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

Fields

9/29/23
Lecture 1

Def: Fields

A field is a set F associated with two binary operations - "addition" ($+ : F \times F \to F$) and "multiplication" ($\cdot : F \times F \to F$) 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 1 (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 ($+ : V \times V \to V$) and scalar multiplication ($\cdot : F \times V \to V$) - 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 $\mathcal{F}(S,F)$ [the set of functions $f: S \to F$] 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 $f: \{1,2,\dots,n\} \to \mathbb{R}$ (mapping an "index" in the coordinate to the coordinate's value [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 $f: n \to a_n$) is a vector space over F (where addition: $(s_n)+(t_n) = (s_n+t_n)$).
  - (*) l^2 ("little el 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 $\mathbb{R}$ ($q \in \mathbb{Q} \to \pi q \notin \mathbb{Q}$, 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).

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

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

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

Notes on Fields

10/3/23
Disc. 2

(*) On Fields

Notation: Although we generally accept the meaning of "$a+b+c+\ldots$" and "$a\cdot b\cdot c\cdot\ldots$" to be the repeated addition/multiplication of $a, b, c, \ldots \in F$, 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 $+(a,+(b,+(c,\ldots)))$ and $\cdot(a,\cdot(b,\cdot(c,\ldots)))$ 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 $(a\cdot 0)+b=0$.
Then $0=(a\cdot 0)+b=(a\cdot 0+a\cdot 0)+b = a\cdot 0+0 = a\cdot 0$. □

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

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 "<u>almost-field</u>", 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
Lecture 3 (cont.) + Lecture 4

Thm: "The Big Theorem"
Let (V, F) be a vector space generated by a finite set of m vectors. Then any set $S \subset V$ with $n > m$ elements is linearly dependent.

(*) 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_1 to $w_1, \dots, w_m \to v_1, w_1, \dots, w_m$. 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 $w_1, \dots, w_m$, then $\exists\, w_i$ s.t. w_i is a linear combination of v_1 and $w_{j \neq i}$. Then we can remove w_i from the set while still generating V: we have "replaced" w_i with v_1. [Remove v_1 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 $v_j \notin \{v_1, \dots, v_m\}$; since $v_1, \dots, v_m$ generates V, then this v_j 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 $x_1v_1 + x_2v_2 + \dots + x_nv_n = 0$ (x_i not all 0). We know $\exists$ a spanning set of vectors $w_1, \dots, w_n$ s.t. $v_j = a_{j1}w_1 + \dots + a_{jm}w_n\ \forall v_j$. Substituting: $x_1(a_{11}w_1 + \dots + a_{1m}w_n) + \dots + (a_{n1}w_1 + \dots + a_{nm}w_m) = 0$.
We can rearrange to get: $(a_{11}x_1 + a_{21}x_2 + \dots + a_{n1}x_n)w_1 + \dots + (a_{1m}x_1 + \dots + a_{nm}x_n)w_m = 0$. We know $\alpha_1w_1 + \dots + \alpha_r w_m = 0$ only if $\alpha_i = 0\ \forall i$; then we want to find $x_1, \dots, x_n$ s.t. $a_{1j}x_1 + \dots + a_{nj}x_n = 0\ \forall j \in \{1, \dots, m\}$. This is a homogeneous linear system with m equations and $n > m$ unknowns; thus, we know $\exists$ a nontrivial solution such that $x_1v_1 + \dots + x_nv_n = 0$, 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 $\to$ set $S = S \cup \{v\}$ & repeat.

Def: Given vector space V, a subspace of V is a subset $V' \subseteq V$ 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 $w_1, \dots, w_k$ of V', we can extend the basis by $(n-k)$ additional vectors $v_1, \dots, v_{n-k}$ 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 $f: V \to W$ such that $f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2)$.

We call the set of linear transformations $V \to W$ $\text{Hom}(V, W)$, 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 5

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 $v_1, \dots, v_n$

(*) 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 U, the assignment $\text{Lin}(V, W) \to \text{Fun}(B, V)$ [Equivalently: $f: T \in \text{Hom}(V, W) \mapsto T|_B$] is a bijection.

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

(*) $\boxed{\text{Axiom of Choice: Every vector space has a basis.}}$

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

(*) Analogously: $T: V \to W$ is surjective iff $\exists$ a "right inverse" linear map $S: W \to V$ such that $T \circ S = $ identity 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 [in F] $1 + 1 + \dots + 1 = 0 \in F$ for some number of 1s. The smallest such positive integer is called the characteristic of F and is prime.
- HW1: $\text{Hom}(V, W) = $ set of linear functions $V \to W$; $C([0,1]) = $ set of continuous functions $[0,1] \to \mathbb{R}$. [$F(V, W) = $ func. $V \to W$]

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

(*) Proof: $T(V) = \{\sum_{j=1}^{n} \alpha_j v_j\}$ is the set of linear combinations of vectors $T(\vec{v}) \in W$. Then $T(v_1), \ldots, T(v_n)$ generate $T(V)$. Except $v_1, \ldots, v_k \in \ker(T)$ and therefore $T(v_1) = \ldots = T(v_k) = 0$; thus, $T(v_{k+1}), \ldots, T(v_n)$ generate $T(V)$.

We want to show that $T(v_{k+1}), \ldots, T(v_n)$ are independent. Assume they are not independent: then $\exists \sum_{j=k+1}^{n} \beta_j T(v_j) = 0 \in W$ for some $\beta_{k+1}, \ldots, \beta_n$ not all 0. Then $\sum_{j=k+1}^{n} \beta_j v_j \in \ker(T)$; therefore $\exists$ a linear combination [nontrivial] of $v_1, \ldots, v_k$ equal to $\sum_{j=k+1}^{n} \beta_j v_j$; but this is impossible, since $v_1, \ldots, v_n$ are a basis and thus independent. Therefore $T(v_{k+1}), \ldots, T(v_n)$ are independent and thus a basis for $T(V)$.

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

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

Notes on Transformations

$(T: V \to W)$

- For a linear transformation T, the associated vector spaces $\ker(T)$, $\mathrm{im}(T)$ are actually subspaces (where $\ker(T) \subset V$, $\mathrm{im}(T) \subset W$) [As proof of subspace, we need only check closure & additive identity]
  - (*) Does not necessarily hold for nonlinear transformations (e.g. range$(f: x \mapsto x^2) = \{y : y > 0\}$ is not closed under addition)
- Things (elements, operations, etc.) may not always be the same across vector spaces
  - (*) "$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$" $\longrightarrow$ "$\alpha v_1 + \beta v_2$" 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 F^n: $e_1, \dots, e_n$: $(1,0,\dots,0), (0,1,\dots,0), \dots, (0,\dots,0,1)$

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

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

Matrix Representations of Linear Transformations

Let $T: F^n \to V$ be a linear transformation, and let $e_1, \dots, e_j$ be the standard basis of 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 $1 \le j \le n$, i.e.:

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

Notes on Matrices of Linear Transformations

- The image $T(F^n)$ of T is generated by the columns $T(e_j)$ of the matrix of T

  $\Rightarrow \dim(T(F^n)) = \dim(\text{span}(T(e_1), \dots, T(e_n)))$

  (*) 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 [concrete]; 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
(cont.)

Def: Dual Spaces

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

Def: Dual Bases

Let V be a finite-dimensional vector space over F, and let $v_1, \ldots, v_n$ be a basis for V. Define linear functions $\theta_1, \ldots, \theta_n \in V^*$, where $\theta_i(v_j) = \{1 \text{ if } i=j;\ 0 \text{ if } i \neq j\}$. Then $\theta_1, \ldots, \theta_n$ are a basis for V^*, called the dual basis of $v_1, \ldots, v_n$. $[\theta_j(\sum_{k=1}^{n} \alpha_k v_k) = \alpha_j]$

(*) Notation [Kronecker delta]: $\theta_i(v_j) = \delta_{ij} = \{1 \text{ if } i=j;\ 0 \text{ if } i \neq j\}$

(*) Proof: Dual Bases

We claim $\theta_1, \ldots, \theta_n$ dual to $v_1, \ldots, v_n$ 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 $\gamma_1, \ldots, \gamma_n$ such that $w = \sum_{j=1}^{n} \gamma_j \theta_j$. Set $\gamma_k = w(v_k)$ for each basis vector v_k; then $\sum_{j=1}^{n} \gamma_j \theta_j(v) = w(v)\ \forall\ v \in V$.

(*) 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 $(v_1, v_2, v_3, \ldots, v_z, 0, 0, 0, \ldots)$).

Then there is no finite set $\theta_1, \ldots, \theta_n$ that could express $\Omega: \Sigma(v_1, v_2, \ldots)$ 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 $v_1, \dots, v_k$ $(k \le n)$ for W [$\dim W = k$]; the 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 $\omega = \sum_{j=1}^{n} a_j \theta_j \in \text{ann}(W)$. Then $\omega(w) = 0 \ \forall w \in W$, therefore $\omega(v_j) = 0$ $\forall 1 \le j \le k$. Then $\omega \in \text{ann}(W)$ iff $a_1 = a_2 = \dots = a_k = 0$, 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 $T: V \to W$ for vector spaces V, W over F. Then we define the adjoint of T to be the linear transformation $T^*: W^* \to V^*$ 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 F is a function $\langle , \rangle : V \times V \to F$ [mapping ordered pairs $(v_1, v_2) \in V^2$ to F] that satisfies:

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

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

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

Inner Products in Geometry

Let $\langle , \rangle : \langle (a_1, \dots, a_n), (b_1, \dots, b_n) \rangle = a_1 b_1 + \dots + a_n b_n$ 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: $\|v\| = \langle v, v \rangle^{1/2}$ and the distance $(v_1, v_2) = \|v_1 - v_2\|$. [Def: Distance/Norm]

Def: If $v_1, v_2 \neq 0$, we define $\cos(\theta) = \frac{\langle v_1, v_2 \rangle}{\|v_1\| \cdot \|v_2\|}$ 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: $\langle v - \text{proj}_w(v), v - \text{proj}_w(v) \rangle = \langle v, v \rangle - 2 \frac{\langle v, w \rangle}{\|w\|^2} \langle v, w \rangle + \frac{\langle v, w \rangle^2}{\|w\|^4} \langle w, w \rangle$

$\to = \langle v, v \rangle - 2 \frac{\langle v, w \rangle^2}{\|w\|^2} + \frac{\langle v, w \rangle^2}{\|w\|^2} = \langle v, v \rangle - \langle v, \frac{w}{\|w\|^2} \rangle \geq 0 \Rightarrow \|v - \text{proj}_w(v)\| \geq 0$. $\blacksquare$

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

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: $x, y \in \ell^2 \to \langle x, y \rangle = \sum |x_j \cdot y_j| \leq \sum (|x_j|^2 + |y_j|^2) < +\infty$

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

Orthonormalization

10/18/23
Lecture 9
+ Lecture 10

Orthonormalization

Let $(V, \mathbb{R})$ be a vector space with inner product $\langle , \rangle$. 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 $\|v_j\| = 1 \ \forall 1 \leq j \leq n$,
$\langle v_i, v_j \rangle = 0 \ \forall i \neq j$. [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) $\langle w_i, w_j \rangle = 0 \ \forall i \neq j$.

(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 $w_1, \dots, w_k$. We set
$w_{k+1} = v_{k+1} - \sum_{j=1}^{k} \text{proj}_{w_j}(v_{k+1}) = v_{k+1} - \sum_{j=1}^{k} \langle w_j, v_{k+1} \rangle w_j$, then divide by its length.

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

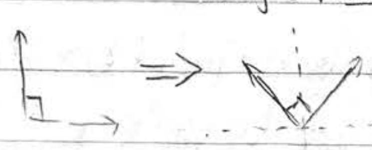
(*) Proof: Let $i < j$. Then: $\langle w_i, w_j \rangle = \langle w_i, v_j - \sum_{k=1}^{j} \frac{\langle w_k, v_j \rangle}{\|v_j\|} w_k \rangle = \langle w_i, v_j \rangle - \langle w_i, \langle w_i, v_j \rangle w_i \rangle$
$\rightarrow = \langle w_i, v_j \rangle - \langle w_i, v_j \rangle = 0$. ∎

(*) 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 analagous to rotating axes in $\mathbb{R}^n$:

Ex $(\mathbb{R}^2)$:  $\{(1,0), (0,1)\} \Rightarrow \{(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})\}$

Orthogonal Transformations

10/20/23

Lecture 10 (cont.)

Orthogonal Transformations

Def: Let $(V, \mathbb{R})$ be a vector space with inner product $\langle , \rangle$. Then a linear transformation $T: V \to V$ 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: "$T: V \to V$ is orthogonal" $\Leftrightarrow$ "$\|T(v)\| = \|v\| \; \forall \, v \in V$". [i.e. T preserves length].

(*) Proof:

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

$B \to A$: Know: $\|v_1 + v_2\| = \|T(v_1 + v_2)\| \Rightarrow \|v_1 + v_2\| = \|T(v_1) + T(v_2)\|$

$\to \|v_1 + v_2\|^2 = \langle v_1 + v_2, v_1 + v_2 \rangle = \|Tv_1 + Tv_2\|^2 = \langle Tv_1 + Tv_2, Tv_1 + Tv_2 \rangle$

$\langle Tv_1 + Tv_2, Tv_1 + Tv_2 \rangle = \langle Tv_1, Tv_1 \rangle + 2\langle Tv_1, Tv_2 \rangle + \langle Tv_2, Tv_2 \rangle$

$= \|Tv_1\|^2 + 2\langle Tv_1, Tv_2 \rangle + \|Tv_2\|^2$

$= \|v_1\|^2 + 2\langle Tv_1, Tv_2 \rangle + \|v_2\|^2$

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

$\to \langle v_1, v_2 \rangle = \langle Tv_1, Tv_2 \rangle$. ∎

(*) 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 $\langle , \rangle$. Then the set of orthogonal transformations $T: V \to V$ 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 [infinite] sequence of prime integers: $p_1=2, p_2=3, p_3=5, \ldots$. Recall (Fundamental Thm): any $m\in\mathbb{N}>1$ has a unique prime factorization.

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

Proof: Suppose $\exists\, a_1,\ldots,a_n \in \mathbb{Q}$ such that $\sum_{j=1}^{n} a_j \log(p_j) = 0$. We want to show $a_1 = \ldots = a_n = 0$.

Recall log properties: $\sum_{j=1}^{n} a_j \log(p_j) = 0 \Rightarrow 0 = \log\left(\prod_{j=1}^{n} p_j^{a_j}\right)$.

Drop the log: $1 = \prod_{j=1}^{n} p_j^{a_j}$. Observe: if $a_1,\ldots,a_n$ all non-negative, then we could conclude that $a_1,\ldots,a_n$ all 0.

Assume $\exists\, a_j < 0$. Then let $I_{(-)} = \{i : a_i < 0\}$, $I_{(+)} = \{i : a_i \geq 0\}$. Multiply both sides to cancel the negative exponents: $\prod_{j\in I_{(-)}} p_j^{|a_j|} = \prod_{i\in I_{(+)}} p_i^{a_i}$, where $p_i \neq p_j$ for any $i\in I_{(+)}, j \in I_{(-)}$.

Let $x = \prod_{j\in I_{(-)}} p_j^{a_j} = \prod_{i\in I_{(+)}} p_i^{a_i}$. We know $p_k > 0\ \forall k\in\mathbb{N}$; then $x>0$. Assume WLOG that $a_1,\ldots,a_n$ 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 $a_1 = \ldots = a_n = 0$. ∎

(*) Prop: Dual Bases

Let (V,F) be a vector space with basis $v_1,\ldots,v_n$ and dual basis $f_1,\ldots,f_n$. 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 $B: V\times V \to F$ 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 $V\to V^*$, i.e.: let $\varphi: \mathrm{Bil}(V) \to \mathrm{Hom}(V,V^*)$, such that for $B\in \mathrm{Bil}(V)$, $\varphi(B) = \varphi_B : V\to V^*$, where $\varphi_B(v\in V) = B(v,\cdot)$ $[(\varphi_B(v))(u) = B(v,u)]$.

Eigenvectors & Diagonalization

10/23/23

Lecture 11

Diagonalization

Let V be a vector space with basis $v_1, \ldots, v_n$. Then for any linear transformation $T: V \to W$, where $\dim W = n$ and $\dim(\operatorname{im}(T)) = k$, we can find a basis $w_1, \ldots, w_n$ 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 $T: V \to V$, where we can only pick a single basis for both sides.

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

Similarity

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

Diagonalization and Similarity

Prop: Given a matrix A and associated transformation $T_A : \mathbb{R}^n \to \mathbb{R}^n$, T_A is similar to some T_{diagonal} if, for some $e_1, \ldots, e_n$, $T_A(e_j) = \alpha_j e_j$ (for some $\alpha_j \in \mathbb{R}$) $\forall j$.

Def: Eigenvectors

Let $T: V \to V$. 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, $\vec{v} \in \ker(T - \lambda I)$.

=> Diagonalization

A linear transformation $T: V \to V$ 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.)

10/24/23

Disc 8

Recall: Isomorphisms

Def: A linear map $T: V \to W$ is called an isomorphism if $\exists\, L: W \to V$ such that $TL = 1_W$ and $LT = 1_V$. V and W are called isomorphic if there exists an isomorphism $V \to W$.

Recall: Bilinear Forms

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

Remark: φ_B is an isomorphism iff φ_B is injective (since $\dim V = \dim V^*$), i.e. iff $\ker(\varphi_B) = \{\vec{0}\}$. A vector $\vec{x} \in \ker(\varphi_B)$ iff $\varphi_B(x)$ is the zero function, i.e. $B(x,y) = 0\ \forall\, y \in V$.

Degeneracy

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

Nondegenerate Forms

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

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

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

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

$\Rightarrow$ Corollary: If $(V, \langle , \rangle)$ is an inner product space with orthonormal basis $v_1, \dots, v_n$, then $x = \sum_{i=1}^{n} \langle x_i, x \rangle x_i\ \forall\, x \in V$.

Self-Adjoint Transformations

10/25/23

Lecture 12 Symmetric Matrices

The 2x2 symmetric matrix $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ has eigenvalues $\lambda = \frac{a+d}{2} \pm \sqrt{(a-d)^2+4b^2}$ [b=0 => diagonal]
→ has two distinct real solutions λ_1, λ_2, if $b \neq 0$

→ $v_1 = \left(1, -\frac{a-\lambda_1}{b}\right)$; $v_2 = \left(1, -\frac{b}{d-\lambda_2}\right)$. Observation: $v_1 \perp v_2$ => 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 $\langle A\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \rangle = \langle \begin{pmatrix} ax_1 + by_1 \\ bx_1 + dy_2 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \rangle$
→ $= ax_1x_2 + b(y_1x_2 + x_1y_2) + d(y_1y_2)$. This result is the same no matter which side of the equation A is on; i.e. $\langle Av, w \rangle = \langle v, Aw \rangle \ \forall v, w$.

→ 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 $T: V \to V$ 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 => relative to all]

Eigenvectors of Self-Adjoint Transformations

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

(*) Proof: $\langle \vec{w}, v \rangle = 0 \Rightarrow \langle Tw, v \rangle = \langle w, Tv \rangle = \langle w, \lambda v \rangle = \lambda \langle w, v \rangle = 0 \to Tw \in v^\perp$. □

Restricting T to $v^\perp$, then the transformation $T: v^\perp \to v^\perp$ is self-adjoint with respect to $v^\perp$. Then we can find an eigenvector for $T_{v^\perp}$ 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 ∃ 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 $im(M) = im(MM^T)$. [im: image]

Disc 9

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

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

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

Let $\vec{x} \in ker(MM^T)$. Then $\vec{0} = \vec{x}^T \cdot \vec{0} = \vec{x}^T(MM^T\vec{x}) = (\vec{x}^T M)(M^T\vec{x})$

$\to (M^T\vec{x})^T(M^T\vec{x}) = 0 \Rightarrow \|M^T\vec{x}\| = 0 \Rightarrow \vec{x} \in ker(M^T)$. $\square$

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

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

This is a linear combination of $e^t, \ldots, e^{nt}$ equaling 0; then per the inductive hypothesis, $a_0, \ldots, a_n = 0$. Then $a_{n+1} = 0$. $\square$

(*) Alternatively: Can show that $e^t, \ldots, e^{nt}$ 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 $(V, \langle , \rangle)$ be a finite-dimensional inner product space. If a transformation $T: V \to V$ 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.

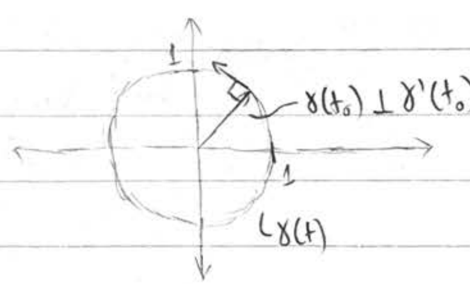
(*) 2-Dimensional Case

We know the set $\{\vec{x} : \|\vec{x}\| = 1\}$ 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 $f(\vec{x}) = \langle A\binom{x}{y}, \binom{x}{y} \rangle = ax^2 + 2bxy + cy^2 \;\forall\, \vec{x} = \binom{x}{y} \in \gamma(t)$. 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
(cont.)

(*) <u>2-Dimensional Case (cont.)</u>

(*) Notation: Let $\gamma(t_0) = x_0 = \begin{pmatrix} x \\ y \end{pmatrix}$.

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

$\Rightarrow \nabla f(\gamma(t)) = \langle \frac{\partial}{\partial x}(ax^2+2bxy+cy^2), \frac{\partial}{\partial y}(ax^2+2bxy+cy^2) \rangle$

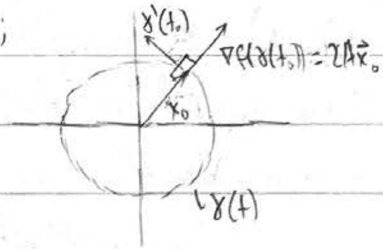
$= 2(ax+by, bx+cy)$. Then $2\langle (ax+by, bx+cy), \gamma'(t) \rangle \big|_{t_0} = 0$; but

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

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

<u>$A\vec{x}_0 \parallel \vec{x}_0$, thus $A\vec{x}_0 = \lambda \vec{x}_0$ [for some λ]</u>

$\Rightarrow$ <u>$\vec{x}_0$ is an eigenvector of A</u>.



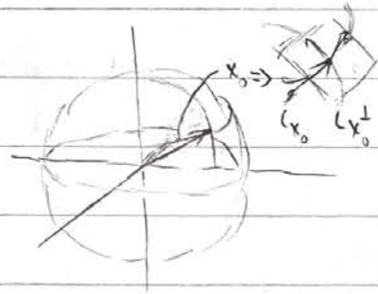
[This process is similar to <u>Lagrange multipliers</u>, or finding the axes of rotated ellipses].

(*) <u>n-Dimensional Case</u>

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

Let $\gamma(t_0) = x_0$ be the point on $\gamma(t)$ where $\langle A\vec{x}, \vec{x} \rangle$ is maximized.

Then the set of vectors orthogonal to x_0, $x_0^\perp$, is a subspace of dimension $(n-1)$. We know $\frac{\partial}{\partial t} \langle A\gamma(t), \gamma(t) \rangle \big|_{t=t_0} = 0$



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

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

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

Since $\vec{x}_0 \perp \gamma'(t_0)$ $\forall$ γ also $\Rightarrow$ <u>$A\vec{x}_0 = \lambda \vec{x}_0$</u>.

Multilinear Forms

10/30/23

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 $f: V^n \to F$ that are skew-symmetric/alternating (i.e. $f_1(v_1, \dots, v_i, v_j, \dots, v_n) = -f(v_1, \dots, v_j, v_i, \dots, v_n)$).

Ex: $w(e_i, e_j)$ on $\mathbb{R}^2$. Know $w(e_i, e_i) = -w(e_i, e_i) \Rightarrow w(e_i, e_i) = 0$. 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 $w(e_1, e_2) \to \det\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = w(v_1, v_2)$].

→ Set of 2-multilinear alternating forms $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ 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] $\{1, \dots, n\} \to \{\sigma(1), \dots, \sigma(n)\}$

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

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

Look at permutation: $\{1, 2, \dots, n\} \mapsto \{\sigma(1), \dots, \sigma(n)\}$. 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 $(-1)^{\#\text{inversions}} = \text{sgn}(\sigma)$.

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

Matrices of Linear Maps

10/31/23

Disc 10

Matrix Representations of Linear Maps

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

→ Define the $m \times n$ matrix $[T]_\beta^\gamma = (a_{ij})$, where $T(x_j) = \sum_{i=1}^n a_{ij} y_i$.

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

(*) Proof: Begin by choosing a basis $(y_1, \dots, y_r)$ for $\mathrm{im}(T) \subset W$. 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 $(x_{r+1}, \dots, x_n) \subset \ker(T)$. Then $\beta = (x_1, \dots, x_n)$ is a basis for V & $[T]_\beta^\gamma$ as desired. □

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

(*) Proof:

($\Rightarrow$) Suppose $\exists \beta$ with $[T]_\beta$ as before. Let $\beta = (x_1, \dots, x_n)$. Then $T(x_j) = \sum_{i=1}^n a_{ij} x_i$ $= a_{jj} x_j$, $a_{jj} \in \{0, 1\}$. We claim that $[T^2]_\beta = [T]_\beta$. $T^2(x_j) = T(T(x_j)) =$ $= T(a_{jj} x_j) = a_{jj} T(x_j) = a_{jj}^2 x_j = T(x_j)$. □

($\Leftarrow$) Suppose $T^2 = T$. Choose a basis $(x_1, \dots, x_r)$ for $\mathrm{im}(T)$ and $(x_{r+1}, \dots, x_n)$ for $\ker(T)$. We claim $\beta = (x_1, \dots, x_n)$ is a basis for V w/ $[T]_\beta$ 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 $T: V \to V$. 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 $\omega : V^n \to \mathbb{R}$ [$\dim V = n$] 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 $w_1, \dots, w_n \in V$: $\omega(w_1, \dots, w_n) = 0 \iff w_1, \dots, w_n$ dependent.

(*) Proof

($\Leftarrow$) follows from multilinearity (dependent $\Rightarrow$ can take 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 WOLOG that $V = \mathbb{R}^n$. We know we can represent each $w_i \in V$ relative to $v_1, \dots, v_n$: $w_i = \sum_{j=1}^{n} a_{ij} v_j$.

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

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

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:

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

$$[= {}^{*}\text{nonzero factors}^{*} \cdot \omega(v_1, \dots, v_n) = {}^{*}\text{nonzero factors}]$$

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

Disc 11

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

(*) 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: WTS $\exists$ an eigenvector.
Recall: square matrix B 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 $\lim_{\lambda \to +\infty} \det(A - \lambda I) = -\infty$; $\lim_{\lambda \to -\infty} \det(A - \lambda I) = +\infty$. 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 $A \in F^{3\times3}$ over F such that $A^3 \neq 0$, but $A^4 = 0$? [A: No]

(*) Proof: Assume, for the sake of contradiction, that $A \in F^{3\times3}$ 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 $0 = A^3 \cdot 0 \Rightarrow A^3\left(\sum_{i=0}^{3} a_i A^i x\right) = 0$; 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_0, 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 $T:\mathbb{R}^n \to \mathbb{R}^n$ with corresponding matrix $A \in \mathbb{R}^{n \times n}$, is it possible to find a basis that makes A diagonal? $\llcorner\left[T(a_1, \dots, a_n) = A\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right]$.

=> Let $T^{\mathbb{C}}: \mathbb{C}^n \to \mathbb{C}^n$ be the linear transformation defined by $T^{\mathbb{C}}(\alpha_1, \dots, \alpha_n) = A\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$.

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

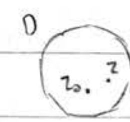
→ We can find an eigenvalue for $T^{\mathbb{C}}$ if we can find a root $\lambda \in \mathbb{C}$ for $\det(A - \lambda I) = 0$.

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 $n > 0$.

→ 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 (i.e. $\exists\, z_0 \in D$ s.t. $|p(z_0)| \le |p(z)|\ \forall z \in D$).



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

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

→ $= p(z_0) + a_k(z - z_0)^k + \dots + a_n(z - z_0)^n$. For z near z_0, higher-order terms drop out → $p(z) \simeq p(z_0) + a_k(z - z_0)^k$, $(z - z_0) \neq 0$. Namely: $\left|\frac{p(z) - (p(z_0) + a_k(z - z_0)^k)}{(z - z_0)^{k+1}}\right| \le M$ for some $M > 0$ in some neighborhood of z_0. Rearranging terms:

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

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

→ $|p(z)| \le |p(z_0) + a_k(z - z_0)^k| + M(z - z_0)^{k+1} < |p(z_0)|$ $\square$

(*) Misc. Notes

<u>Claim</u> Matrix $M(T^*)$ of $T^*: W^* \to V^*$ is $\underline{M(T^*) = (M(T))^T}$, $T: V \to W$.

<u>Proof</u>: Let $e_1, \ldots, e_n$ standard basis for V, $\Theta_1, \ldots, \Theta_m$ dual basis for standard basis on W.

$\to (M(T))_{ij} = \Theta_i(T(e_j)) = \Theta_i(\varphi_j(e_j))$ [$\varphi_1, \ldots, \varphi_n$ dual basis for $e_1, \ldots, e_n$ on V]

Observe: $\Theta_i(T(e_j)) = (T^*(\Theta_i))(e_j) = \varphi_j(T^*(\Theta_i)) = (M(T^*))_{ji}$. $(M(T))_{ij} = (M(T^*))_{ji}$. $\square$

<u>Claim</u>. <u>Row rank = column rank</u>.

<u>Proof</u>: Know: column rank of $T = \dim(\operatorname{im}(T))$. [$T: V \to W$, $\dim V = n$, $\dim W = m$]

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

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

$\to \dim(\operatorname{im}(T^*)) = \dim W^* - \dim \ker T^*$

$= \dim W^* - \dim(\operatorname{ann}(T(V)))$

$= \dim W^* - (\dim W - \dim(\operatorname{im}(T))) = \dim(\operatorname{im}(T))$. $\square$

(preview: Jordan form

<u>Claim</u>: Any 2×2 matrix in $\mathbb{C}$ is either diagonalizable or can be placed into the form $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

<u>Proof</u>: <u>Case 1</u>: Matrix A has 2 distinct eigenvalues $\to$ diagonalizable.

<u>Case 2</u>: Matrix A has only 1 eigenvalue.

<u>2(a)</u>: A has 2 independent eigenvectors $\to$ diagonalizable.

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

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

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

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

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

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

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

Diagonalization (cont.)

11/8/23

Lecture 17

> <u>Lemma</u>: Let $\vec{v}_1, \dots, \vec{v}_k$ be eigenvectors associated with <u>distinct eigenvalues $\lambda_1, \dots, \lambda_k$</u>. Then <u>$\vec{v}_1, \dots, \vec{v}_k$ are linearly independent.</u>

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

(*) <u>Proof</u>: 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 $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$ $[k \leq n]$. We know: $T(\alpha_1 v_1 + \dots + \alpha_k v_k) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_k \lambda_k v_k = 0$. Also know: $(\lambda_1)(\alpha_1 v_1 + \dots + \alpha_k v_k) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_k \lambda_1 v_k$. Subtract $(T - \lambda_1 I)(\alpha_1 v_1 + \dots + \alpha_k v_k)$.
$\alpha_2(\lambda_2 - \lambda_1)v_1 + \dots + \alpha_k(\lambda_k - \lambda_1)v_k = 0$. Then $\alpha_1 v_1 + \dots + \alpha_k v_k$ was not the shortest relation. □

(*) Note: we know $\lambda_i - \lambda_1 \neq 0 \; \forall i \neq i$ because we assumed λs distinct

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

(*) <u>Diagonalization</u>

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

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

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

→ Applied cases: Every matrix B, 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: $\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda \end{pmatrix}$ [λ some eigenvalue]

→ For any $T^{\mathbb{C}}$, can find a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}$ in Jordan form:

$$[T^{\mathbb{C}}]_{\mathcal{B}} = \begin{pmatrix} \boxed{\lambda_1} & 0 & 0 & 0 \\ 0 & \boxed{\lambda_2} & 0 & 0 \\ 0 & 0 & \boxed{\lambda_3} & 0 \\ 0 & 0 & 0 & \boxed{\lambda_4} \end{pmatrix}$$

Jordan blocks for eigenvalues $\lambda_1, \dots, \lambda_4$

(*) Ex:

$$\begin{array}{c} \lambda = 2 \\ \begin{pmatrix} \boxed{2} & 0 & 0 \\ 0 & \boxed{\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{matrix}} & 0 \\ 0 & 0 & \boxed{\begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix}} \end{pmatrix} \\ \lambda = 3 \end{array}$$

[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 $T: V \to V$, 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 $\ker(A - \lambda_i I)$, $\ker(A - \lambda_i I)$, ... = generalized eigenspace for λ_i

(*) Eigenspace Examples

11/9/23

Disc 12

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

(a) Observe: $A^2 - I = (A+I)(A-I) = 0 \to$ one of $ker(A+I)$, $ker(A-I)$ nontrivial. Per Rank-Nullity: $n = rk(A-I) + N(A-I) \leq N(A+I) + N(A-I)$. Since $I \neq -I$ [$char(F) \neq 2$], then these are distinct eigenspaces w/ eigenbases β_1, β_2 s.t. $\beta_1 \cup \beta_2$ are independent and span F^n; then $\exists$ basis of eigenvectors of $A \to A$ diagonalizable. □

(b) Consider $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$: then $A^2 = I$, but Jordan block size $> 1 \Rightarrow$ not diagonalizable.

(U4A Basic

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

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

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

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

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

Then $B(E_\lambda) \subset E_\lambda \neq \emptyset$. Then $B|_{E_\lambda}$ [$B : E_\lambda \to E_\lambda$] 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. □

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. $\dim E_\lambda$'s)

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

$\to null(CD) = null(D) = null(DC)$.

Compare $null(AB - \lambda I)$, $null(BA - \lambda I)$. $null(BA - \lambda I) = null(A(BA - \lambda I))$

$= null(A(BA - \lambda I)A^{-1}) = null(AB - \lambda I)$ 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 F form a ring of expressions $p(x) = a_m x^m + \dots + a_0$.

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

$$\boxed{p(x) = a_m x^m + \dots + a_0 \longmapsto p(T) = a_m T^m + \dots + a_1 T + a_0 I_n}$$

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 $T: V \to V$. 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 $v_1, Tv_1, \dots, T^n v_1$. Since there are $n+1$ vectors in the set (and $\dim V = n$), therefore $\exists\, a_0, \dots, a_n$ not all 0 such that $a_0 v_1 + \dots + a_n T^n v_1 = 0$. Set the polynomial $p_1(x) = a_0 + a_1 x + \dots + a_n x^n$; then $p_1(T)v_1 = \vec{0}$ [but $p_1(x) \neq 0$].

We can find such a polynomial p_i for every vector v_i, $1 \le i \le n$. Define $p_\alpha(x) = p_1(x)p_2(x)\dots p_n(x)$ to be the product of the polynomials $p_1, \dots, p_n$. Then $p_\alpha(x) \neq 0$, but for any v_i, $p_\alpha(A)v_i = (p_1(A)\dots p_n(A))(p_i(A)v_i) = 0$. □

(*) Note: $\deg(p_\alpha) = n^2$; 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 $\deg Q = 0 \Rightarrow Q$ is a scalar multiple of P. $\square$

(Proof continues until p.42)

(*) Diagonalization Examples

11/14/23

Disc 13

<u>Problem: Suppose $T: V \to V$ is a diagonalizable linear operator on fin. dim. V.S. V over F. If $W \subset V$ is a nonzero T-invariant subspace, show that $T|_W$ is diagonalizable.</u>

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 $x_1, \ldots, x_r \in V$ be T-eigenvectors with distinct eigenvalues. Then if $\sum_{i=1}^{r} x_i \in W$, then $x_1, \ldots, x_r \in W$.

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

Through induction, we can keep reducing to find $x_2, \ldots, x_r \in W$. Then $(\lambda_2 - \lambda_1)x_2, \ldots, (\lambda_r - \lambda_1)x_r \in W \Rightarrow x_1 \in W$. □

Choose a basis $y_1, \ldots, y_m$ for W; since T is diagonalizable on V, then for each y_i $(1 \le i \le m)$, we can find eigenvectors $x_{i_1}, \ldots, x_{i_{r_i}}$ w/ distinct eigenvalues s.t. $y_i = \sum_{j=1}^{r} x_{i_j}$; then $x_{i_j} \in W$ $\forall i, j$. Then $W = \text{span}\{y_1, \ldots, y_m\} = \text{span}\{x_{1_1}, \ldots, x_{m_{r_m}}\}$, i.e. W is spanned by eigenvectors of T. □

<u>Problem: Let $T, S: V \to V$ be diagonalizable lin. ops. such that $T \circ S = S \circ T$. Show that S, T are simultaneously diagonalizable, i.e. $\exists$ basis β for V s.t. $[S]_\beta$, $[T]_\beta$ are both diagonal.</u>

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

<u>Observe</u>: 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 $\lambda_1, \ldots, \lambda_r$ 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 $\forall\ 1 \le i \le r$. For each $1 \le i \le r$, we can choose a basis $\{x_{i_1}, \ldots, x_{i_{s_i}}\}$ for E_{λ_i} of eigenvectors for $T|_{E_{\lambda_i}}$. Then $\beta = (x_{1_1}, \ldots, x_{r_{s_r}})$ is a basis for V. Since each vector in the basis is an eigenvector of S, then β is an eigenbasis for S; since each vector is an eigenvector for $T|_{E_{\lambda_i}}$ for some i, then β is an eigenbasis for T. □

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 $\exp(A)$:

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

Intuition: e^A eventually converges => 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: $\|A\| \geq |a_{i,j}|$ $\forall a_{ij}$ in A ($\|A\|$ is a "bound" on the elements in A)

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

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

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

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

(*) Proof: We know $\|A^n\| \leq \|A\|^n$ $\forall n \in \mathbb{N}$ from the proposition. Additionally, we know $(A^n)_{ij} \leq \|A^n\|$ $\forall n \in \mathbb{N}$, any element i,j. Then we can show each element converges:

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

(*) 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. $e^A = B$] only if e^A is invertible, and B is a matrix of the form $B = I + C$ [$\|C\|$ small; $\|C\| < 1$, 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

Disc 14

<u>Problem</u>: Suppose $T, S : V \to V$ 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.

<u>Recall</u>: T, S commute $\Rightarrow$ preserve eigenspaces

<u>Observe</u>: $\dim E_{\lambda_i} = 1 \ \forall \lambda_1, \dots, \lambda_n$ $[\dim \bigoplus_{i=1}^{n} E_{\lambda_i} = \sum_{i=1}^{n} \dim E_{\lambda_i} = n]$

<u>Claim</u>: T-eigenvectors span 1D S-dimensional subspaces.

(*) <u>Proof</u>: Pick T-eigenbasis $\{v_1, \dots, v_n\}$ for V s.t. $T(v_i) = \lambda_i v_i$. We know $E_{\lambda_i} = \operatorname{span}\{v_i\}$ and $S(E_{\lambda_i}) \subset E_{\lambda_i} = \operatorname{span}\{v_i\} \to S(v_i) \in \operatorname{span}\{v_i\} \to v_i$ is an eigenvector for S. $\square$

<u>Jordan Canonical Form</u>

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

<u>Recall</u>: Given composable linear maps $U \xrightarrow{T} V \xrightarrow{S} W$ and finite bases α, β, γ for U, V, W (respectively) $\to [S \circ T]_\alpha^\gamma = [S]_\beta^\gamma \cdot [T]_\alpha^\beta$.

(*) <u>Recall</u>: For any basis β of fin. dim. V, $[Id_V]_\beta = I$.

$\to$ <u>Lemma</u>: Let β, γ bases for fin. dim. V; then $[Id_V]_\gamma^\beta = ([Id_V]_\beta^\gamma)^{-1}$. [Change of coordinates]

<u>Problem</u>: 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 $\beta_i = \{x_{i_1}, \dots, x_{i_s}\}$ and $\exists \lambda_i \in \mathbb{C}$ such that:

$$Ax_{ij} \begin{cases} \lambda_i x_i, & j = 1 \\ x_{i,j-1} + \lambda_i x_{ij}, & j > 1. \end{cases}$$

Let $Q = [Id_A]_S^\beta$, where S is the standard basis. (β_i, reversed)

$A = [L_A]_S^S = [Id]_\beta^S [A]_\beta^\beta [Id]_S^\beta = Q [L_A]_\beta^\beta Q^{-1}$. Let $\gamma = \gamma_1 \cup \dots \cup \gamma_r$, $\gamma_i = (x_{i s_i}, \dots, x_{i,1})$.

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

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

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 $D = A_1P_1 + A_2P_2$ for some A_1, A_2 polynomials (Notation: $P \mid Q \Rightarrow Q = A \cdot P$)

(*) Euclid's Algorithm for Polynomials

Let P_1, P_2 polynomials; $\deg P_1 \geq \deg P_2$. ($\deg P_2 \leq \deg P_1$)

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

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

Corollary: $\gcd(P_1, P_2) = 1$ (no common factors) ("relatively prime") $\Rightarrow 1 = A_1P_1 + A_2P_2$

(*) 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 $v \in \mathbb{R}^n$: $v = A_1(T)P_1(T)v + A_2(T)P_2(T)v$.

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

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.: $\deg P = n \Rightarrow P(x) = a(x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_n)$ $[a \in \mathbb{C} \setminus \{0\}]$ 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. $p \mid ab \Rightarrow p \mid a$ or $p \mid b$.

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

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

(Base case). We know $p_1 \mid P = q_1 q_2 \ldots q_k$. Then $p_1 = q_j$ for some j (we know that p_1 is irreducible).

(Induc. step). Assume we have found $p_1 = q_j$ for some j. Then $\prod_{i=2}^{k} p_i = \prod_{\substack{i=1 \\ i \neq j}}^{k} q_i$. Take $P' = \prod_{i=2}^{k} p_i$; then we can find (same argument as prev.) that $p_i' = q_j'$ for some j.

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

Matrix Decomposition

11/20/23

Lecture 21

(cont.)

Matrix Decomposition

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

→ $p(x)$ can be expressed as the product of irreducibles: $p(x) = p_1(x) \cdot \ldots \cdot p_k(x)$.

Assume $F = \mathbb{C}$; then all factors are linear: $p(x) = (x-\lambda_1)^{k_1} \cdot \ldots \cdot (x-\lambda_\ell)^{k_\ell}$. ~ in $\mathbb{C}$: $(x-\lambda)$ a factor $\Rightarrow (x-\bar{\lambda})$ a factor

Observe: The zeroes 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: $\mathbb{C} = G_A(\lambda_1) \oplus \ldots \oplus G_A(\lambda_k)$, i.e. the generalized eigenspaces of A span $\mathbb{C}$.
>
> (*) Consequence: Let B_i be a basis for each G.E. $G_A(\lambda_i)$. Then $B = B_1 \cup \ldots B_k$ 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 $p(x) = (x-\lambda_1)^{k_1} \cdot \ldots \cdot (x-\lambda_\ell)^{\lambda_\ell}$ be the factorization of $p(x)$. Define $P_1 = (x-\lambda_1)^{k_1} \cdot \ldots \cdot (x-\lambda_{\ell-1})^{\lambda_{\ell-1}}$, $P_2 = (x-\lambda_\ell)^{k_\ell}$. We note that P_1, P_2 are relatively prime, i.e. $\gcd(P_1, P_2) = 1$.

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

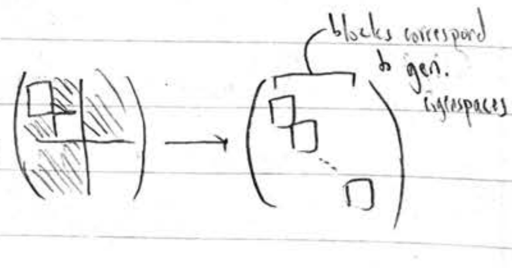
Lemma: $\ker P_1(T), \ker P_2(T)$ are T-invariant.

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

Then $P_1(A) \cdot Av = P_1(A) \lambda_i v = \lambda_i (P_1(A) v)) = 0$. □

Let $v_1, \ldots, v_m$ be a basis for $\ker P_1(T)$, and let $v_{m+1}, \ldots, v_n$ be a basis for $\ker P_2(T)$. Then $B = v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ is a basis for $\mathbb{C}^n$ such that:

$$[T_A]_B = \left(\begin{array}{c|c} A(\ker P_1(T)) & 0 \\ \hline 0 & A(\ker P_2(T)) \end{array} \right) \xrightarrow{\text{decompose } P_1}$$



40

(*) Cyclic Subspaces

11/21/23
Disc 15

Def: Cyclic Subspaces

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

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

(*) <u>Proof</u>: ($\Rightarrow$) Let $\lambda_1, \ldots, \lambda_n \in F$ be eigenvalues with a corresponding eigenbasis $\beta = (x_1, \ldots, x_n)$. Let $y = \sum_{i=1}^{n} x_i$, and let $y_j = T^{j-1} y$ $(1 \le j \le n)$. 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 $D = [[y_1]_\beta \ldots [y_n]_\beta] = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}$ $\rightarrow$ nonzero determinant $\Rightarrow$ invertible $\Rightarrow V = \text{span}\{y_1, \ldots, y_n\} = \text{span}\{y, Ty, \ldots, T^{n-1}y\}$.

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

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

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

$0 = g(T)(x) = T^r x + a_{r-1} T^{r-1} x + \ldots + a_0 x_0 \Rightarrow T^{r+k} x \in \text{span}\{T^{r-1}x, \ldots, x\} \; \forall k \ge 0$. □

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

(*) <u>Proof</u>: ($\Rightarrow$) Suppose V is T-cyclic, with $x_0 \in V$ s.t. $V = \text{span}\{x_0, Tx_0, \ldots\}$. 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 C-H). We know $g_T(T)(x_0) = 0$; then $V = \text{span}\{x_0, Tx_0, \ldots\} = \text{span}\{x_0, \ldots, T^{r-1}x_0\}$.

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

<u>Prop.</u> 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 <u>$p(t) \mid g_T(t)$</u>.

Matrix Polynomials

11/27/23

Lecture 22

Recall: Generalized Eigenspaces

Let $T_A: \mathbb{C}^n \to \mathbb{C}^n$, with eigenvalues $\lambda_1, \dots, \lambda_k$ ($k \leq n$). Then, given λ_j, the generalized eigenspace of λ_j is the set $\{v : (A - \lambda_j I)^\ell v = 0 \text{ for some } \ell \geq 1\} = \lim_{\ell \to +\infty} \ker (A - \lambda_j I)^\ell$.

Minimal Polynomials

Lemma. Let p be the minimal polynomial for $T_A: \mathbb{C}^n \to \mathbb{C}^n$; then the factorization of $p(x)$ is an expression of the form $p(x) = (x - \lambda_1)^{\ell_1} \dots (x - \lambda_k)^{\ell_k}$, where $\lambda_1, \dots, \lambda_k$ are eigenvalues of A.

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

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

(*) Proof: Suppose (contradiction) $\exists\, \lambda_0$ 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 $P(A) = (A - \lambda_0 I) Q(A) = 0 \Rightarrow (A - \lambda_0 I) Q(A) v = \vec{0} \ \forall\, v \in \mathbb{C}^n$. Then $Q(A)v \in \ker(A - \lambda_0 I)$; but λ_0 not an eigenvalue $\Rightarrow \ker(A - \lambda_0 I) = \{0\} \Rightarrow Q(A)v = \vec{0} \ \forall\, v \in \mathbb{C}^n$. Then $Q(A) = 0$, but $\deg Q < \deg P$; then $Q(x)$ can only be 0. □

Claim. Let $p(x) = (x - \lambda_1)^{\ell_1} \cdot \dots (x - \lambda_k)^{\ell_k}$ be the minimal polynomial. We claim that $1 \leq \ell_j \leq d_j \ \forall\, 1 \leq j \leq k$, where $d_j = \dim G_A(\lambda_j)$.

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

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.

Ex: (Group theory) Groups can be represented as invertible matrices via homomorphism $G \mapsto M_G$ (used in physics, e.g.)

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

Complex Inner Products

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

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

$\rightarrow$ Define inner product on $\mathbb{C}$: $\boxed{\langle (\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) \rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}}$ (*) Alt. $\langle\langle -,- \rangle\rangle$

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. $\langle v, \lambda_1 w_1 + \lambda_2 w_2 \rangle = \overline{\lambda_1}\langle v, w_1 \rangle + \overline{\lambda_2}\langle v, w_2 \rangle$
- Is conjugate-symmetric, i.e. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

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

Recall: Adjoints

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

(*) Also called a Hermitian adjoint

Hermitian Matrices

11/27/23
Lecture 27
(cont.)

Determining T^*

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: $\langle v, Te_j \rangle$ known $\rightarrow \langle T^* v, e_j \rangle$ known.

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

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

Diagonalizability

Recall: Let $A \in \mathbb{R}^{n \times n}$. Then $A = A^T$ [i.e. A symmetric] $\Rightarrow A$ 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 $\langle \lambda v, w \rangle = \langle T_A v, w \rangle = \overline{\langle w, Tv \rangle}$.
Know: $\langle w, Tv \rangle = \langle T^* w, v \rangle$. Recall: $\langle T^* v, v \rangle = \langle v, Tv \rangle = \bar{\lambda} \langle v, v \rangle$.
Since $T = T^*$: $\langle T^* v, v \rangle = \overline{\langle Tv, v \rangle} = \lambda \langle v, v \rangle$. Then $\lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}$. [since $|v| \neq 0$]
(*) To find additional eigenvectors: look at $v^{\perp} = \{ w : \langle v, w \rangle = 0 \}$ [T-invariant subsp.]. $\square$

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

Adjoints

11/28/23
Disc 16

Recall: Adjoints

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

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

(*) Proof: Since $\langle -,- \rangle$ is positive-definite, then for any vector $v_0 \in V$, $\langle v_0, v \rangle = 0$ $\forall v \in V$ if and only if $v = \vec{0}$. Equivalently: $\langle v, v_0 \rangle = \langle v, v_1 \rangle$ $\forall v \in V$ iff $v_0 = v_1$.

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 $S_1 y = S_2 y$; can repeat $\forall y \in W$. $\square$

Prop. Let $T: V \to V$ be a linear operator on V a fin.-dim. IPS over $\mathbb{R}$ ($IPS/\mathbb{R}$). Then if $\beta = (e_1, \dots, e_n)$ is an orthonormal basis, then $[T^*]_\beta = ([T]_\beta)^T$.

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

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

We know: $Te_j = \sum_{i=1}^n \langle Te_j, e_i \rangle e_i = \sum_{i=1}^n \langle e_j, T^* e_i \rangle e_i \to A_{ij} = \overline{\langle T^* e_i, e_j \rangle}$. Let $B = [T^*]_\beta$; then $A_{ij} = B_{ji} \Rightarrow A = B^T$. $\square$

Upper Triangularization

11/20/23

Lecture 28 (*) <u>Derivatives as Transformations</u>

<u>Recall</u>: $\int u\,dv = uv - \int v\,du$ [Integration by Parts] $\Rightarrow \int f\,dg = -\int df\,g$ (+ boundary term)

Define $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx \rightarrow \langle f, \frac{dg}{dx}\rangle = \langle -\frac{df}{dx}, g\rangle + \underbrace{\text{boundary term}}$ goes to 0

$\rightarrow \int_{-\infty}^{\infty} \frac{d}{dx}\left(\frac{d}{dx}f\right)\cdot g = -\int_{-\infty}^{\infty} \frac{df}{dx}\cdot\frac{dg}{dx} = -\left(-\int f \cdot \frac{d}{dx}\left(\frac{d}{dx}g\right)\right) = \int_{-\infty}^{\infty} f\,\frac{d^2g}{dx^2}$

↳ Observe: $\frac{d}{dx}$, $\frac{d^2}{dx^2}$ akin to self-adjoint operators

<u>Upper Triangularization [Schur's Theorem]</u>

<u>Recall</u>: $T_A^* = (\overline{T_A})^T$, i.e. $A^* = (\overline{A})^T$

<u>Claim</u>: Let $T_A : \mathbb{C}^n \to \mathbb{C}^n$. We claim that the matrix A is <u>upper-triangular</u> for some <u>basis</u> (besides Jordan form).

(*) <u>Proof</u>: We know A is a matrix of the form: $\begin{pmatrix} \uparrow & & \uparrow \\ v_1 & \cdots & v_n \\ \downarrow & & \downarrow \end{pmatrix}$

$\longrightarrow$ Set v_n to be an <u>eigenvector of T^*</u>, s.t. $T^* v_n = \lambda_n v_n$. Assume WLOG that $\|v_n\| = 1$, i.e. $\langle v_n, v_n\rangle = \underline{1}$

$\rightarrow$ <u>Claim</u>. $\underline{v_n^\perp = \{v \in \mathbb{C}^n : \langle v, v_n\rangle = 0\}}$ is a <u>T-invariant subspace</u>.

(*) <u>Proof</u>: 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 $v_1, \ldots, v_n$) is a matrix of the form: $\boxed{\left(\begin{array}{c|c} T(v_n^\perp) & \begin{matrix}\uparrow \\ v_n\end{matrix} \\ \hline 0 & \downarrow \end{array}\right)}$

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

<u>Observe</u>: Values on the diagonals are <u>eigenvalues of A</u>. ((*) Process only works in $\mathbb{R}^n$ if all eigenvalues real).

(*) Midterm Review Problems

11/30/23
Disc 17

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 $= \dim(\ker(A-\lambda I))$ and (b) # 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. $Be_i = \{0$ if $i \in \{i_1, \dots, i_k\}$, e_{i-1} otherwise$\}$.

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

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

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

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

$$Bx = B\left(\sum_{i=1}^{n} c_i e_i\right) = \sum_{i=1}^{n} c_i B(e_i) = \sum_{i \notin \beta} c_i B(e_i) \Rightarrow c_i = 0 \;\; \forall i \notin \beta \text{ [Since } e_1, \dots, e_n \text{ independent]}. \square$$

(b) Look at B^2: Blocks size 1 (0) in B $\to$ look the same in B^2 (0)

Blocks size ≥ 2 $\begin{pmatrix} 0&1&0\\0&0&1\\0&0&0 \end{pmatrix}$ in B $\to$ one smaller in B^2 $\begin{pmatrix} 0&0&1\\0&0&0\\0&0&0 \end{pmatrix}$

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

$\to$ 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 $T: \mathbb{C}^n \to \mathbb{C}^n$ is linear and self-adjoint $[T^* = T]$, then T has only real eigenvalues.

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

$\to$ $\lambda\|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, T^* v\rangle = \langle v, Tv\rangle = \overline{\langle Tv, v\rangle}$

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

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

Proof of Jordan Form

12/6/23

Lecture 29

__Recall__: Given $T: V \to V$, 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 $T: V \to V$ linear; let E_{λ_i} be a gen. eigenspace of T. Then $\exists$ a matrix representation of $T|_{E_{\lambda_i}}$ in Jordan form.
↳ true for any T-invariant subspace

(*) __Proof__: By induction on $\dim E_{\lambda_i}$. Base case: $\dim E_{\lambda_i} = 1 \to T|_{E_{\lambda_i}} = (\lambda_i)$ [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 $\vec{0}$.
↳ image $(T - \lambda_i I)$ on E_{λ_i}

If we look at (on E_{λ_i}) $\ker(T-\lambda_i I)$, $\operatorname{im}(T - \lambda_i I)$:

__Case 1__: $\ker(T-\lambda_i I) \cap \operatorname{im}(T-\lambda_i I) = \{\vec{0}\}$. Per Rank-Nullity: $\dim E_{\lambda_i} = \dim(\ker(T-\lambda_i I)) + \dim(\operatorname{im}(T-\lambda_i I))$; then $E_{\lambda_i} = \ker(T+\lambda_i I) \oplus \operatorname{im}(T+\lambda_i I)$. 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 $T|_{\operatorname{im}(T+\lambda_i I)}$ has Jordan form by induction.

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

We observe that for every vector $\vec{v} \in (\ker(T-\lambda_i I) \cap \operatorname{im}(T-\lambda_i I))$, $\exists\ \vec{v}' \in \operatorname{im}(T-\lambda_i I) \setminus \ker(T-\lambda_i I)$ such that $(T-\lambda_i I)\vec{v}' = \vec{v}$. Then for every $\vec{v}_0 \in \ker(T-\lambda_i I) \cap \operatorname{im}(T-\lambda_i I)$, we can find a Jordan chain of vectors $\vec{v}, (T-\lambda I)\vec{v}, \dots, (T-\lambda_i I)^s \vec{v} = \vec{v}_0$ [$s < \dim(\operatorname{im}(T-\lambda_i I))$] in E_{λ_i}. We observe that the span of the vectors in the chain are $(T-\lambda I)$-invariant and a Jordan basis for $(T-\lambda I)|_{\operatorname{span}\{v, \dots, (T-\lambda_i I)^s \vec{v}\}}$ ↳ w/ eigenvalue 0; then they are a Jordan basis for $T|_{\operatorname{span}\{v, \dots, (T-\lambda_i I)^s v\}}$. Per definition of E_{λ_i}, every vector not in $\ker(T-\lambda_i I)$ belongs to exactly one Jordan chain; then $E_{\lambda_i} = \ker(T-\lambda_i I) \oplus \operatorname{span}\{v_0, \dots, (T-\lambda_i I)^{s_0} v_0\} \oplus \dots \oplus \operatorname{span}\{v_r, \dots, (T-\lambda_i I)^{s_r} v_r\}$. $\square$