v_1, \dots, v_k\}$.
- Given finite-dimensional vector space V , if a set of vectors w_1, \dots, w_k generates V and \nexists any set of vectors v_1, \dots, v_n ($n < k$) finitely generating V , then $\dim(V) = k$ and w_1, \dots, w_k are linearly independent

Dependence & Dimension

10/4/23

Thm: "The Big Theorem"

Let (V, F) be a vector space generated by a finite set of m vectors. Then any set $S \subseteq V$ with $n > m$ elements is linearly dependent.

Lecture 3

(cont.)

+ Lecture 4

(*) Proof 1 (Halmos)

Let w_1, \dots, w_m be a set of generating vectors for V , and let v_1, \dots, v_n be a set of linearly independent vectors. Assume, for the sake of contradiction, that $n > m$.

Now, "attach" v_i to $w_1, \dots, w_m \rightarrow v_1, w_1, \dots, w_m$. Since w_1, \dots, w_m is a set of generating vectors, v_1 is a linear combination of w_1, \dots, w_m ; thus v_1, w_1, \dots, w_m is dependent and generating. Since v_1 is a linear comb. of v_1, \dots, v_n , then $\exists v_i$ s.t. v_i is a linear combination of v_1, \dots, v_n . Then we can remove v_i from the set while still generating V : we have "replaced" w_i with v_i . [Remove v_i from v_1, \dots, v_n].

Continue replacing vectors w_i with vectors v_j until all vectors w_i have been replaced. Then we have a set of vectors v_1, \dots, v_m that generates V . But since $n > m$, \exists at least one $v_i \notin \{v_1, \dots, v_m\}$; since v_1, \dots, v_m generates V , then this v_i must be a linear combination of v_1, \dots, v_m . But we stated that v_1, \dots, v_n was linearly independent; thus we have a contradiction. \square

(*) Proof 2

Reminder: If # unknowns > # equations in a homogeneous system of linear equations, \exists a nontrivial solution.

Let v_1, \dots, v_n be a set of n vectors in a vector space of dimension m ($n > m$). To prove v_1, \dots, v_n are linearly dependent, we want to find $x_1, v_1 + x_2 v_2 + \dots + x_n v_n = 0$ (x_i not all 0). We know \exists a spanning set of vectors w_1, \dots, w_m s.t. $v_j = a_{j1}w_1 + \dots + a_{jm}w_m \forall v_j$. Substituting $x_1(a_{11}w_1 + \dots + a_{1m}w_m) + \dots + (x_n a_{n1}w_1 + \dots + a_{nm}w_m) = 0$.

We can rearrange to get: $(a_{11}x_1 + a_{21}x_2 + \dots + a_{n1}x_n)v_1 + \dots + (a_{1m}x_1 + \dots + a_{nm}x_n)v_m = 0$. We know $a_{ij}w_1 + \dots + a_{im}w_m = 0$ only if $a_{ij} = 0 \forall j$; then we want to find x_1, \dots, x_n s.t. $a_{ij}x_1 + \dots + a_{nj}x_n = 0 \forall j \in \{1, \dots, m\}$. This is a homogeneous linear system with m equations and $n > m$ unknowns; thus, we know \exists a nontrivial solution such that $x_1, \dots + x_n v_n = 0$, therefore v_1, \dots, v_n are linearly dependent.

Bases, Subspaces, and Linear Transformations

10/9/23

Lecture 5

Bases and Subspaces

Def: Given finitely generated vector space V , a basis of V is a linearly independent set of vectors that generates V .

(*) Finding a Basis

Take an independent set S (e.g. a set of just 1 vector). Either S generates V , or $\exists v \in V$ not generated by S . Then we know $S \cup \{v\}$ is still independent \rightarrow set $S = S \cup \{v\}$ & repeat.

Def: Given vector space V , a subspace of V is a subset $V' \subset V$ such that V' is itself a vector space (with the same operations as V).

(*) If $\dim(V) = n$ and $\dim(V') = k < n$, then for any basis v_1, \dots, v_k of V' , we can extend the basis by $(n-k)$ additional vectors v_{k+1}, \dots, v_n to obtain a basis of V .

(*) Note: representations of vectors as combinations of bases are unique.

Def: Linear Transformations

A linear transformation is a linear function $f: V \rightarrow W$ such that $f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2)$.

We call the set of linear transformations $V \rightarrow W$ $\text{Hom}(V, W)$, and for each $T \in \text{Hom}(V, W)$ we define two vector spaces "attached" to T :

(*) Def: Kernel and Image

(1) Kernel/null space of $T := \ker(T) = \{v \in V : T(v) = \vec{0} \in W\}$

(2) Image/range $T(V)$ of $T := \text{im}(T) = \{T(\vec{v}) : \vec{v} \in V\}$

THM: If V is finite dimensional (W need not be), then both $\ker(T)$ and $\text{im}(T)$ are finite dimensional and $\dim(\text{im}(T)) + \dim(\ker(T)) = \dim(V)$.

(*) Vector Spaces & Transformations

10/10/23

Discussion 4

Discussion 4

Def: Given vectors v_1, \dots, v_n their span $[\text{span}(v_1, \dots, v_n)]$ is the set of all linear combinations of v_1, \dots, v_n , or the intersection of all subspaces $V' \subset V$ containing all of v_1, \dots, v_n

+ Discussion 5

(*) Equivalently, $\text{span}(v_1, \dots, v_n)$ is the smallest subspace containing all of v_1, \dots, v_n

(*) A vector space V is finite-dimensional if it is spanned by a finite list of vectors

Propositions: (i) Given finite-dimensional vector space V with $\dim V = n$, then any set S that spans V and has $|S| = n$ is a basis for V .

(ii) Similarly, any linearly independent T where $|T| = n$ is a basis for V .

Discussion 5

Prop: Let V be a f.d.v.s. with basis $B = \{x_1, \dots, x_n\}$. Given another vector space W , the assignment $\text{Lin}(V, W) \rightarrow \text{Fun}(B, W)$ [equivalently, $\mathcal{F}: \text{Hom}(V, W) \xrightarrow{\sim} T|_B$] is a bijection.

(*) The notion of "linearity" between V, W implicitly requires V, W over the same field

(*) Axiom of Choice: Every vector space has a basis.

Prop: Assume V, W are vector spaces. Then a linear map $T: V \rightarrow W$ is injective iff \exists a "left inverse" linear map $S: W \rightarrow V$ such that $S \circ T = \text{id}$ in V .

(*) Analogously: $T: V \rightarrow W$ is surjective iff \exists a "right inverse" linear map $S: W \rightarrow V$ such that $T \circ S = \text{id}$ in W .

Notes: Homework 1 & Lecture

• Lecture: The "big theorem" [about lin. indep.] is a fairly special [i.e. not universal] property of vector spaces in finite group theory.

• HW1: If F is a finite field, then $\text{Lin}(F) \quad 1+1+\dots+1=0 \in F$ for some number of 1s. The smallest such positive integer is called the characteristic of F and is prime.

• HW1: $\text{Hom}(V, W) = \text{set of linear functions } V \rightarrow W; C([0, 1]) = \text{set of continuous functions } [0, 1] \rightarrow \mathbb{R}, [F(v, w) = \text{func. } V \rightarrow W]$

Rank-Nullity

10/11/23

Lecture 6

Rank-Nullity

Reminder [Rank-Nullity theorem]: If vector space V is finite-dimensional, then $T(V)$ is finite-dimensional and $\dim(T(V)) = \dim V - \dim N(V)$.

(*) Proof: Let V be a vector space, $T: V \rightarrow W$ a linear transformation. Let $\dim V = n$.

Choose basis $v_1, \dots, v_k \in V$ ($k \leq n$) as a basis for $\ker(T) \subset V$. Then we need $(n-k)$ additional independent vectors v_{k+1}, \dots, v_n to form a basis for V . We claim that $T(v_{k+1}), \dots, T(v_n)$ is a basis for $T(V)$.

(**) Proof: $T(V) = \left\{ \sum_{j=1}^n \alpha_j v_j \mid v_j \in V \right\}$ is the set of linear combinations of vectors $T(v_j) \in W$.

Then $T(v_1), \dots, T(v_n)$ generate $T(V)$. Except $v_1, \dots, v_k \in \ker(T)$ and therefore $T(v_1) = \dots = T(v_k) = 0$; thus, $T(v_{k+1}), \dots, T(v_n)$ generate $T(V)$.

We want to show that $T(v_{k+1}), \dots, T(v_n)$ are independent. Assume they are not independent; then $\exists \sum_{j=k+1}^n \beta_j T(v_j) = 0 \in W$ for some $\beta_{k+1}, \dots, \beta_n$ not all 0.

Then $\sum_{j=k+1}^n \beta_j v_j \in \ker(T)$; therefore \exists a linear combination [not trivial] of v_1, \dots, v_k equal to $\sum_{j=k+1}^n \beta_j v_j$; but this is impossible, since v_1, \dots, v_n are a basis and thus independent. Therefore $T(v_{k+1}), \dots, T(v_n)$ are independent and thus a basis for $T(V)$.

Then $\dim(V) = n = k + (n-k) = |\{v_1, \dots, v_k\}| + |\{v_{k+1}, \dots, v_n\}| = \dim(\ker(T)) + \dim(T(V))$. \blacksquare

(**) Corollary: $\ker(T) = \{0\}$ iff $\text{im}(T) = V$ [$\dim\{0\} = 0$]

Notes on Transformations $(T: V \rightarrow W)$

- For a linear transformation T , the associated vector spaces $\ker(T)$, $\text{im}(T)$ are actually subspaces (where $\ker(T) \subset V$, $\text{im}(T) \subset W$) [As proof of subspace, we need only check closure & additive identity]

(*) Does not necessarily hold for nonlinear transformations (e.g. range($f: x \mapsto x^2$) = $\{y: y \geq 0\}$ is not closed under addition)

- Things (elements, operations, etc.) may not always be the same across vector spaces

(*) " $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$ " \rightarrow " $\alpha v_1 + \beta v_2$ " addition in V ; " $\alpha T(v_1) + \beta T(v_2)$ " addition in W

Matrix Representations of Linear Transformations

10/13/23

Lecture 7

Bases & Linear Transformations

Notation: Standard bases in \mathbb{F}^n : $e_1, \dots, e_n : (1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 0, 1)$

$$\Rightarrow \vec{v} \in \mathbb{F}^n = \sum_{j=1}^n \alpha_j e_j \Rightarrow T(\vec{v}) = T\left(\sum_{j=1}^n \alpha_j e_j\right) = \sum_{j=1}^n \alpha_j T(e_j) [T: \mathbb{F}^n \rightarrow V]$$

(*) Implication: knowing $T(e_j)$ for all e_1, \dots, e_n = knowing T for all $\vec{v} \in \mathbb{F}^n$

Matrix Representations of Linear Transformations

Let $T: \mathbb{F}^n \rightarrow V$ be a linear transformation, and let e_1, \dots, e_n be the standard basis of \mathbb{F}^n . Let $\dim(V) = m$. Then the matrix of T is an $m \times n$ matrix, where the j^{th} column is the vector $T(e_j) \in V$ for all $1 \leq j \leq n$, i.e.:

$$[T] = [T(e_1) \ T(e_2) \ \dots \ T(e_n)] \Rightarrow m \text{ rows, } n \text{ columns}$$

Notes on Matrices of Linear Transformations

- The image $T(\mathbb{F}^n)$ of T is generated by the columns $T(e_j)$ of the matrix of T
 $\Rightarrow \dim(T(\mathbb{F}^n)) = \dim(\text{span}(T(e_1), \dots, T(e_n)))$

(*) This is also equal to the dimension of the space generated by the rows

• Alternate notation:

$$L(x) = [T(e_1) \ \dots \ T(e_n)] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} \quad (\text{for some scalars } \alpha_1, \dots, \alpha_m)$$

(*) The choice to represent $T(e_j)$ as columns of $[T]$ is somewhat arbitrary - we can express $T(e_j)$ as rows instead if we so chose, for instance

(*) Linear algebra began as the study of matrices [concreted]; would only later transition into more abstract topics (e.g. vector spaces)

• Matrices & Linear Transformations: $T(\vec{x}) = [T] \vec{x}$

Dual Spaces

10/13/23

Lecture 7

Def: Dual Spaces

(cont.)

Given a vector space V over F , we define the dual [vector] space of (V, F) to be the vector space $V^* = \text{Hom}(V, F)$, i.e. the space of linear transformations $T: V \rightarrow F$.

Def: Dual Bases

Let V be a finite-dimensional vector space over F , and let v_1, \dots, v_n be a basis for V . Define linear functions $\Theta_1, \dots, \Theta_n \in V^*$, where $\Theta_i(v_j) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$. Then $\Theta_1, \dots, \Theta_n$ are a basis for V^* , called the dual basis of v_1, \dots, v_n . $[\Theta_i(\sum_{k=1}^n \alpha_k v_k) = \alpha_i]$

(*) Notation [Kronecker delta]: $\Theta_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$

(*) Proof: Dual Bases

We claim $\Theta_1, \dots, \Theta_n$ dual to v_1, \dots, v_n are a basis for V^* .

(i) Independence: Suppose $\sum_{j=1}^n \gamma_j \Theta_j = 0$ [zero transformation]. Then for all $v_k: (\sum_{j=1}^n \gamma_j \Theta_j)(v_k) = 0$
 $\Rightarrow \gamma_k = 0$. Then $\gamma_j = 0$ for all j .

(ii) Generativity: Let $w \in V^*$. We want to find $\gamma_1, \dots, \gamma_n$ such that $w = \sum_{j=1}^n \gamma_j \Theta_j$. Set
 $\gamma_k = w(v_k)$ for each basis vector v_k ; then $\sum_{j=1}^n \gamma_j \Theta_j(v) = w(v) \quad \forall v \in V$.

(*) Dual Spaces in Infinite Dimensions

Although dual spaces and functions analogous to dual basis functions can be defined in infinite-dimensional spaces, it is not possible to find [finite] dual bases in infinite dimensions.

Ex: Let V be the space of infinite sequences that are eventually all 0 (i.e. sequences of the form $(v_1, v_2, v_3, \dots, v_n, 0, 0, 0, \dots)$).

Then there is no finite set $\Theta_1, \dots, \Theta_n$ that could express $\Omega: \Sigma(v_1, v_2, \dots)$ in terms of Θ . [If there were such a set, we would simply find a sequence with (n+1) nonzero elements as a counterexample].

Annihilator & Adjoint

10/16/23

Lecture 8

Def: Annihilator

Let $S \subset V$ be a subspace. Then we define the annihilator of S , $\text{ann}(S) \subset V^*$ to be the subspace

$$\text{ann}(S) := \{\omega \in V^* : \omega(s) = 0 \ \forall s \in S\}$$

Theorem: Dimension & Annihilator

If $W \subset V$ is a finite-dimensional subspace, then $\text{ann}(W) \subset V^*$ is also finite-dimensional and $\dim(\text{ann}(W)) = \dim(V) - \dim(W)$.

Proof: Choose a basis v_1, \dots, v_k ($k \leq n$) for W [$\dim W = k$]; then we want v_{k+1}, \dots, v_n to obtain a basis for V [$\dim V = n$]. Let $\theta_1, \dots, \theta_n$ be the dual basis of v_1, \dots, v_n . (Claim: $\text{ann}(W)$ is the set of all linear combinations of $\theta_{k+1}, \dots, \theta_n$, thus $\theta_{k+1}, \dots, \theta_n$ is a basis for $\text{ann}(V)$.)

Proof: Suppose $\omega = \sum_{j=1}^n \alpha_j \theta_j \in \text{ann}(W)$. Then $\omega(w) = 0 \ \forall w \in W$, therefore $\omega(v_i) = 0 \ \forall 1 \leq i \leq k$. Then $\omega \in \text{ann}(W)$ iff $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, i.e. when $\omega = \sum_{j=k+1}^n \alpha_j \theta_j$.

Then $\dim(\text{ann}(V)) = (n-k) = \dim V - \dim W$.

Def: Adjoint

Let $T: V \rightarrow W$ for vector spaces V, W over F . Then we define the adjoint of T to be the linear transformation $T^*: W^* \rightarrow V^*$ defined by $\omega \mapsto \omega \circ T$.

Notes on Adjoints

Observation: $\ker(T^*) = \text{ann}(T(V))$

Proof: Let $\omega \in W^*$. $\omega \in \ker(T^*) \Rightarrow (T^*\omega)(v) = 0 \ \forall v \in V \Rightarrow \omega(T(v)) = 0 \ \forall v \in V \Rightarrow \omega \in \text{ann}(T(V))$.

Consequence: $\dim(N(T^*)) = \dim(\text{ann}(T(V))) = \dim(W) - \dim(T(V))$.

(*) Notation: Adjoint may also be referred to as algebraic adjoint, transpose, or dual [map] of T

Inner Product

10/18/23

Lecture 9

Def: Inner Product

An inner product on a vector space V over \mathbb{F} is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ [mapping ordered pairs $(v_1, v_2) \in V^2$ to \mathbb{F}] that satisfies:

(i) $\langle x, y \rangle = \langle y, x \rangle$ [Symmetry]

(ii) $\langle x, x \rangle \geq 0 \quad \forall x \in V; \langle x, x \rangle = 0 \text{ only if } x = \vec{0}$ [Positive definite]

(iii) $\langle \cdot, \cdot \rangle$ is linear in both variables (i.e. $x \mapsto \langle x, y \rangle$ is linear $\forall y \in V$)

Inner Products in Geometry

Def $\langle \cdot, \cdot \rangle : ((a_1, \dots, a_n), (b_1, \dots, b_n)) \mapsto a_1 b_1 + \dots + a_n b_n$ be an inner product on \mathbb{R}^n .

Def: Let (V, \mathbb{R}) be a vector space with inner product. Then for every $v_1, v_2 \in V$, we define the

magnitude of v : $\|v\| = \sqrt{\langle v, v \rangle}$ and the distance $(v_1, v_2) = \|v_1 - v_2\|$. [Def: Distance/Norm]

Def: If $v_1, v_2 \neq \vec{0}$, we define $\cos(\theta) = \frac{\langle v_1, v_2 \rangle}{\|v_1\| \|v_2\|}$ for some $0 \leq \theta \leq \pi$. [Def: Angles]

Prop (Cauchy-Schwarz): $|\langle v, w \rangle| \leq \|v\| \cdot \|w\|$

Def: Let $v, w \in V$. Then we define the projection of v onto w : $\text{proj}_w(v) = \frac{\langle v, w \rangle}{\|w\|^2} w$

(Claim: $\|v\| \geq \|\text{proj}_w(v)\|$)

(*) Proof: $\langle v - \text{proj}_w(v), v - \text{proj}_w(v) \rangle = \langle v, v \rangle - 2 \frac{\langle v, w \rangle}{\|w\|^2} \langle v, w \rangle + \frac{\langle v, w \rangle^2}{\|w\|^2} \langle w, w \rangle$
 $\rightarrow \langle v, v \rangle - 2 \frac{\langle v, w \rangle^2}{\|w\|^2} + \frac{\langle v, w \rangle^2}{\|w\|^2} = \langle v, v \rangle - \langle v, \frac{w}{\|w\|^2} w \rangle \geq 0 \Rightarrow \|v - \text{proj}_w(v)\| \geq 0$. \square

(*) Proof: Assume $v, w \neq \vec{0}$ (otherwise, trivial). Then: $\|v\| \geq \|\text{proj}_w(v)\| = \left| \frac{\langle v, w \rangle}{\|w\|^2} \right| \|w\| = \frac{|\langle v, w \rangle|}{\|w\|}$. \square

Corollary (Triangle inequality): $\|v + w\| \leq \|v\| + \|w\|$

(*) Proof: $\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + 2 \langle v, w \rangle + \langle w, w \rangle \leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$

(*) Note: Cauchy-Schwarz does not stipulate finite dimensions.

Ex: Inner product on ℓ^2 : $x, y \in \ell^2 \rightarrow \langle x, y \rangle = \sum |x_i \cdot y_i| \leq \sum (|x_i|^2 + |y_i|^2) < \infty$

(*) Notation: $\langle x, y \rangle$ may also be written $\langle x | y \rangle$.

Orthonormalization

10/18/23

Lecture 9

+ Lecture 10

Orthonormalization

Let (V, \mathbb{R}) be a vector space with inner product $\langle \cdot, \cdot \rangle$. Let v_1, \dots, v_n be a finite basis for V .

Problem: We want v_1, \dots, v_n to be similar to $e_1, \dots, e_n \in \mathbb{R}^n$; i.e. we want $\|v_i\| = 1 \forall 1 \leq i \leq n$, $\langle v_i, v_j \rangle = 0 \forall i \neq j$. [We want v_i, v_j orthogonal/perpendicular].

Theorem (Gram-Schmidt):

Let v_1, \dots, v_n be linearly independent. Then $\exists w_1, \dots, w_n \in V$ such that:

$$(i) \|w_i\| = 1 \forall 1 \leq i \leq n.$$

$$(ii) \langle w_i, v_j \rangle = 0 \forall i \neq j.$$

$$(iii) \text{span}(w_1, \dots, w_n) = \text{span}(v_1, \dots, v_n).$$

We can construct w_1, \dots, w_n inductively:

Base case: Let $w_1 = \frac{v_1}{\|v_1\|}$. Then $\|w_1\| = 1$.

Ind. Step: We want w_{k+1} such that w_{k+1} is orthogonal to v_1, \dots, v_k . We set

$$w_{k+1} = v_{k+1} - \sum_{j=1}^k \text{proj}_{w_j}(v_{k+1}) = v_{k+1} - \sum_{j=1}^k \frac{\langle w_j, v_{k+1} \rangle}{\|w_j\|^2} w_j, \text{ then divide by its length.}$$

Then w_1, \dots, w_n are an orthonormal basis.

(*) Proof: Let $i \neq j$. Then: $\langle w_i, w_j \rangle = \langle w_i, v_j - \sum_{k=1}^{i-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_k \rangle = \langle w_i, v_j \rangle - \langle w_i, \sum_{k=1}^{i-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_k \rangle$
 $\rightarrow = \langle w_i, v_j \rangle - \langle w_i, v_j \rangle = 0. \blacksquare$

(*) Consequences

(i) If (V, \mathbb{R}) is a finite-dimensional vector space, then (V, \mathbb{R}) has an orthonormal basis.

(ii) If $W \subset V$ (V as above), $\dim W = k$ and $\dim V = n$, then \exists an orthonormal basis w_1, \dots, w_n for V such that w_1, \dots, w_k is a basis for W .

(*) Changing Orthonormal Bases analogous to rotating axes in \mathbb{R}^n :

$$\text{Ex } (\mathbb{R}^2): \begin{cases} \downarrow & \Rightarrow & \begin{matrix} i \\ j \end{matrix} & \{ (1,0), (0,1) \} \Rightarrow \{ (\sqrt{2}/2, \sqrt{2}/2), (-\sqrt{2}/2, \sqrt{2}/2) \} \end{cases}$$

Orthogonal Transformations

10/20/23

Lecture 10

(cont.)

Orthogonal Transformations

Def: Let (V, \mathbb{R}) be a vector space with inner product $\langle \cdot, \cdot \rangle$. Then a linear transformation $T: V \rightarrow V$ is called orthogonal if $\langle T(v_1), T(v_2) \rangle = \langle v_1, v_2 \rangle \forall v_1, v_2 \in V$ [i.e. T preserves inner products].

Orthogonal Transformations

Lemma: " $T: V \rightarrow V$ is orthogonal" \Leftrightarrow " $\|T(v)\| = \|v\| \forall v \in V$ ". [i.e. T preserves length].

(*) Proof:

$$A \rightarrow B: \|T(v)\| = \langle T(v), T(v) \rangle^{1/2} = \langle v, v \rangle^{1/2} = \|v\|. \blacksquare$$

$$B \rightarrow A: \text{Know: } \|v_1 + v_2\| = \|T(v_1 + v_2)\| \Rightarrow \|v_1 + v_2\| = \|T(v_1) + T(v_2)\|$$

$$\rightarrow \|v_1 + v_2\|^2 = \langle v_1 + v_2, v_1 + v_2 \rangle = \|T(v_1) + T(v_2)\|^2 = \langle T(v_1) + T(v_2), T(v_1) + T(v_2) \rangle$$

$$\langle T(v_1) + T(v_2), T(v_1) + T(v_2) \rangle = \langle T(v_1), T(v_1) \rangle + 2\langle T(v_1), T(v_2) \rangle + \langle T(v_2), T(v_2) \rangle$$

$$= \|T(v_1)\|^2 + 2\langle T(v_1), T(v_2) \rangle + \|T(v_2)\|^2$$

$$= \|v_1\|^2 + 2\langle T(v_1), T(v_2) \rangle + \|v_2\|^2$$

$$= \langle v_1 + v_2, v_1 + v_2 \rangle = \|v_1\|^2 + 2\langle v_1, v_2 \rangle + \|v_2\|^2$$

$$\rightarrow \langle v_1, v_2 \rangle = \langle T(v_1), T(v_2) \rangle. \blacksquare$$

(*) Def: Groups

A group is a set with an operation satisfying: associativity, presence of an identity element, and the existence of an inverse element for every element in the set.

(*) The Orthogonal Group

Let (V, \mathbb{R}) be a finite-dimensional vector space with inner product $\langle \cdot, \cdot \rangle$. Then the set of orthogonal transformations $T: V \rightarrow V$ has a group structure with the operation of function composition; this is called the orthogonal group.

(*) Bilinear Forms

10/19/23

Disc 7

(*) Exercise: Infinite-Dimensional Spaces

We know \mathbb{R} is an infinite-dimensional vector space over \mathbb{Q} . Let p_n be the [finite] sequence of prime integers: $p_1=2, p_2=3, p_3=5, \dots$. Recall (Fundamental Thm): any $m \in \mathbb{N} > 1$ has a unique prime factorization.

Claim: $\log(p_1), \log(p_2), \dots$ is linearly independent over \mathbb{Q} .

Proof: Suppose $\exists a_1, \dots, a_n \in \mathbb{Q}$ such that $\sum a_i \log(p_i) = 0$. We want to show $a_1 = \dots = a_n = 0$.

Recall log properties: $\sum a_i \log(p_i) = 0 \Rightarrow 0 = \log\left(\prod_{i=1}^n p_i^{a_i}\right)$.

Drop the log: $1 = \prod_{i=1}^n p_i^{a_i}$. Observe: if a_1, \dots, a_n all non-negative, then we could conclude that a_1, \dots, a_n all 0.

Assume $\exists a_i < 0$. Then let $I_{(0)} = \{i : a_i < 0\}$, $I_{(+)} = \{i : a_i \geq 0\}$. Multiply both sides to

cancel the negative exponents: $\prod_{i \in I_{(0)}} p_i^{|a_i|} = \prod_{i \in I_{(+)}} p_i^{a_i}$, where $p_i \neq p_j$ for any $i \in I_{(+)}, j \in I_{(-)}$.

Let $x = \prod_{i \in I_{(0)}} p_i^{|a_i|} = \prod_{i \in I_{(0)}} p_i^{a_i}$. We know $p_i > 0 \forall i \in \mathbb{N}$; then $x > 0$. Assume WLOG that a_1, \dots, a_n are integers. Then $x \in \mathbb{N}$. If $x > 1$, then x is an integer with two different prime factorizations; this is impossible, therefore $x = 1$. Therefore $a_1 = \dots = a_n = 0$. \blacksquare

(*) Prop: Dual Bases

Let (V, F) be a vector space with basis v_1, \dots, v_n and dual basis f_1, \dots, f_n . Then, for any $v \in V$, $v = \sum_{j=1}^n f_j(v) \cdot v_j$. Additionally, for any $g \in V^*$, $g = \sum_{j=1}^n g(v_j) \cdot f_j$.

Def: Bilinear Forms

Let V be a vector space over F . Then a bilinear form is a function $B: V \times V \rightarrow F$ that is linear in each component, separately. (Ex: Dot product in \mathbb{R}^2)

Remark: Bilinear Forms

Bilinear forms on V are analogous to linear maps $V \rightarrow V^*$, i.e.: let $\ell: \text{Bil}(V) \rightarrow \text{Hom}(V, V^*)$, such that for $B \in \text{Bil}(V)$, $\ell(B) = \ell_B: V \rightarrow V^*$, where $\ell_B(v)(u) = B(v, u)$ \square

Eigenvectors & Diagonalization

10/13/23

Lecture 11

Diagonalization

Let V be a vector space with basis v_1, \dots, v_n . Then for any linear transformation $T: V \rightarrow W$, where $\dim W = n$ and $\dim(\text{im}(T)) = k$, we can find a basis w_1, \dots, w_n for W such that $T(v_j) = w_j$ if $j \leq k$ and $T(v_j) = 0$ if $j > k$. Then the matrix of T becomes:

$$\begin{pmatrix} I^k & 0 \\ 0 & 0 \end{pmatrix}$$

A more interesting case: let $T: V \rightarrow V$, where we can only pick a single basis for both sides.

Given $T: V \rightarrow V$, can we find a basis for V such that the matrix of T is diagonal?

Similarity

Def: We say that two linear transformations $T_1, T_2: V \rightarrow V$ are similar if \exists bases v_1, \dots, v_n and u_1, \dots, u_n of V such that the matrix of T_1 with respect to the bases u_1, \dots, u_n is equal to the matrix of T_2 with respect to v_1, \dots, v_n .

Diagonalization and Similarity

Prop: Given a matrix A and associated transformation $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$, T_A is similar to some T_{diagonal} if, for some e_1, \dots, e_n , $T_A(e_i) = \alpha_i e_i$ (for some $\alpha_i \in \mathbb{R}$) $\forall i$.

Def: Eigenvectors.

Let $T: V \rightarrow V$. A vector $\vec{v} \neq 0 \in V$ is an eigenvector of T if $T(\vec{v}) = \lambda \vec{v}$ for some scalar λ . Then \vec{v} is an eigenvector with the associated eigenvalue λ ; notably, $\vec{v} \in \ker(T - \lambda I)$.

\Rightarrow Diagonalization

A linear transformation $T: V \rightarrow V$ is diagonalizable only if the eigenvectors of T form a basis for V . Otherwise, if $\#$ eigenvectors $< \dim V$, we say T is non-diagonalizable.

(*) Bilinear Forms (cont.)

10/24/17

Disc 8

Recall: Isomorphisms

Def: A linear map $T: V \rightarrow W$ is called an isomorphism if $\exists L: W \rightarrow V$ such that $TL = I_W$ and $LT = I_V$.

V and W are called isomorphic if there exists an isomorphism $V \rightarrow W$.

Recall: Bilinear Forms

A bilinear form $B: V \times V \rightarrow F$ gives rise to a linear transformation $f_B: V \rightarrow V^*$, where $(\rho(B))(x) = B(x, \cdot): V \rightarrow F$. We note that all transformations $V \rightarrow V^*$ arise in this way uniquely.

Remark: f_B is an isomorphism iff f_B is injective (since $\dim V = \dim V^*$), i.e. iff $\ker(f_B) = \{\vec{0}\}$.

A vector $\vec{x} \in \ker(f_B)$ iff $f_B(\vec{x})$ is the zero function, i.e. $B(\vec{x}, y) = 0 \forall y \in V$.

Degeneracy

Def: A bilinear form $B: V \times V \rightarrow F$ is called nondegenerate if $\forall x \neq \vec{0} \in V, \exists y \in V$ such that $B(x, y) \neq 0$. Equivalently: f_B is bijective.

Nondegenerate Forms

Prop: Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product space over \mathbb{R} . Then the inner product $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ is a nondegenerate bilinear form.

Corollary: $\forall g \in V^*, \exists$ unique $x \in V$ such that $g(y) = \langle x, y \rangle \forall y \in (V, \langle \cdot, \cdot \rangle)$.

Prop: Take V as before, and let $f: V \rightarrow V^*$ be the isomorphism determined by the inner product $\langle \cdot, \cdot \rangle$. Then a basis v_1, \dots, v_n for V is orthonormal iff its dual basis is $(f(v_1), \dots, f(v_n))$.

Recall: If v_1, \dots, v_n is a basis for V with dual basis f_1, \dots, f_n , then $x = \sum_{i=1}^n f_i(x)v_i \quad \forall x \in V$.

\Rightarrow Corollary: If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space with orthonormal basis v_1, \dots, v_n , then

$$x = \sum_{i=1}^n \langle x, v_i \rangle v_i \quad \forall x \in V.$$

Self-Adjoint Transformations

10/25/23

Lecture 12 Symmetric Matrices

The 2×2 symmetric matrix $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ has eigenvalues $\lambda = \frac{a+d}{2} \pm \sqrt{(a-d)^2 + 4b^2}$ [if $b=0 \Rightarrow$ diagonal] \rightarrow has two distinct real solutions λ_1, λ_2 , if $b \neq 0$.

$\rightarrow v_1 = (1, -\frac{a-\lambda_1}{b})$; $v_2 = (1, -\frac{b}{d-\lambda_2})$. Observation: $v_1 \perp v_2 \Rightarrow$ eigenvectors for $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ are orthogonal. Then we can find an orthonormal basis of eigenvectors.

More generally: let $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$. Then $\langle A(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}), (\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) \rangle = \langle (\begin{pmatrix} ax_1 + by_1 \\ bx_1 + dy_1 \end{pmatrix}), (\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) \rangle$
 $\rightarrow ax_1x_2 + b(y_1x_2 + x_1y_2) + dy_1y_2$. This result is the same no matter which side of the equation A is on; i.e. $\langle Av, w \rangle = \langle v, Aw \rangle \forall v, w$.

\rightarrow Let T_A be the linear transformation assoc. with A . Then $\langle T_A v, w \rangle = \langle v, T_A w \rangle$.

Def: Self-Adjoint

A linear transformation $T: V \rightarrow V$ is called self-adjoint if $\langle Tv, w \rangle = \langle v, Tw \rangle \forall v, w \in V$.

Theorem: A linear transformation T is self-adjoint if and only if its matrix representation relative to any orthonormal basis is symmetric. [Relative to one \Rightarrow relative to all]

Eigenvectors of Self-Adjoint Transformations

Suppose $T: V \rightarrow V$ self-adjoint, $\dim V = n$. Let $\vec{v} \in V$ be an eigenvector for T ($v \neq 0$). Then $\vec{v}^\perp = \{w: \langle w, v \rangle = 0, w \in V\}$ has the property that $T_{\vec{v}^\perp} \subset \vec{v}^\perp$.

(+) Pmt: $\langle \vec{w}, v \rangle = 0 \Rightarrow \langle Tv, v \rangle = \langle w, T_v \rangle = \langle w, \lambda v \rangle = \lambda \langle w, v \rangle = 0 \rightarrow T_w \in \vec{v}^\perp \square$

Restricting T to \vec{v}^\perp , then the transformation $T: \vec{v}^\perp \rightarrow \vec{v}^\perp$ is self-adjoint with respect to \vec{v}^\perp . Then we can find an eigenvector for $T_{\vec{v}^\perp}$ orthogonal to \vec{v} ; we can do this n times. Then if we can prove T has at least 1 real eigenvalue, we can prove \exists an orthonormal basis of eigenvectors of T by induction.

(*) Linear Transformation Examples

UCLA Basic Exam

10/26/23

Problem (S17): Let $M \in \mathbb{R}^{n \times n}$, with M^T . Show that $\text{im}(M) = \text{im}(MM^T)$. [im = image]

Disc 9

Proof: We know $\text{im}(MM^T) \subset \text{im}(M)$. Then it suffices to show $\dim(\text{im}(MM^T)) = \dim(\text{im}(M))$; since $\text{rk}(M) = \text{rk}(M^T)$, then we need only show $\dim(\text{im}(MM^T)) = \dim(\text{im}(M^T)) = \text{rank}(M^T)$.

To do this, we want to show $\text{null}(MM^T) = \text{null}(M^T)$. (Claim: $\ker(MM^T) = \ker(M^T)$).

(*) Proof: We know $\ker(MM^T) \supset \ker(M^T)$; we want to show $\ker(MM^T) \subset \ker(M^T)$.

$$\begin{aligned} \text{Let } \vec{x} \in \ker(MM^T). \text{ Then } \vec{0} = \vec{x}^T \cdot \vec{0} = \vec{x}^T (MM^T \vec{x}) = (\vec{x}^T M)(M^T \vec{x}) \\ \rightarrow (M^T \vec{x})^T (M^T \vec{x}) = 0 \Rightarrow \|M^T \vec{x}\| = 0 \Rightarrow \vec{x} \in \ker(M^T). \square \end{aligned}$$

Problem (S18): Let $n \in \mathbb{N}_+$. Show that $e^t, e^{2t}, \dots, e^{nt}$ are linearly independent.

Proof: We want to show that claim holds $\forall n \in \mathbb{N}_+$; proceed by induction.

Base case: $n=1$; then $e^t \neq 0 \Rightarrow \{e^t\}$ is independent.

Ind. step: Assume the claim holds for $e^t, e^{2t}, \dots, e^{nt}$.

Suppose $a_1, \dots, a_{n+1} \in \mathbb{R}$ are s.t. $\sum_{m=1}^{n+1} a_m e^{mt} = 0$. Applying the linear map $((n+1) \text{Id} - \frac{d}{dt})(a_n e^{nt})$, then $0 = ((n+1) \text{Id} - \frac{d}{dt}) \left(\sum_{m=1}^{n+1} a_m e^{mt} \right) = \sum_{m=1}^n a_m (n+1-m) e^{mt} = 0$.

This is a linear combination of e^t, \dots, e^{nt} equating 0; then per the inductive hypothesis, $a_1, \dots, a_n = 0$. Then $a_{n+1} = 0$. \square

(*) Alternatively: Can show that e^t, \dots, e^{nt} are eigenvectors (w/ distinct eigenvalues) for the linear map $\frac{d}{dt}$. [namely, $\frac{d}{dt} e^{nt} = n(e^{nt})$].

The Spectral Theorem

10/27/23

Lecture 13

The Spectral Theorem

Thm: Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product space. If a transformation $T: V \rightarrow V$ is self-adjoint, then \exists an orthonormal basis for V of eigenvectors e_1, \dots, e_n of T .

Corollary: T is diagonalizable with respect to an orthonormal basis.

(*) Proof

Recall: We proved that a symmetric matrix A [corresponding to a self-adjoint transformation T_A] has an orthonormal basis of eigenvectors if A has at least 1 real eigenvalue. Then it suffices to show that A has at least 1 real eigenvalue.

2 approaches:

- (1) Prove that if \exists a complex eigenvalue for A , then that eigenvalue must be real.
- (2) Look at the map $\vec{x} \mapsto \langle A\vec{x}, \vec{x} \rangle$; we claim that the vector $\vec{x} \in \{\vec{x} : \|\vec{x}\|=1\}$ that maximizes $\langle A\vec{x}, \vec{x} \rangle$ is an eigenvector for A with a corresponding real eigenvalue.

(*) 2-Dimensional Case

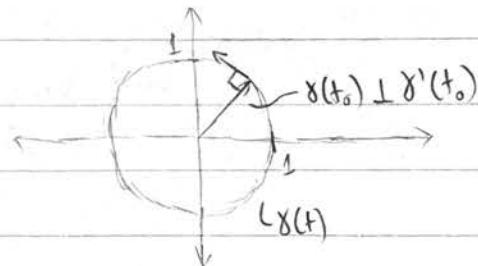
We know the set $\{\vec{v} : \|\vec{v}\|=1\}$ is the unit sphere.

Let $\gamma(t)$ be a parametrization of the unit circle.

Let A be the 2×2 symmetric matrix:

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix};$$

and let $f(\vec{x}) = \langle A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle = ax^2 + 2bx + cy^2$ $\forall \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \gamma(t)$. Let $\gamma(t_0)$ be the point in $\gamma(t)$ such that $f(\gamma(t_0))$ is the maximum value for $f(\vec{x})$ on $\gamma(t)$. Then we know $\frac{d}{dt} f(\gamma(t)) = 0$ at $t=t_0$ (since $\gamma(t_0)$ is a peak/maximum).



The Spectral Theorem (cont.)

10/27/23

Lecture 13

(*) 2-Dimensional Case (cont.)

(*) Notation: Let $\gamma(t_0) = \vec{x}_0 = \begin{pmatrix} x \\ y \end{pmatrix}$,

(cont.)

$$\text{We know } \left. \frac{\partial}{\partial t} f(\gamma(t)) \right|_{t=t_0} = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle \Big|_{t=t_0} = 0. \text{ We know } \left. \nabla f(\gamma(t)) \right|_{t_0} = \nabla \langle A\vec{x}(t), \vec{x}(t) \rangle \Big|_{t_0} = \nabla (ax^2 + 2bx + cy^2).$$

$$\Rightarrow \nabla f(\gamma(t)) = \left(\frac{\partial}{\partial x} (ax^2 + 2bx + cy^2), \frac{\partial}{\partial y} (ax^2 + 2bx + cy^2) \right)$$

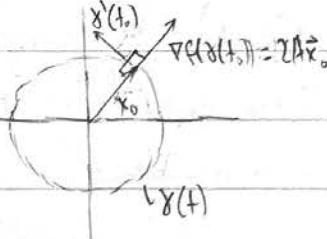
$$= 2(ax + by, bx + cy). \text{ Then } 2 \langle (ax + by, bx + cy), \gamma'(t) \rangle \Big|_{t_0} = 0; \text{ but}$$

$$(ax + by, bx + cy) = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A\vec{x}_0. \text{ Then } \langle A\vec{x}_0, \gamma'(t) \rangle = 0;$$

Hence $A\vec{x}_0 \perp \gamma'(t_0)$. But we know $\vec{x}_0 \perp \gamma'(t_0)$; therefore

$$A\vec{x}_0 \parallel \vec{x}_0, \text{ thus } A\vec{x}_0 = \lambda \vec{x}_0 \text{ [for some } \lambda]$$

$\Rightarrow \vec{x}_0$ is an eigenvector of A .

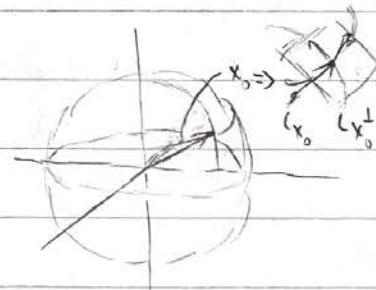


[This process is similar to Lagrange multipliers, or finding the axes of rotated ellipses].

(*) n-Dimensional Case

Let $\gamma(t)$ be the unit sphere in n dimensions.

Let $\gamma(t_0) = \vec{x}_0$ be the point on $\gamma(t)$ where $\langle A\vec{x}, \vec{x} \rangle$ is maximized.



Then the set of vectors orthogonal to \vec{x}_0 , \vec{x}_0^\perp , is a subspace of dimension $(n-1)$. We know $\left. \frac{\partial}{\partial t} \langle A\gamma(t), \gamma(t) \rangle \right|_{t=t_0} = 0$

$$\Rightarrow \left. \langle \frac{\partial}{\partial t} A\gamma(t), \gamma(t) \rangle + \langle A\gamma(t), \frac{\partial}{\partial t} \gamma(t) \rangle \right|_{t=t_0} = 0. \text{ [per chain rule]}$$

$= \langle A\gamma'(t), \gamma(t) \rangle + \langle A\gamma(t), \gamma'(t) \rangle$, Since A symmetric [$\langle v, Av \rangle = \langle Av, v \rangle$], we can move A :

$$\Rightarrow \left. 2 \langle \gamma'(t), A\gamma(t) \rangle \right|_{t=t_0} = 0. \text{ Then } A\vec{x}_0 \perp \gamma'(t_0) \wedge \{ \gamma : \gamma(t_0) = \vec{x}_0, \gamma(t) = \text{unit sphere} \}.$$

Since $\vec{x}_0 \perp \gamma'(t_0) \wedge \gamma'(t_0) \perp \vec{x}_0 \Rightarrow A\vec{x}_0 = \lambda \vec{x}_0$.

Multilinear Forms

10/30/27

Lecture 14 Multilinear Forms

Def: A multilinear form is a function that is linear in each variable, separately.

Interesting case: Want to look at multilinear forms on V (dim $V = n$) of the form $f: V^n \rightarrow F$ that are skew-symmetric/alternating (i.e. $f(v_1, \dots, v_i, v_j, \dots, v_n) = -f(v_1, \dots, v_j, v_i, \dots, v_n)$).

Ex: $w(e_1, e_2)$ on \mathbb{R}^2 . Know $w(e_1, e_1) = -w(e_2, e_2) \Rightarrow w(e_1, e_2) = 0$. Then $w(e_1, e_2)$ determines $w: w(ae_1 + be_2, ce_1 + de_2) = (ad - bc)(w(e_1, e_2)) = \det(A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}) w(e_1, e_2)$.
 (We have "defined" a determinant without using matrices, i.e. take $w(e_1, e_2) \mapsto \det\left(\frac{v_1}{v_2}\right) = w(v_1, v_2)$).

\rightarrow Set of 2-multilinear alternating forms $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is one-dimensional, determined by $w(e_1, e_2)$; determinant $= w(v_1, v_2)$, if we take $w(e_1, e_2) = 1$.
 \Rightarrow n -multilinear alternating forms determined by $w(v_1, \dots, v_n)$ [v_1, \dots, v_n independent vectors chosen arbitrarily in \mathbb{R}^n].

(*) Signum

Motivation: $w(e_1, e_2, e_3), w(e_3, e_2, e_1)$: different permutations of e_1, e_2, e_3 [after swaps]

\Rightarrow Define $\text{sgn}(\sigma)$: 1 if the # of swaps is even, -1 if the # of swaps is odd

$$\{1, \dots, n\} \xrightarrow{\quad} \{\sigma(1), \dots, \sigma(n)\}$$

(*) Proof: σ is well-defined (i.e. for any given permutation, \exists only 1 value for σ).

Look at permutation: $\{1, 2, \dots, n\} \xrightarrow{\quad} \{\sigma(1), \dots, \sigma(n)\}$. Then the number of inversions is the number of instances where $i < j$ but $\sigma(i) > \sigma(j)$ [alt: $\sigma(i) > \sigma(i+1)$] for the permutation $\sigma(1), \dots, \sigma(n)$. Then $(-1)^{\# \text{ inversions}} = \text{sgn}(\sigma)$.

$$\rightarrow w(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \underbrace{\text{sgn}(\sigma) w(e_1, \dots, e_n)}_{\text{determines } w}$$

Matrices of Linear Maps

10/31/23

Matrix Representations of Linear Maps

Disc 10

Let $T: V \rightarrow W$ be a linear map; let V, W be vector spaces with ordered bases $B = (x_1, \dots, x_n)$ and $\gamma = (y_1, \dots, y_r)$, respectively.

→ Define the $m \times n$ matrix $[T]_B^\gamma = (a_{ij})$, where $T(x_i) = \sum_{j=1}^r a_{ij} y_j$.

(*) Ex: Let $T: V \rightarrow W$. (Claim: \exists ordered bases B, γ such that $[T]_B^\gamma$ is diagonal with only 0s and 1s on the diagonal. [Note: $[T]_B^\gamma$ need not be square; call $[T]_B^\gamma$ "diagonal" if $a_{ij}=0 \forall i \neq j$])

(*) Proof: Begin by choosing a basis (y_1, \dots, y_r) for $\text{im}(T) \subset W$. Then its extension $\gamma = (y_1, \dots, y_n)$ is a basis for W . Choose (x_1, \dots, x_r) in V such that $y_i = T(x_i) \forall 1 \leq i \leq r$. Next, choose a basis (x_{r+1}, \dots, x_n) for $\ker(T)$. Then $B = (x_1, \dots, x_n)$ is a basis for V & $[T]_B^\gamma$ as desired. □

(*) Ex: Let $T: V \rightarrow V$. (Claim: \exists a basis B such that $[T]_B$ (i.e. $[T]_B^B$) as before, if and only if $T^2 = T$ [T is a linear projection].

(*) Proof:

(\Rightarrow) Suppose $\exists B$ with $[T]_B$ as before. Let $B = (x_1, \dots, x_n)$. Then $T(x_i) = \sum_{j=1}^n a_{ji} x_j = a_{ii} x_i, a_{ji} \in \{0, 1\}$. We claim that $[T^2]_B = [T]_B$. $T^2(x_i) = T(T(x_i)) = T(a_{ii} x_i) = a_{ii} T(x_i) = a_{ii}^2 x_i = T(x_i)$. □

(\Leftarrow) Suppose $T^2 = T$. (Choose a basis (x_1, \dots, x_r) for $\text{im}(T)$ and (x_{r+1}, \dots, x_n) for $\ker(T)$). We claim $B = (x_1, \dots, x_n)$ is a basis for V w/ $[T]_B$ as desired. Let $z_i \in V$ be such that $T(z_i) = x_i$. Then $x_i = T(z_i) = T^2(z_i) = T(x_i)$. □

(*) Def: Let $T: V \rightarrow V$. Then a subspace $W \subset V$ is called T -invariant if $T(W) \subset W$.

Multilinear Forms (cont.)

11/1/23

Lecture 15 Multilinear Forms (cont.)

Let $\omega: V^n \rightarrow \mathbb{R}$ [dim $V = n$] be an alternating multilinear form. Let v_1, \dots, v_n be a basis for V , and assume $\omega(v_1, \dots, v_n) = 1$. [$\Rightarrow \omega$ analogous to \det in \mathbb{R}^n].

Proposition: For any $w_1, \dots, w_n \in V$: $\omega(w_1, \dots, w_n) = 0 \Leftrightarrow w_1, \dots, w_n$ dependent.

(\Leftarrow) Proof

(\Leftarrow) follows from multilinearity ($\text{dependent} \Rightarrow$ contains a term in $\omega(w_1, \dots, w_n)$ to 0).

(\Rightarrow) Want to prove contrapositive: w_1, \dots, w_n independent $\Rightarrow \omega(w_1, \dots, w_n) \neq 0$.

Assume WLOG that $V = \mathbb{R}^n$. We know we can represent each $w_i \in V$

relative to v_1, \dots, v_n : $w_i = \sum_{j=1}^n a_{ij} v_j$.

$$\Rightarrow \omega(w_1, \dots, w_n) = \omega((a_{11}, \dots, a_{1n}), \dots, (a_{n1}, \dots, a_{nn}))$$

\rightarrow Analogous to a matrix: $(\omega) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$.

We know: (i) Interchanging "rows" [vectors] preserves nonzero ω

(ii) Adding multiples of rows to each other does not affect ω

(iii) Scalar multiplying rows, scalar multiplies ω

We can reduce $\omega((a_{11}, \dots, a_{1n}), \dots, (a_{n1}, \dots, a_{nn}))$ to "RREF" [since w_1, \dots, w_n are independent] with the above operations. Then:

$$\omega(w_1, \dots, w_n) = \underset{\text{"nonzero factors"} \cdot}{\omega(I_n)} \neq 0. \quad \square$$

$$[\underset{\text{"nonzero factors"} \cdot}{\omega(v_1, \dots, v_n)} = \underset{\text{"nonzero factors"} \cdot}{\omega(I_n)}]$$

Consequence: The following are equivalent: (i) A is invertible,

(ii) $\det(A) \neq 0$,

(iii) The rows of A are independent,

(iv) The columns of A are independent.

(*) Matrix Examples

11/2/23

Problem: Show that for every 3×3 orthogonal matrix A , $\exists x \in \mathbb{R}^3$ s.t. $x \neq 0$ and $A\hat{x} = \pm \hat{x}$.

Disc 11

(*) Proof: Recall: "Orthogonal" \Leftrightarrow preserves inner products, lengths. (Note: Orthogonal matrix \Rightarrow real)

We want to find an eigenvector for A with eigenvalue ± 1 . First: VTS \exists an eigenvector.

Recall: square matrix singular (not invertible) iff $\det B = 0$. Since eigenvectors are nonzero solutions to the equations $(A - \lambda I)\vec{v} = 0$, then A has eigenvectors iff $\det(A - \lambda I)$ has roots.

Observe: $\det(A - \lambda I)$ is a cubic polynomial w/ leading coefficient 1. Then

$\lim_{\lambda \rightarrow -\infty} \det(A - \lambda I) = -\infty$; $\lim_{\lambda \rightarrow \infty} \det(A - \lambda I) = +\infty$. Then per the Intermediate Value Theorem

[calculus], $\exists \lambda_0$ s.t. $\det(A - \lambda I) = 0 \Rightarrow \exists$ an eigenvector \vec{v} .

A orthogonal $\Rightarrow \langle v, v \rangle = \langle v, Av \rangle \Rightarrow \langle v, v \rangle = \langle v, \lambda v \rangle \Rightarrow \lambda = \pm 1$. \square

Problem: Is there a 3×3 matrix $A \in F^{3 \times 3}$ over F such that $A^3 \neq 0$, but $A^4 = 0$? [A: No]

(*) Proof: Assume, for the sake of contradiction, that $A \in F^{3 \times 3}$ is such a matrix. Since $A^3 \neq 0$, $\exists x \in F^3$, $x \neq 0$ such that $A^3 x \neq 0$. [Ex: Take e_1, e_2, e_3 ; one of them will satisfy $A^3 e_i \neq 0$].

(Claim: (x, Ax, A^2x, A^3x) are linearly independent.)

(*) Proof: Suppose $a_0, \dots, a_3 \in F$ s.t. $\sum_{i=0}^3 a_i A^i x = 0$. Then $0 = A^3 \cdot 0 \Rightarrow A^3 (\sum_{i=0}^3 a_i A^i x) = 0$; since $A^3 (A^i x) = 0 \forall i \geq 1$, then $a_0 A^0 x = 0$. But $A^0 x \neq 0 \Rightarrow a_0 = 0$. Repeat argument for a_1, a_2, a_3 .

Then (x, Ax, A^2x, A^3x) are an independent set of size 4 inside a 3-dimensional space [contradiction]. \square

Fundamental Theorem of Algebra

11/3/23

Lecture 16

Q: Given linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ with corresponding matrix $A \in \mathbb{R}^{n \times n}$, is it possible to find a basis that makes A diagonal? $\left[T(a_1, \dots, a_n) = A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \right]$

\Rightarrow Let $T^A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ be the linear transformation defined by $T^A(a_1, \dots, a_n) = A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$.

We know we can find an eigenvalue $\lambda \in \mathbb{R}$ for T if we can find a real root for $\det(A - \lambda I) = 0$

\rightarrow We can find an eigenvalue for T^A if we can find a root $\lambda \in \mathbb{C}$ for $\det(A - \lambda I) = 0$.

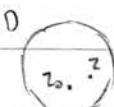
Prop: For any $A \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$), $\exists \lambda \in \mathbb{C}$ such that λ is an eigenvalue for A .

(*) Reason: We know $\det(A - \lambda I)$ is a polynomial $p(\lambda)$ of degree $n > 0$.

\rightarrow Per the Fundamental Theorem of Algebra, $p(\lambda)$ has a complex root.

(*) Proof: Let D be a closed disk with radius R large enough that p attains its minimum on D ($\exists z_0 \in D$ s.t. $|p(z)| \leq |p(z)| \forall z \in D$)

Claim: $p(z_0) = 0$. $|p(z_0)|$



(*) Proof: Assume $p(z_0) \neq 0$. Then for $z \in D$: $p(z) = p((z-z_0) + z_0)$

$\rightarrow p(z) = p(z_0) + a_k(z-z_0)^k + \dots + a_n(z-z_0)^n$. For z near z_0 , higher-order terms drop

out $\rightarrow p(z) \approx p(z_0) + a_k(z-z_0)^k$, $(z-z_0) \neq 0$. Namely: $\left| \frac{p(z) - p(z_0) + a_k(z-z_0)^k}{(z-z_0)^{k+1}} \right| \leq M$ for

some $M > 0$ in some neighborhood of z_0 . Rearranging terms:

$$|p(z)| \leq |p(z_0) + a_k(z-z_0)^k| + M|(z-z_0)|^{k+1}$$

For $(z-z_0)$ sufficiently small, can find z s.t. $a_k(z-z_0)^k < 0$, $|a_k(z-z_0)^k| > |M(z-z_0)|^{k+1}|$

$$\rightarrow |p(z)| \leq |p(z_0) + a_k(z-z_0)^k| + M|(z-z_0)|^{k+1} < |p(z_0)|. \square$$

(*) Misc. Notes

MT Review

Misc. Notes

Claim: Matrix $M(T^*)$ of $T^*: W^* \rightarrow V^*$ is $M(T^*) = (M(T))^T$, $T: V \rightarrow W$.

Proof: Let e_1, \dots, e_n standard basis for V , $\theta_1, \dots, \theta_m$ dual basis for standard basis on W .

$$\rightarrow (M(T))_{ij} = \theta_i(T(e_j)) = \theta_i(\varphi_i(e_j)) \quad [\varphi_1, \dots, \varphi_n \text{ dual basis for } e_1, \dots, e_n \text{ on } V]$$

$$\text{Observe: } \theta_i(T(e_j)) = (T^*(\theta_i))(e_j) = \varphi_i(T^*(\theta_i)) = (M(T^*))_{ji}, \quad (M(T))^T_{ji} = (M(T^*))_{ji}. \quad \square$$

Claim: Row rank = column rank.

Proof: Know: column rank of $T = \dim(\text{im}(T))$, $[T: V \rightarrow U, \dim V = n, \dim U = n]$

Row rank of $T = \text{column rank of } T^* = \dim(\text{im}(T^*))$.

Recall: $\ker(T^*) = \text{ann}(T(V))$, $\dim(\text{ann}(T(V))) = \dim W - \dim T(V)$

$$\begin{aligned} \rightarrow \dim(\text{im}(T^*)) &= \dim W^* - \dim \ker T^* \\ &= \dim V^* - \dim(\text{ann}(T(V))) \\ &= \dim V^* - (\dim U - \dim(\text{im}(T))) = \dim(\text{im}(T)). \quad \square \end{aligned}$$

r preview: Jordan form

Claim: Any 2×2 matrix in \mathbb{C} is either diagonalizable or can be placed into the form $\begin{pmatrix} \lambda^2 \\ 0 \lambda \end{pmatrix}$.

Proof: Case 1: Matrix A has 2 distinct eigenvalues \rightarrow diagonalizable.

Case 2: Matrix A has only 1 eigenvalue.

2(a): A has 2 independent eigenvectors \rightarrow diagonalizable.

2(b): A has only 1 eigenvector. Let v be the eigenvector w/ eigenvalue λ .

\rightarrow Can find $y \notin \ker(A - \lambda I)$; observe: v, y independent \rightarrow form a basis.

$\rightarrow (A - \lambda I)y = av + bw$ for some $a, b \in \mathbb{C} \rightarrow (A - \lambda I)^2 y = b(A - \lambda I)y$

$\rightarrow A(A - \lambda I)y = (b + \lambda)(A - \lambda I)y \rightarrow b + \lambda$ eigenvalue for A .

But λ is the only eigenvalue $\Rightarrow b = 0 \Rightarrow (A - \lambda I)y = av$.

Let $\bar{y} = \frac{1}{a}y \rightarrow (A - \lambda I)\bar{y} = v \Rightarrow A\bar{y} = v + \lambda\bar{y}$.

\rightarrow Matrix of A : $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. \square

Diagonalization (cont.)

11/8/23

Lecture 17

Lemma: Let $\vec{v}_1, \dots, \vec{v}_n$ be eigenvectors associated with distinct eigenvalues $\lambda_1, \dots, \lambda_n$. Then $\vec{v}_1, \dots, \vec{v}_n$ are linearly independent.

Corollary: If a matrix $A \in \mathbb{R}^{n \times n}$ has n distinct eigenvalues (i.e. $\det(A - \lambda I)$ has n distinct roots), then A is diagonalizable. [Note: converse does not hold!]

(*) Proof: Let v_1, \dots, v_n be eigenvectors for A , with $\lambda_1, \dots, \lambda_n$ all distinct. Assume, for the sake of contradiction, that v_1, \dots, v_n are dependent. Then we can find a shortest relation $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ [$k \leq n$]. We know: $T(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_n \lambda_n v_n = 0$.

Also know: $(\lambda_1)(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_n \lambda_1 v_n$. Subtract $(T - \lambda_1 I)(\alpha_1 v_1 + \dots + \alpha_n v_n)$.

$\alpha_1 (\lambda_1 - \lambda_1) v_1 + \dots + \alpha_n (\lambda_1 - \lambda_1) v_n = 0$. Then $\alpha_1 v_1 + \dots + \alpha_n v_n$ was not the shortest relation. \square

(*) Note: we know $\lambda_i - \lambda_j \neq 0 \forall i \neq j$ because we assumed λ s distinct

Def: Eigenspace

Let λ be an eigenvalue for a matrix A . Then we define the eigenspace for λ : $E_\lambda = \ker(A - \lambda I)$, containing all linear combinations of eigenvectors of A associated with eigenvalue λ .

(*) Diagonalization

Fact: for any matrix A that is not diagonalizable, we can find another matrix A' arbitrarily close to A that is diagonalizable. [$A, A' \in \mathbb{R}^{n \times n}$]

(*) Formally: for any $\epsilon > 0$, $\exists A'$ such that $|a_{ij} - \bar{a}_{ij}| < \epsilon \forall i, j$ and A' has n distinct eigenvalues.

(*) Consequence: Non-diagonalizable matrices are "sparse".

→ Applied cases: Every matrix is, for practical purposes, "diagonalizable"; we want perfect solutions to remove uncertainty/error, avoid unstable equilibria.

(*) Ex: Eigenvectors for "nearly the same" eigenvalues (only very slightly different); may "converge"

Jordan Form

11/8/23

Lecture 17

(cont.)

Representing Non-Diagonalizable Matrices

In the case of non-diagonalizable matrices, we can still represent them in an "almost diagonal" form: Jordan canonical form.

↳ Builds matrices from Jordan blocks: $\begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}$ [λ some eigenvalue]

→ For any T^c , can find a basis B such that $[T]_B$ in Jordan form:

$$[T^c]_B = \begin{pmatrix} \lambda_1 & & 0 & 0 & 0 \\ 0 & \lambda_2 & & 0 & 0 \\ 0 & 0 & \ddots & & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

Jordan blocks for eigenvalues $\lambda_1, \dots, \lambda_n$

(*) Ex: $\begin{pmatrix} 2 & 0 & 0 \\ 0 & \begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{matrix} & 0 \\ 0 & 0 & \begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix} \end{pmatrix}$ [matrix in Jordan form]

(*) Properties of Jordan Form

- (i) Each block is independent of the others
- (ii) Ordering does not matter - blocks can be switched around freely, e.g.
- (iii) Every eigenvalue has at least 1 associated block, may have multiple

Finding Jordan Forms

In order to find a Jordan form for $T: V \rightarrow V$, need to split $V = \mathbb{C}^n$ into pieces [Jordan blocks]/subspaces corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$:

Idea: Use eigenspaces $E_{\lambda_i} = \ker(A - \lambda_i I)$.

We know: $\ker(A - \lambda_i I) \subset \ker(A - \lambda_i I)^2 \subset \dots$ eventually stabilizes to some matrix

→ union of $\ker(A - \lambda_1 I), \ker(A - \lambda_2 I), \dots$ = generalized eigenspace for λ :

(*) Eigenspace Examples

11/01/23

Disc 12 Problem: Let $A \in F^{n \times n}$ w/ $\text{char}(F) \neq 2$. Suppose $A^2 = I$. Show that A is diagonalizable; check for $\text{char}(F) = 2$.

(a) Observe: $A^2 - I = (A + I)(A - I) = 0 \rightarrow \text{one of } \ker(A + I), \ker(A - I) \text{ nontrivial. Per}$

Rank-Nullity: $n = \text{rk}(A - I) + N(A - I) \leq N(A + I) + N(A - I)$. Since $I \neq -I$ [$\text{char}(F) \neq 2$], then these are distinct eigenspaces w/ eigenvalues P_1, P_2 s.t. $P_1 \cup P_2$ are independent and span F^n ; then \exists basis of eigenvectors of $A \rightarrow A$ diagonalizable. \square

(b) Consider: (''): then $A^2 = I$, but Jordan block size $> 1 \Rightarrow$ not diagonalizable.

V4A Basic

F21 #7: Let $A, B \in C^{n \times n}$ such that $AB = BA$. Show that A, B share an eigenvector.

Since $F = C$, then per F.T. of Algebra: $\det(A - \lambda I), \det(B - \lambda I)$ have at least 1 root [eigenvalue].

Let $x_0 \in C^n$ be an eigenvector of A ($Ax_0 = \lambda x_0$). matrices commute \Rightarrow preserve eigenspaces

Let E_λ be the eigenspace $E_\lambda = \{x \in C : Ax = \lambda x\}$. (Claim: $B(E_\lambda) \subseteq E_\lambda$ (i.e. E_λ B -invariant).)

(* Proof: Let $x \in E_\lambda$. Then $A(Bx) = B(Ax) = B(\lambda x) = \lambda Bx \rightarrow Bx \in E_\lambda$. \square

Then $B(E_\lambda) \subset E_\lambda \neq \emptyset$. Then $B|_{E_\lambda} [B : E_\lambda \rightarrow E_\lambda]$ has an eigenvector (per F.T. of Alg.) $x \in E_\lambda$, such

that $Bx = \lambda_2 x$. Since $x \in E_\lambda$, x is an eigenvector for both A, B . \square

F17 #2: Let $A, B \in F^{n \times n}$ s.t. A invertible; show that AB, BA have the same eigenvalues

w/ the same geometric multiplicities (i.e. dim E_λ 's).

Observation: Let C, D square, C invertible; then $\ker(CD) = \ker(D) \wedge \ker(DC) = C^{-1} \cdot \ker(D)$.

$\rightarrow \text{null}(CD) = \text{null}(D) = \text{null}(DC)$.

Compare $\text{null}(AB - \lambda I), \text{null}(BA - \lambda I)$, $\text{null}(BA - \lambda I) = \text{null}(A(BA - \lambda I))$

$= \text{null}(A(BA - \lambda I)A^{-1}) = \text{null}(AB - \lambda I)$ for any $\lambda \in F$.

Matrix Polynomials

11/13/23

Lecture 18

Polynomials of Matrices

We know we can take exponents of linear transformations, e.g. $T: \vec{x} \mapsto A\vec{x} \Rightarrow T^2: \vec{x} \mapsto A^2\vec{x}$.

Recall: polynomials with coefficients in \mathbb{F} form a ring of expressions $p(x) = a_nx^n + \dots + a_0$.

→ Can define a map $\{p(a)\} \rightarrow \text{Hom}(V, V)$ by, given $T: V \rightarrow V$, mapping:

$$p(x) = a_nx^n + \dots + a_0 \mapsto p(T) = a_nT^n + \dots + a_1T + a_0I_n$$

is a ring homomorphism

→ Q: Does $\exists p(x) \neq 0$ such that $p(T) = 0$? [Answer: Yes]

Theorem (Cayley-Hamilton): Let A be an $n \times n$ matrix, and let $p(\lambda) = \det(A - \lambda I)$.
Then $p(A) = 0$.

(*) Proof (pt. 1)

Claim: For any matrix A , \exists a polynomial $p(x)$ such that $p(x) \neq 0$, but $p(A)$ is the zero transformation.

Proof: Let A be the matrix corresponding to $T: V \rightarrow V$. Let v_1, \dots, v_n be a basis for V .

For each v_i , we want a polynomial $p_i(x)$ such that $p_i \neq 0$, but $p_i(A)v_i = \vec{0}$.

Look at the set of vectors $v_i, Tv_i, \dots, T^n v_i$. Since there are $n+1$ vectors in the set (and $\dim V = n$),

therefore $\exists a_0, \dots, a_n \neq 0$ such that $a_0v_i + \dots + a_nT^n v_i = \vec{0}$. Set the polynomial

$$p(x) = a_0 + a_1x + \dots + a_nx^n; \text{ then } p_i(T)v_i = \vec{0} \quad [\text{but } p_i(x) \neq 0]$$

We can find such a polynomial p_i for every vector v_i , $1 \leq i \leq n$. Define

$p_A(x) = p_1(x)p_2(x)\dots p_n(x)$ to be the product of the polynomials p_1, \dots, p_n .

Then $p_A(x) \neq 0$, but for any v_i , $p_A(A)v_i = (p_1(A)\dots p_n(A))(p_i(A)v_i) = \vec{0}$. \square

(*) Note: $\deg(p_A) = n^2$; eventually, we want a polynomial of degree n .

Matrix Polynomials (cont.)

11/17/23

Lecture 18 (*) Proof (Pt. 2)

Claim: The minimal polynomial [polynomial of least degree] $p(x) \neq 0$ s.t. $p(A) = 0$ is unique.

(*) Proof

Lemma: Suppose $P(x)$ is the minimal polynomial. Let $Q(x)$ be a polynomial such that $Q(x) \neq 0$, but $Q(A) = 0$. Then $P \mid Q$, i.e. $\exists Q_1(x)$ such that $Q = Q_1 \cdot P$.

Observation: $(Q \cdot P)(T) = 0$ for any polynomial Q [$(Q \cdot P)(T) = Q(P(T)) = Q(0)$].

(*) Polynomial Division

Let $P(x), Q(x)$ be two polynomials such that $\deg Q \geq \deg P$. Then \exists polynomials $Q_1(x)$, $R(x)$ ($\deg R < \deg P$) such that $Q = Q_1 P + R$,

(*) Proof

Case 1: $\deg Q = \deg P \Rightarrow$ trivial

Case 2: $\deg Q > \deg P \Rightarrow$ take $Q_1 = \alpha x^k$ ($k = \deg Q - \deg P$)

Let $P(x)$ be the minimal polynomial. Let $Q(x) \neq 0$ be a polynomial such that $Q(A) = 0$; then $\deg Q \geq \deg P$. Then \exists polynomials Q_1, R such that $Q = Q_1 P + R$.

Since $Q(T) = 0 = Q_1(P(T)) + R(T) = 0 + R(T)$, therefore $R(x) = 0$ [since $\deg R < \deg P$].

Then $Q = Q_1 P$; then either $\deg Q_1 > 0 \Rightarrow \deg Q > \deg P$ (i.e. Q is not a minimal polynomial), or $\deg Q_1 = 0 \Rightarrow Q_1$ is a scalar multiple of P . \square

(Proof continues until p. 42)

(*) Diagonalization Examples

11/14/23

Problem: Suppose $T: V \rightarrow V$ is a diagonalizable linear operator on fin. dim. V.S. V over \mathbb{F} . If $W \subset V$ is a nonzero T -invariant subspace, show that $T|_W$ is diagonalizable.

We want to find a basis for W of eigenvectors for $T|_W$. Then it suffices to find any finite set of eigenvectors for $T|_W$ spanning W . Claim: Let $x_1, \dots, x_r \in V$ be T -eigenvectors with distinct eigenvalues. Then if $\sum_{i=1}^r x_i \in W$, then $x_1, \dots, x_r \in W$.

(*) Proof: By induction. Base case: $r=1$ [trivial]. Inductive step: Since $T(W) \subset W$ and $(\lambda_1 \cdot I)(W) \subset W$, then $(T - \lambda_1 \cdot I)(W) \subset W$. Let $x \in W$ s.t. $x = \sum_{i=1}^r x_i$. Then $(T - \lambda_1 \cdot I)(\sum_{i=1}^r x_i) = \sum_{i=1}^r (T - \lambda_1 \cdot I)(x_i) = \sum_{i=2}^r (\lambda_i - \lambda_1)x_i = \sum_{i=2}^r (\lambda_i - \lambda_1)x_i$ (we have reduced the sum by 1 term).

Through induction, we can keep reducing to find $x_2, \dots, x_r \in W$. Then $(\lambda_2 - \lambda_1)x_2, \dots, (\lambda_r - \lambda_1)x_r \in W \Rightarrow x \in W$. \square

Choose a basis y_1, \dots, y_m for W ; since T is diagonalizable on V , then for each y_i ($1 \leq i \leq m$), we can find eigenvectors $x_{i,1}, \dots, x_{i,r_i}$ w/ distinct eigenvalues s.t. $y_i = \sum_{j=1}^{r_i} x_{i,j}$; then $x_{i,j} \in W \forall i, j$. Then $W = \text{span}\{y_1, \dots, y_m\} = \text{span}\{x_{1,1}, \dots, x_{m,r_m}\}$, i.e. W is spanned by eigenvectors of T . \square

Problem: Let $T, S: V \rightarrow V$ be diagonalizable lin. ops. such that $T \circ S = S \circ T$. Show that S, T are simultaneously diagonalizable, i.e. \exists basis B for V s.t. $[S]_B, [T]_B$ are both diagonal.

Recall: S, T commute $\Rightarrow S, T$ preserve each other's eigenspaces.

Observe: Let x be an eigenvector of S , s.t. $Sx = \lambda x$. Then $S(Tx) = TSx = \lambda Tx$.

Since S is diagonalizable, we know \exists distinct $\lambda_1, \dots, \lambda_r$ s.t. $V = \bigoplus_{i=1}^r E_{\lambda_i}$, and that each E_{λ_i} is T -invariant. Since T diagonalizable, then $T|_{E_{\lambda_i}}$ is diagonalizable $\forall 1 \leq i \leq r$. For each $1 \leq i \leq r$, we can choose a basis $\{x_{i,1}, \dots, x_{i,s_i}\}$ for E_{λ_i} of eigenvectors for $T|_{E_{\lambda_i}}$. Then $B = (x_{1,1}, \dots, x_{r,s_r})$ is a basis for V . Since each vector in the basis is an eigenvector of S , then B is an eigenbasis for S ; since each vector is an eigenvector for $T|_{E_{\lambda_i}}$ for some i , then B is an eigenbasis for T . \square

Matrix Exponentials

11/15/23

Lecture 19 Matrix Exponentials

Recall: $e^x [\exp(x)] = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$

→ Def. Let $A \in \mathbb{R}^{n \times n}$. Then we define the matrix exponential $\exp(A)$:

$$e^A = I_n + A + \frac{1}{2!} A^2 + \dots + \frac{1}{n!} A^n + \dots \quad \left[= \sum_{i=0}^{\infty} \frac{1}{i!} A^i \right]$$

Intuition: e^A eventually converges \Rightarrow where?

Operator Norm

Def. Let $A \in \mathbb{R}^{n \times n}$. Then we define the operator norm of A : $\|A\| := \sup\{\|T_A \vec{v}\| : \|\vec{v}\|=1\}$

Observation: $\|A\| \geq |a_{ij}| \forall a_{ij}$ in A ($\|A\|$ is a "bound" on the elements in A)

Prop. $\|AB\| \leq \|A\| \cdot \|B\|$.

(*) Proof: $\|AB\| \leq \|A(B\vec{v})\|$. Know: $\|B\vec{v}\| \leq \|B\| \cdot \|\vec{v}\|$.

$$\Rightarrow \|AB\| \leq \|A\| \cdot \|B\| \leq \|A\| \cdot \|B\| \cdot \|\vec{v}\|.$$

Corollary: e^{tA} converges ($t \in \mathbb{R}$).

(*) Proof: We know $\|A^n\| \leq \|A\|^n \quad \forall n \in \mathbb{N}$ from the proposition. Additionally, we know

$(A^n)_{ij} \leq \|A^n\| \quad \forall n \in \mathbb{N}, \text{ any element } i, j$. Then we can show each element converges:

$$(e^{tA})_{ij} = I_{ij} + (tA)_{ij} + \dots + \frac{1}{n!} (tA)^n_{ij} \leq 1 + \|A\| + \dots + \frac{1}{n!} \|A\|^n = e^{\|A\| t}. \square$$

(*) Note: Each element of e^{tA} is a differentiable function of t . [Analysis]

(*) Matrix Logarithms

Let $A \in \mathbb{R}^{n \times n}$. Then $\exists B = \log(A)$ [i.e. $B \in \mathbb{R}^{n \times n}$ s.t. $e^B = A$] only if e^A is invertible, and

B is a matrix of the form $B = I + C$ [$\|C\|$ small; $\|C\| \leq 1$, e.g.].

(*) Analogy: $\ln(1+x)$ only exists if $|x| < 1 \Rightarrow \ln(1+x) = 1 + x + \frac{1}{2}x^2 + \dots$ (apply to matrices)

(*) Jordan Form & Diagonalization Examples

11/16/23

Problem: Suppose $T, S: V \rightarrow V$ are linear operators on fin. dim. V over F s.t. T has n distinct eigenvalues and T, S commute. Show that S is diagonalizable.

Disc 14

Recall: T, S commute \Rightarrow preserve eigenspaces

$$\text{Observe: } \dim E_{\lambda_i} = 1 \quad \forall \lambda_1, \dots, \lambda_n \quad [\dim \bigoplus E_{\lambda_i} = \sum_{i=1}^n \dim E_{\lambda_i} = n]$$

Claim: T -eigenvectors span 1D S -dimensional subspaces.

(*) Proof: Pick T -eigenbasis $\{x_1, \dots, x_n\}$ for V s.t. $T(x_i) = \lambda_i x_i$. We know $E_{\lambda_i} = \text{span}\{x_i\}$

and $S(E_{\lambda_i}) \subset E_{\lambda_i} = \text{span}\{x_i\} \rightarrow S(x_i) \in \text{span}\{x_i\} \rightarrow x_i$ is an eigenvector for S . \square

Jordan Canonical Form

Recall: Given matrix $A \in F^{N \times N}$, we obtain a linear map $L_A: F^n \rightarrow F^n$ defined by $L_A(x) = Ax$. If $S = \{e_1, \dots, e_n\}$ standard basis for F^n , then $[L_A]_S = A$.

Recall: Given composable linear maps $U \xrightarrow{T} V \xrightarrow{S} W$ and finite bases α, β, γ for U, V, W (respectively)

$$\rightarrow [S \circ T]_{\gamma}^{\alpha} = [S]_{\beta}^{\gamma} \cdot [T]_{\alpha}^{\beta}$$

(*) Recall: For any basis β of fin. dim. V , $[I \otimes v]_{\beta} = I$.

\rightarrow Lemma: Let β, γ bases for fin. dim. V ; then $[I \otimes v]_{\gamma}^{\beta} = ([I \otimes v]_{\beta}^{\gamma})^{-1}$. [Change of coordinates]

Problem: Let $A \in \mathbb{C}^{N \times N}$. Show that A, A^T are similar.

Choose Jordan basis $\beta = \beta_1 \cup \dots \cup \beta_r$, where $\beta_i = \{x_{i1}, \dots, x_{is_i}\}$ and $\exists \lambda_i \in \mathbb{C}$ such that

$$Ax_{ij} \begin{cases} \lambda_i x_{i,j+1} & i = 1 \\ x_{i,j-1} + \lambda_i x_{i,j} & j > 1 \end{cases} \quad \text{Let } Q = [I \otimes \beta]_{\beta}^{\beta}, \text{ where } \beta \text{ is the standard basis.} \quad \beta_i \text{ reversed}$$

$$A = [L_A]_{\beta}^{\beta} = [I \otimes \beta]_{\beta}^{\beta} [A]_{\beta}^{\beta} [I \otimes \beta]_{\beta}^{\beta} = Q [L_A]_{\beta}^{\beta} Q^{-1}. \quad \text{Let } \gamma = \gamma_1 \cup \dots \cup \gamma_r, \quad \gamma_i = (x_{i1}, \dots, x_{i-1}).$$

$$\text{Want: } A_{ij} = Ax_{i(j-1)} \begin{cases} \lambda_i x_{i,j+1} & i = 1 \\ x_{i,j-1} + \lambda_i x_{i,j} & j > 1 \end{cases} \rightarrow [L_A]_{\gamma}^{\gamma} = ([L_A]_{\beta}^{\beta})^T = (Q^{-1}AQ)^T$$

$$\rightarrow = Q^T A^T (Q^T)^{-1}$$

Polynomial Algebra

11/17/23

Lecture 20

Polynomial Algebra

Recall (Euclid's Algorithm): $\gcd(a, b) [a > b] = \gcd(b, a \bmod b)$

→ Consequence: $\gcd(a, b) = ax + by$ for some $x, y \in \mathbb{R}$.

Euclid's Algorithm works for polynomials: $\gcd(P_1, P_2) = D(x)$, such that $D = A_1 P_1 + A_2 P_2$
for some A_1, A_2 polynomials (Notation: $P_1 | Q \Rightarrow Q = A \cdot P_1$)

(*) Euclid's Algorithm for Polynomials

Let P_1, P_2 polynomials; $\deg P_1 \geq \deg P_2$.
 $\left(\begin{array}{l} \text{by } P_2 \leq \deg P_1 \\ \text{by } P_2 < \deg P_2 \end{array} \right)$

→ $P_1 = QP_2 + R$ ($\deg R < \deg P_2$) $\Rightarrow \gcd(P_1, P_2) = \gcd(P_2, R)$

→ Continue until $R=0$; then preceding R is the gcd (up to constant scalar multiple).

"relatively prime"

Corollary: $\gcd(P_1, P_2) = 1$ (no common factors) $\Rightarrow 1 = A_1 P_1 + A_2 P_2$

(*) Ex: $P_1 = (x-\lambda_1)^k, P_2 = (x-\lambda_2)^k$ [$\lambda_1 \neq \lambda_2$] $\Rightarrow 1 = A_1(x) \cdot (x-\lambda_1)^k + A_2(x) \cdot (x-\lambda_2)^k$

Polynomials & Matrices

Suppose P_1, P_2 relatively prime (i.e. $\gcd(P_1, P_2) = 1$), $A \in \mathbb{R}^{n \times n}$

$\Rightarrow I = A_1(T_A)P_1(T_A) + A_2(T_A)P_2(T_A)$. Then, for $v \in \mathbb{R}^n$: $v = A_1(T)v + A_2(T)P_2(T)v$.

Then $v \in (\ker P_1(T)) \cap \ker P_2(T) \Rightarrow v = 0$; then $\ker P_1(T) \cap \ker P_2(T) = \{0\}$.

$\Rightarrow \ker P_1(T), \ker P_2(T)$ are disjoint.

\Rightarrow Corollary: Let P_1, P_2 be relatively prime polynomials such that $P_1(T) \cdot P_2(T) = 0$.

Then $V = \ker P_1(T) \oplus \ker P_2(T)$. (**) Proof is left as an exercise to the reader.

(*) Def: Direct Sum

We say that $V = V_1 \oplus V_2$ if, $\forall v \in V, \exists$ a unique representation $v = v_1 + v_2$ ($v_1 \in V_1, v_2 \in V_2$).

(**) Ex: $\mathbb{R}^3 = \{(x, y, z) : x, y \in \mathbb{R}\} \oplus \{(0, 0, z) : z \in \mathbb{R}\}$

Polynomial Algebra (cont.)

11/20/23

Lecture 21

Polynomial Algebra (cont.)

Fact: Every polynomial is the product of linear factors over \mathbb{C} .

i.e.: $\deg P = n \Rightarrow P(x) = a(x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_n)$ [$a \in \mathbb{C} \setminus \{0\}$] is a unique representation for $P(x)$.

(*) Def: An irreducible polynomial is a polynomial that cannot be factored.

Ex: $(x - \lambda)$ in \mathbb{C} ; $(x - \lambda)$, $(x^2 + bx + c)$ [w/ no real roots] in \mathbb{R}

(*) Note: Polynomials in an arbitrary field F can still be expressed as the product of irreducibles, but the factors may not be linear. (Ex: $(x^2 + 1)$ in \mathbb{R}).

Claim: The factorization of a polynomial into irreducible factors is unique in any field F .

(*) Proof

Lemma: $p | ab \Rightarrow p | a$ or $p | b$.

(*) Proof: Suppose $p \nmid a$. Then $\gcd(p, a) = 1 \Rightarrow 1 = ai + pj$ for some i, j . Then $b = abi + pbj$. We know $p | ab \Rightarrow p | abi$, and $p | p \Rightarrow p | pbj$; then $p | b$. \square

Let $P(x) = p_1 \cdot p_2 \cdots p_k$ be a factorization of P . Let $P = q_1 \cdot q_2 \cdots q_k$ be another factorization. We claim that for every p_i , $p_i = q_j$ (down to scalar multiple) for some $1 \leq j \leq k$.

(Base case). We know $P \mid P = q_1 q_2 \cdots q_k$. Then $p_i = q_j$ for some j (we know that p_i is irreducible).

(Induc. step). Assume we have found $p_i = q_j$ for some j . Then $\prod_{i=2}^k p_i = \prod_{j=1}^{k-1} q_i$. Take

$P' = \prod_{i=2}^k p_i$; then we can find (same argument as prev.) that $p'_i = q'_i$ for some j .

Then every factor p_i is represented in $q_1 \cdot q_2 \cdots q_k$, therefore the factorizations are equivalent. \square

Matrix Decomposition

11/20/23

Lecture 21 Matrix Decomposition

(cont.)

Let $T_A: \mathbb{C}^n \rightarrow \mathbb{C}^n / \mathbb{R}^n \rightarrow \mathbb{R}^n$. Let $p(x)$ be the minimal polynomial such that $p(x) \neq 0$, but $p(T_A) = 0$.

$\rightarrow p(x)$ can be expressed as the product of irreducibles: $p(x) = p_1(x) \cdots p_k(x)$.

Assume $F = \mathbb{C}$; then all factors are linear: $p(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_n)^{k_n}$.

in \mathbb{C} :
 $(x - \lambda)$ a factor
 $\Rightarrow (x - \bar{\lambda})$ a factor

Observe: The zeros of the factors $(A - \lambda_i I)^{k_i}$ correspond to the generalized eigenspaces $G_A(\lambda_i)$ of A . Additionally, the functions $p_i(x) = (x - \lambda_i)^{k_i}$ are all relatively prime.

Corollary: $\mathbb{C} = G_A(\lambda_1) \oplus \dots \oplus G_A(\lambda_n)$, i.e. the generalized eigenspaces of A span \mathbb{C} .

(*) **Consequence:** Let P_i be a basis for each G.E. $G_A(\lambda_i)$. Then $B = [P_1 \mid \dots \mid P_n]$ is a basis for \mathbb{C} .

(*) Ex: Decomposition

Let $A \in \mathbb{C}^{n \times n}$. Let $p(x)$ be the minimal polynomial of A [of T_A], and let

$p(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_n)^{k_n}$ be the factorization of $p(x)$. Define $P_1 = (x - \lambda_1)^{k_1} \cdots (x - \lambda_{n-1})^{k_{n-1}}$,

$P_2 = (x - \lambda_n)^{k_n}$. We note that P_1, P_2 are relatively prime, i.e. $\gcd(P_1, P_2) = 1$.

$\rightarrow \mathbb{C} = \ker P_1(T) \oplus \ker P_2(T)$.

Lemma: $\ker P_1(T), \ker P_2(T)$ are T -invariant.

(*) **Proof ($\ker P_1(T)$):** Let $v \in \ker P_1(T)$. Then $Av = \lambda_1 v$ for some $1 \leq i \leq n-1$.

Then $P_1(A) \cdot Av = P_1(A) \lambda_1 v = \lambda_1 (P_1(A)v) = 0$. \square

Let v_1, \dots, v_m be a basis for $\ker P_1(T)$, and let v_{m+1}, \dots, v_n be a basis for $\ker P_2(T)$. Then

$B = [v_1 \mid \dots \mid v_m \mid v_{m+1} \mid \dots \mid v_n]$ is a basis for \mathbb{C}^n such that:

$$[T_A]_B = \left(\begin{array}{c|c} A(\ker P_1(T)) & 0 \\ \hline 0 & A(\ker P_2(T)) \end{array} \right) \xrightarrow{\text{decompose } P_1} \left(\begin{array}{c|c} \boxed{P_1} & \boxed{0} \\ \hline \boxed{0} & \boxed{P_2} \end{array} \right) \xrightarrow{\text{blocks correspond to gen. eigenspaces}} \left(\begin{array}{c|c} \boxed{I_m} & 0 \\ \hline 0 & \boxed{I_{n-m}} \end{array} \right)$$

(*) Cyclic Subspaces

11/21/23

Disc 15

Def: Cyclic Subspaces

Let $T: V \rightarrow V$ be a linear operator. We call a subspace $W \subset V$ "T-cyclic" if $\exists x \in W$ s.t. $W = \text{span}\{x, Tx, \dots\}$.

(*) Problem: Suppose $T: V \rightarrow V$ is a diagonalizable lin. op. on F -D VS V over F . Show that T has $n = \dim V$ distinct eigenvalues iff V is T-cyclic.

(*) Proof: (\Rightarrow) Let $\lambda_1, \dots, \lambda_n \in F$ be eigenvalues with a corresponding eigenbasis $B = (x_1, \dots, x_n)$.

Let $y = \sum_{i=1}^n x_i$, and let $y_j = T^{j-1} y$ ($1 \leq j \leq n$). Then $y_j = T^{j-1} \left(\sum_{i=1}^n x_i \right) = \sum_{i=1}^n \lambda_i^{j-1} x_i$.

Then $D = [y_1]_B \dots [y_n]_B = \begin{bmatrix} \lambda_1 & \lambda_1^{n-1} \\ \lambda_2 & \lambda_2^{n-1} \\ \vdots & \vdots \\ \lambda_n & \lambda_n^{n-1} \end{bmatrix} \rightsquigarrow$ nonzero determinant \Rightarrow invertible.

$$\Rightarrow V = \text{span}\{y_1, \dots, y_n\} = \text{span}\{y, Ty, \dots, T^{n-1} y\}.$$

(\Leftarrow) Suppose T has $r \leq n$ distinct eigenvalues. (Claim: $\dim(\text{span}\{x, Tx, \dots\}) \leq r$.)

(*) Proof: Let $x \in V$. T diagonalizable $\Rightarrow V = \bigoplus_{i=1}^r E_{\lambda_i} \Rightarrow x = \sum_{i=1}^r x_i, x_i \in E_{\lambda_i}$.

Then if $g(t) = \prod_{i=1}^r (t - \lambda_i)$, $g(T)(x) = \sum_{i=1}^r (\pi_{i,j} (T - \lambda_i I) \circ (T - \lambda_i I) x_i) = 0$, i.e. by induction

$$0 = g(T)(x) = T^r x + a_{r-1} T^{r-1} x + \dots + a_0 x \Rightarrow T^k x \in \text{span}\{T^{r-1} x, \dots, x\} \forall k \geq 0. \square$$

Generalizing: Let T, V as before. Let $f_T(t), g_T(t)$ be the characteristic and minimal polynomials of T , respectively; then V is T-cyclic iff $g_T(t) \mid f_T(t)$.

(*) Proof: (\Rightarrow) Suppose V is T-cyclic, with $x_0 \in V$ s.t. $V = \text{span}\{x_0, Tx_0, \dots\}$. Let $r = \deg(g_T)$, and let $n = \dim(V) = \deg(f_T)$. We know g_T, f_T are both monic and $g_T \mid f_T$ (per (-H)).

We know $g_T(T)(x_0) = 0$; then $V = \text{span}\{x_0, Tx_0, \dots\} = \text{span}\{x_0, \dots, T^{r-1} x_0\}$.

(\Leftarrow) Suppose $g_T \neq f_T$. Then per monic & $g_T \mid f_T$, $\deg(g_T) \leq \deg(f_T)$. Then per same argument as before: $\dim(\text{span}\{x, Tx, \dots\}) \leq \deg(g_T) < \deg(f_T) = \dim V \Rightarrow V$ is not T-cyclic. \square

Prop. Let T, V as before. Let $f_T(t), g_T(t)$ denote the characteristic & minimal polynomials, respectively.

Then for any prime [irreducible] factor $p(t)$ of $f_T(t)$, we have $p(t) \mid g_T(t)$.

Matrix Polynomials

11/17/23

Lemma 22

Result: Generalized Eigenspaces

Let $T_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$, with eigenvalues $\lambda_1, \dots, \lambda_k$ ($k \leq n$). Then, given λ_i , the generalized eigenspace of λ_i is the set $\{v: (A - \lambda_i I)^l v = 0 \text{ for some } l \geq 1\} = \bigcap_{l=1}^{\infty} \ker(A - \lambda_i I)^l$.

Minimal Polynomials

Lemma. Let p be the minimal polynomial for $T_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$; then the factorization of $p(x)$ is an expression of the form $p(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_k)^{d_k}$, where $\lambda_1, \dots, \lambda_k$ are eigenvalues of A .

(*) Proof: Let v be an eigenvector for eigenvalue λ . We know $p(A)v = 0$; then $p(A)v = p(\lambda)v = 0$. But $v \neq 0$ (by defn.), therefore $p(\lambda) = 0$. \square

Corollary. $p(\lambda) = 0$ if and only if λ is an eigenvalue of A .

(*) Proof: Suppose $\exists \lambda_0$ not an eigenvalue such that $p(\lambda_0) = 0$. Then $(x - \lambda_0) \mid p(x)$; then $p(x) = (x - \lambda_0)Q(x)$ for some $Q(x)$. Claim: $Q(x) = 0$. We know $P(A) = 0$; then $P(A) = (A - \lambda_0 I)Q(A) = 0 \Rightarrow (A - \lambda_0 I)Q(A)v = 0 \forall v \in \mathbb{C}^n$. Then $Q(A)v \in \ker(A - \lambda_0 I)$; but λ_0 not an eigenvalue $\Rightarrow \ker(A - \lambda_0 I) = \{0\} \Rightarrow Q(A)v = 0 \forall v \in \mathbb{C}^n$. Then $Q(A) = 0$, but $\deg Q < \deg P$; then $Q(x)$ can only be 0. \square

Claim. Let $p(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_k)^{d_k}$ be the minimal polynomial. We claim that $1 \leq d_i \leq \dim \ker(\lambda_i)$, $1 \leq i \leq k$, where $d_i = \dim \ker(\lambda_i)$.

(*) Proof: Know $p(\lambda_i) = 0 \forall 1 \leq i \leq k \Rightarrow d_i \geq 1$. If $d_i > d_j$, since $(A - \lambda_i I)^{d_i} = (A - \lambda_j I)^{d_j} \forall d_j > d_i$, therefore we would be able to find a smaller polynomial than $p(x)$ [contradiction]. \square

Corollary: Let $C(x)$ be the characteristic polynomial of A . Then $p(x) \mid C(x)$; then $C(A) = 0$. [C-H]

Complex Inner Products

11/27/23

Lecture 23

(*) Linear Algebra

Linear algebra used in / "touched by" all other branches of math. ↗ used in physics, e.g.

Ex: (Group theory) Groups can be represented as invertible matrices via homomorphism $G \rightarrow M_G$

Ex: Finite groups mapped to permutations [Permutation \Rightarrow invertible matrix].

Complex Inner Products

Recall: Inner product $\langle \cdot, \cdot \rangle$ on \mathbb{R}^n is symmetric, \mathbb{R} -linear

Problem: Cannot simply apply $\langle \cdot, \cdot \rangle$ to \mathbb{C}^n (ex: $\langle v, v \rangle > 0 \Rightarrow \langle iv, iv \rangle < 0$ [not positive-def.])

→ Define inner product on \mathbb{C}^n :

$$\langle (\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n) \rangle = \sum_{i=1}^n \alpha_i \bar{\beta}_i$$

(*) Alt. $\langle \cdot, \cdot \rangle$

Properties:

- Is still positive definite
- Is no longer \mathbb{C} -linear [ex: $\langle v, iw \rangle \neq i\langle v, w \rangle$]
- (*) Is still "sesquilinear", i.e. $\langle v, \lambda_1 w_1 + \lambda_2 w_2 \rangle = \bar{\lambda}_1 \langle v, w_1 \rangle + \bar{\lambda}_2 \langle v, w_2 \rangle$
- Is conjugate-symmetric, i.e. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

$$(*) v = (\alpha_1, \dots, \alpha_n) \rightarrow \langle v, v \rangle = \sum_{i=1}^n |\alpha_i|^2$$

Recall: Adjoints

Let $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$ be a \mathbb{C} -linear transformation. Then the adjoint of T [associated with T] is the linear map $T^*: \mathbb{C}^n \rightarrow \mathbb{C}^n$ determined by $\langle Tv, w \rangle = \langle v, T^*w \rangle$.

(*) Also called a Hermitian adjoint

Hermitian Matrices

11/27/23

Lecture 27

Determining T^*

(cont.)

T^* is uniquely determined by T .

(*) Intuition: Let $\vec{v} \in \mathbb{C}^n$. If $\langle \vec{v}, \vec{w} \rangle$ is known $\forall \vec{w} \in$ basis of $\mathbb{C}^{n*} \Rightarrow \vec{v}$ is known.

\rightarrow Let $T_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$. Then for any $v \in \mathbb{C}^n$ s.t. Tv is known: $\langle v, Te_i \rangle$ known $\Rightarrow \langle T^*v, e_i \rangle$ known.

$\rightarrow \forall e_i, e_i$ [standard basis]: $\langle T^*e_i, e_i \rangle = \langle e_i, Te_i \rangle = \overline{\langle Te_i, e_i \rangle} = [(A^*)_{ii}] = A_{ii}$

$$\rightarrow A^* = (\bar{A})^T$$

Diagonalizability

(ie. Asymmetric)

Recall: Let $A \in \mathbb{R}^{n \times n}$. Then $A = A^T \Rightarrow A$ diagonalizable over \mathbb{R} .

Def: Hermitian Matrices

Let $A \in \mathbb{C}^{n \times n}$. We say A is Hermitian if $A = A^*$. If A is Hermitian (alt: self-adjoint), then A can be diagonalized into a real diagonal matrix.

Corollary: If A is real and $A = A^T [= (\bar{A})^T]$, then A is diagonalizable.

(*) Proof: Let $A = A^* [= (\bar{A})^T]$ be the matrix associated with $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$. Let v be an eigenvector of A , i.e. $T_A v = \lambda v$. Then $\langle \lambda v, w \rangle = \langle T_A v, w \rangle = \overline{\langle w, T_A v \rangle}$.

Know: $\langle v, T_A v \rangle = \langle T^* v, v \rangle$. Recall: $\langle T^* v, v \rangle = \langle v, T v \rangle = \overline{\lambda} \langle v, v \rangle$.

Since $T = T^*$: $\langle T^* v, v \rangle = \overline{\langle T v, v \rangle} = \overline{\lambda} \langle v, v \rangle$. Then $\lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}$. [since $|v| \neq 0$]

(**) To find additional eigenvectors: look at $v^\perp = \{w: \langle v, w \rangle = 0\}$ [T -invariant subsp.]. \square

[roots]

(*) Recall: Every polynomial in \mathbb{C}^n has solutions \rightarrow every matrix in \mathbb{C} has an eigenvalue.

Adjoints

11/28/23

Disc 16

Recall: Adjoints

Let $T: V \rightarrow W$. Then an [the] adjoint for T is a linear map $S: W \rightarrow V$ such that $\langle Tx, y \rangle = \langle x, Sy \rangle$, $\forall x \in V, y \in W$.

Claim: If \exists an adjoint for T , then it is unique.

(*) Proof: Since $\langle \cdot, \cdot \rangle$ is positive-definite, then for any vector $v_0 \in V$, $\langle v_0, v \rangle = 0 \iff v \in V$ if and only if $v = \vec{0}$. Equivalently: $\langle v, v_0 \rangle = \langle v, v \rangle \iff v_0 = v$.

Let S_1, S_2 be adjoints for T ; then for any $x \in V, y \in W$: $\langle x, S_1 y \rangle = \langle x, S_2 y \rangle$.

By fixing x , we find $S_1 y = S_2 y$; we can repeat $\forall y \in W$. \square

Prop. Let $T: V \rightarrow W$ be a linear operator on V a fin.-dim. IPS over \mathbb{R} (IPS/\mathbb{R}). Then if

$B = (e_1, \dots, e_n)$ is an orthonormal basis, then $[T^*]_B = ([T]_B)^T$.

Rmk: This can be generalized to \mathbb{C} [over \mathbb{C}]: $[T^*]_B = ([T]_B)^* = \overline{([T]_B)^T}$.

(*) Proof: Recall: $\forall x \in V, x = \sum_{i=1}^n \langle x, e_i \rangle e_i$. Let $A = [T]_B$; then $T(e_j) = \sum_{i=1}^n A_{ij} e_i$.

We know: $T e_j = \sum_{i=1}^n \langle T e_j, e_i \rangle e_i = \sum_{i=1}^n \langle e_j, T^* e_i \rangle e_i \rightarrow A_{ij} = \overline{\langle T^* e_i, e_j \rangle}$. Let $B = [T^*]_B$;

then $A_{ij} = B_{ji} \Rightarrow A = B^T$. \square

Upper Triangularization

11/20/23

Lecture 28

(*) Derivatives as Transformations

Recall: $\int u dv = uv - \int v du$ [Integration by Parts] $\Rightarrow \int f dx = - \int f' g dx$ (+ boundary term)

Define $\langle f, g \rangle = \int_0^{\infty} f(x)g(x)dx \rightarrow \langle f, \frac{df}{dx} \rangle = \langle -\frac{d^2f}{dx^2}, g \rangle + \text{boundary term} \approx 0$
 $\rightarrow \int_0^{\infty} \frac{d}{dx}(f \cdot g) \cdot g = - \int_0^{\infty} f \cdot \frac{d^2g}{dx^2} = -(\int_0^{\infty} f \cdot \frac{d^2g}{dx^2}) = \int_0^{\infty} f \cdot \frac{d^2g}{dx^2}$

Observe: f_x, f_{x^2} akin to self-adjoint operators

Upper Triangularization [Schur's Theorem]

Recall: $T_A^* = (\bar{T}_A)^T$, i.e. $A^* = (\bar{A})^T$

Claim: Let $T_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$. We claim that the matrix A is upper-triangular for some basis (hence Jordan form).

(*) Proof: We know A is a matrix of the form:

$$\begin{pmatrix} & & \\ & \uparrow & \uparrow \\ & v_1 & \dots & v_n \\ & \downarrow & & \downarrow \end{pmatrix}$$

→ Set v_n to be an eigenvector of T^* ,

s.t. $T^*v_n = \lambda_n v_n$. Assume WLOG that $\|v_n\|=1$, i.e. $\langle v_n, v_n \rangle = 1$

→ Claim. $v_n^\perp = \{v \in \mathbb{C}^n : \langle v, v_n \rangle = 0\}$ is a T -invariant subspace.

(*) Proof: Let $w \in v_n^\perp$. Then $\langle w, v_n \rangle = 0 \Rightarrow \langle w, T^*v_n \rangle = \langle w, \lambda_n v_n \rangle = 0$.

Then $\langle Tw, v_n \rangle = 0$. \square

Then our matrix (over v_1, \dots, v_n) is a matrix of the form:

$$\begin{pmatrix} & & & \\ & T(v_n^\perp) & \uparrow & \\ & \vdots & v_n & \\ & 0 & \downarrow & \end{pmatrix}$$

→ Can find an orthonormal basis (and thus an upper-triangular matrix) via induction on v_n^\perp . \square

Observe: Values on the diagonals are eigenvalues of A . ((*) Proofs only works in \mathbb{R}^n if all eigenvalues real))

(*) Midterm Review Problems

11/30/22

Problem 1: Let $A \in \mathbb{C}^{n \times n}$ be in Jordan form w/ single eigenvalue $\lambda \in \mathbb{C}$. Show that: (a) # blocks

$$= \dim(\ker(A - \lambda I)) \text{ and (b) } \# \text{ blocks of size } \geq 2 = \dim(\ker((A - \lambda I)^2)) - \dim(\ker(A - \lambda I)).$$

(a) Write $B = A - \lambda I$; Then $\exists 1 = i_1 < i_2 < \dots < i_k \leq n$ s.t. $B e_i = \begin{cases} 0 & \text{if } i \in \{i_1, \dots, i_k\}, \\ e_i & \text{otherwise.} \end{cases}$

Observe: # blocks = k . Let $B = \{e_{i_1}, \dots, e_{i_k}\}$; want to show that B is a basis for $\ker B$.

We know that $B \subset \ker B$, and that e_{i_1}, \dots, e_{i_k} are linearly independent.

Claim: $\ker B = \text{span}(B)$

Proof: Let $x \in \ker B$. We know $x = \sum_{i=1}^k c_i e_{i_1}$, and that $Bx = 0$. Since $B e_i = 0 \forall i \in \{i_1, \dots, i_k\}$:

$$Bx = B\left(\sum_{i=1}^k c_i e_{i_1}\right) = \sum_{i=1}^k c_i B(e_{i_1}) = \sum_{i \notin B} c_i B(e_i) \Rightarrow c_i = 0 \forall i \notin B \text{ [since } e_{i_1}, \dots, e_n \text{ independent].} \quad \square$$

(b) Look at B^2 : Blocks size 1 \Rightarrow look the same in B^2 (a)

Blocks size ≥ 2 $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ in $B \rightarrow$ one smaller in B^2 $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Can name Jordan blocks of size ≥ 2 : integers $1 \leq j_1 \leq \dots \leq j_r \leq k$ s.t. $B e_{j_i} \neq 0$ but $B^2 e_{j_i} = 0$.

\rightarrow Then show that $\{e_{i_1}, \dots, e_{i_k}, e_{j_1}, \dots, e_{j_r}\}$ is a basis for $\ker(B^2)$. [Same arg. as previous]

Problem 7: Show that if $T: \mathbb{C} \rightarrow \mathbb{C}$ is linear and self-adjoint [$T^* = T$], then T has only real eigenvalues.

Let $v \in \mathbb{C} \setminus \{0\}$ be an eigenvector of T with eigenvalue λ .

$$\begin{aligned} \rightarrow \lambda \|v\|^2 &= \lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, T^* v \rangle = \langle v, T v \rangle = \overline{\langle T v, v \rangle} \\ &\rightarrow = \overline{\langle \lambda v, v \rangle} = \overline{\lambda \|v\|^2} = \bar{\lambda} \|v\|^2. \end{aligned}$$

Since $\|v\|^2 \neq 0$: $\lambda = \bar{\lambda}$. \square

Disc 17

Proof of Jordan Form

12/6/23

Lecture 29

Recall: Given $T: V \rightarrow V$, can represent V as a direct sum $V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_s}$ of generalized eigenspaces of T . Then, to show T has a Jordan form, suffices to show $T|_{E_{\lambda_i}}$ has Jordan form $\forall i$.

Claim: Let $T: V \rightarrow V$ linear; let E_{λ_i} be a gen. eigenspace of T . Then \exists a matrix representation of $T|_{E_{\lambda_i}}$ in Jordan form.

↳ some λ invariant subspace

(*) Proof: By induction on $\dim E_{\lambda_i}$. Base case: $\dim E_{\lambda_i} = 1 \rightarrow T|_{E_{\lambda_i}} = (\lambda_i)$ [trivial].

Inductive step: Look at $(T - \lambda_i I) E_{\lambda_i}$; we observe that $\dim((T - \lambda_i I)(E_{\lambda_i})) < \dim E_{\lambda_i}$, since $T - \lambda_i I$ sends at least one vector [an eigenvector] to $\vec{0}$. \rightarrow $\text{Image}(T - \lambda_i I)$ on E_{λ_i} .

If we look at $(\ker(T - \lambda_i I)) \cap \text{im}(T - \lambda_i I)$:

Case 1: $\ker(T - \lambda_i I) \cap \text{im}(T - \lambda_i I) = \{\vec{0}\}$. Per Rank-Nullity: $\dim E_{\lambda_i} = \dim(\ker(T - \lambda_i I)) + \dim(\text{im}(T - \lambda_i I))$; then $E_{\lambda_i} = \ker(T + \lambda_i I) \oplus \text{im}(T + \lambda_i I)$. Since $\ker(T + \lambda_i I)$

consists entirely of eigenvectors of T (with eigenvalue λ_i), therefore $T|_{\ker(T + \lambda_i I)}$ is diagonalizable; and we claim $T|_{\text{im}(T + \lambda_i I)}$ has Jordan form by induction.

Case 2: $\ker(T - \lambda_i I) \cap \text{im}(T - \lambda_i I) \neq \{\vec{0}\}$. We know that $T|_{\ker(T - \lambda_i I) \cap \text{im}(T - \lambda_i I)}$ is diagonalizable; then it remains to find a Jordan basis for $T|_{\text{im}(T - \lambda_i I)}$.

We observe that for every vector $\vec{v} \in (\ker(T - \lambda_i I) \cap \text{im}(T - \lambda_i I))$, $\exists \vec{v}' \in \text{im}(T - \lambda_i I) \setminus \ker(T - \lambda_i I)$ such that $(T - \lambda_i I)\vec{v}' = \vec{v}$. Then for every $\vec{v}_0 \in \ker(T - \lambda_i I) \cap \text{im}(T - \lambda_i I)$, we can find a Jordan chain of vectors $\vec{v}_0, (T - \lambda_i I)\vec{v}_0, \dots, (T - \lambda_i I)^s \vec{v}_0 = \vec{v}_0$, $[s < \dim(\text{im}(T - \lambda_i I))]$ in E_{λ_i} . We observe that the span of the vectors in the chain are $(T - \lambda_i I)$ -invariant and a Jordan basis for $(T - \lambda_i I)|_{\text{span}\{\vec{v}_0, (T - \lambda_i I)\vec{v}_0, \dots, (T - \lambda_i I)^s \vec{v}_0\}}$, and eigenvalue 0.

then they are a Jordan basis for $T|_{\text{span}\{\vec{v}_0, (T - \lambda_i I)\vec{v}_0, \dots, (T - \lambda_i I)^s \vec{v}_0\}}$. Per definition of E_{λ_i} , every vector not in $\ker(T - \lambda_i I)$ belongs to exactly one Jordan chain;

then $E_{\lambda_i} = \ker(T - \lambda_i I) \oplus \text{span}\{\vec{v}_0, (T - \lambda_i I)\vec{v}_0, \dots, (T - \lambda_i I)^s \vec{v}_0\} \oplus \dots \oplus \text{span}\{\vec{v}_r, (T - \lambda_i I)^s \vec{v}_r\}$. \square