

Math 115B: Linear Algebra

2023-24

Spring 24

Instructor: Linfeng Li (+ Chuyin Jiang)

Textbook: S. Friedberg, A. Insel, & L. Spence - Linear Algebra (5th ed.)

Topics: Linear transformations & duality; invariant subspaces & the Cayley-Hamilton Theorem; unitary & orthogonal operators; projections & the Spectral Theorem; the Jordan canonical form

Table of Contents

- 1) Math 115A Review - 52
- 2) Dual Spaces - 59
- 3) Invariant Subspaces & The Cayley-Hamilton Theorem - 64
 - (i) Invariant Subspaces - 65
 - (ii) The Cayley-Hamilton Theorem - 66
- 4) Inner Product Spaces & Orthogonality - 67
 - (i) Inner Product Spaces - 67
 - (ii) Adjoints of Linear Operators - 70
 - (iii) Unitary & Orthogonal Operators - 70
 - (iv) Rigid Motions - 74
- 5) Projections & The Spectral Theorem - 76
 - (i) Projections & Orthogonal Complement - 77
 - (ii) Orthogonal Projections - 78
 - (iii) The Spectral Theorem - 81
 - (iv) The Geometry of Orthogonal Operators - 85
- 6) The Jordan Canonical Form - 87
 - (i) The Jordan Decomposition Theorem - 90
 - (ii) The Minimal Polynomial - 95

Vector Spaces (Review)

4/1/24

Lecture 1

Vector Spaces (Review)

(generally: \mathbb{R} [real nos.], \mathbb{C} [complex nos.])

Def: A vector space V over field F is a set with associated operations of addition & scalar multiplication such that $[\forall x, y \in V; a, b \in F]$:

$$(i) x+y = y+x$$

$$(v) 1 \cdot x = x \quad [1 \in F]$$

$$(ii) (x+y)+z = x+(y+z)$$

$$(vi) a(bx) = (ab)x$$

$$(iii) \exists 0_v \in V \text{ s.t. } x+0_v = x$$

$$(vii) a(x+y) = ax + ay$$

$$(iv) \forall x \in V, \exists -x \in V \text{ s.t. } x+(-x) = 0_v$$

$$(viii) (a+b)x = ax + bx$$

(*) Notable vector spaces: (i) All fields (e.g. \mathbb{R}) over themselves

(ii) \mathbb{R}^n : vectors [real-number tuples] of length n

(iii) $P_n(\mathbb{R})$: polynomials w/ degree $\leq n$, real coefficients

(*) $P(\mathbb{R})$: polynomials w/ real coefficients

(iv) $M_{n \times n}(\mathbb{R})$: $n \times n$ real matrices

Def (Subspace): Given V a vector space [v.s.] over field F , a subset $W \subseteq V$ is called a subspace of V if W is itself a vector space over F .

Thm: Let V a v.s. over F , then a subset $W \subseteq V$ is a subspace of V iff W satisfies:

$$(i) 0_v \in W$$

$$(ii) x, y \in W \Rightarrow x+y \in W \quad [W \text{ closed under addition}]$$

$$(iii) c \in F, x \in W \Rightarrow cx \in W \quad [W \text{ closed under scalar multiplication}]$$

→ Corollary: W is a subspace of V iff W satisfies: (i) $0_v \in W$

$$(ii) ax+ay \in W \quad \forall x, y \in W; a \in F$$

Linear Independence (Review)

4/1/24

Lecture 1

(cont.)

Thm: Let V a vector space over \mathbb{F} with W_1, \dots, W_n subspaces of V ; then $\bigcap_{i=1}^n W_i$ [intersection, not union] is a subspace of V . [$\bigcap_{i=1}^n W_i \subseteq V$]

Linear Combinations (Review)

Def: Let V a v.s. over \mathbb{F} , $S \subseteq V$ nonempty:

→ (i) A vector $v \in V$ is called a linear combination of vectors in S if v can be written as $v = a_1v_1 + \dots + a_nv_n$ for some $v_1, \dots, v_n \in S$; $a_1, \dots, a_n \in \mathbb{F}$

(ii) The span of S is defined as the set of all linear combinations of vectors in S :

$$\text{span}(S) := \left\{ a_1v_1 + \dots + a_nv_n : a_1, \dots, a_n \in \mathbb{F}; v_1, \dots, v_n \in S \right\}$$

Span: - Let $S \subseteq V$ nonempty $\Rightarrow \text{span}(S)$ is a subspace of V

- Any subspace of V containing S ($S \subseteq U$) contains $\text{span}(S)$; in particular, $\text{span}(S)$ is the smallest subspace of V that contains S

Def: Let V a v.s., $S \subseteq V$; we say that S generates V if $\text{span}(S) = V$.

[To prove, show $\text{span}(S) \subseteq V$ and $\text{span}(S) \supseteq V$]

Def: Linear Independence

A subset $S \subseteq V$ is called linearly dependent if \exists distinct vectors $v_1, \dots, v_n \in S$ and scalars

$a_1, \dots, a_n \in \mathbb{F}$ such that:

(i) $a_1v_1 + a_2v_2 + \dots + a_nv_n = 0_V$, and

(ii) at least one a_i is nonzero.

We say S is linearly independent if S is not linearly dependent.

Bases & Dimension (Review)

4/3/24

Lecture 2

Prop: A subset $S \subseteq V$ is linearly independent iff " $v_1, \dots, v_n \in V; a_1v_1 + \dots + a_nv_n = 0_V \Rightarrow a_1 = a_2 = \dots = a_n = 0$ ".

Bases & Dimension (Review)

Def: Let V a v.s.; a basis for V is a linearly independent subset of V that spans/generates V .

- (*) Notable bases:
 - (i) Standard basis for \mathbb{R}^n : $B = \{e_1, \dots, e_n\}$, where $e_i = (0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$
 - (ii) $B = \{1, x, \dots, x^n\}$ for $P_n(\mathbb{R})$

Thm: Let V a vector space and $u_1, u_2, \dots, u_n \in V$; then $\{u_1, \dots, u_n\}$ is a basis for V iff every $v \in V$ can be uniquely written as a linear combination of $\{u_1, \dots, u_n\}$.

Prop. Let V be a vector space with a finite basis of size n ; then every basis for V will have size n .

→ Def: A vector space is called finite-dimensional if it has a finite basis; we call the size of this basis, the dimension of the vector space.

Dimension:

- A vector space is called infinite-dimensional if it is not finite-dimensional [\exists a finite basis]; denote this as $\dim V = \infty$ (vs finite-dimensional: $\dim V < \infty$)

- We define $\dim(\{0\}) = 0$ [dimension 0]

(*) Notation: " $\dim V = n \Rightarrow V$ is finite-dimensional with dimension n "

Prop: Let V a vector space w/ $\dim(V) = n$. Then:

- (i) Any [finite] generating set for V has at least n elements
- (ii) Any generating set for V of size n [exactly n] is necessarily a basis for V
- (iii) Any linearly independent subset of V with size n is necessarily a basis for V
- (iv) Every linearly independent subset of V can be extended to a basis for V

Linear Transformations (Review)

4/3/24

Lecture 2

Linear Transformations (Review)

domain \rightarrow codomain

Def: Let V, W be vector spaces over a field F ; then a linear transformation from V to W is a

map $T: V \rightarrow W$ that is linear, i.e. that satisfies:

$$(i) T(x+y) = T(x) + T(y) \quad [\forall x, y \in V; c \in F]$$

$$(ii) T(cx) = c \cdot T(x)$$

(cont.)

Linear transformations: (i) T linear \Rightarrow (a) $T(0_v^{(v)}) = 0_w^{(w)}$ [$0_v^{(v)} \in V, 0_w^{(w)} \in W$]

$$(b) T\left(\sum_{i=1}^n c_i v_i\right) = \sum_{i=1}^n c_i T(v_i) \quad [\text{via induction}]$$

(ii) For proofs: T is linear iff $T(cx+y) = cT(x) + T(y) \quad \forall x, y \in V; c \in F$

(*) Notation: T may also be called a linear operator/linear map.

Def: Let $T: V \rightarrow W$ be linear; then we say:

(i) T is injective/one-to-one if $T(u) = T(v) \Rightarrow u = v \quad \forall u, v \in V$

(ii) T is surjective/onto if $\forall w \in W, \exists v \in V$ s.t. $T(v) = w$

(iii) T is bijective/invertible if T is both injective and surjective

Def: Let $T: V \rightarrow W$ be linear; then we define the (i) null space and (ii) range space of T :

$$(i) \text{ Null space } \text{null}(T): \boxed{N(T) = \{v \in V : T(v) = 0_w\}}$$

$$(ii) \text{ Range/image } \text{im}(T): \boxed{R(T) = \{w \in W : w = T(v) \text{ for some } v \in V\}}$$

Prop: (i) $N(T) \subseteq V$ is a subspace of V

(ii) $\text{im}(T) \subseteq W$ is a subspace of W

Linear Transformations (Review)

4/3/24

Lecture 2

(cont.)

Theorem (Rank-Nullity): Let $\dim V = n$ and $T: V \rightarrow W$ linear; then:

$$\Rightarrow \boxed{\dim [\text{im}(T)] + \dim [\text{null}(T)] = \dim(V) [= n]}$$

(*) Notation/Def. (i) Rank: $\text{rank}(T) = \dim(\text{im}(T))$

(ii) Nullity: $\text{nullity}(T) = \dim(\text{null}(T))$

Matrix Representations of Linear Maps

Def: Let $\dim V < \infty$. An ordered basis for V is a basis equipped with an order (e.g. e_1, e_2, \dots).

Recall: Let $B = \{v_1, \dots, v_n\}$ be an ordered basis for V ; then any vector $\vec{v} \in V$ can be expressed uniquely as $\vec{v} = a_1v_1 + a_2v_2 + \dots + a_nv_n$ for some scalars $a_1, \dots, a_n \in F$

→ Def: Define the coordinates of \vec{v} relative to B as:

$$\boxed{[\vec{v}]_B = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad ([\vec{v}]_B \in \mathbb{R}^n)}$$

Def: Let $T: V \rightarrow W$ linear and $B = \{v_1, v_2, \dots, v_n\}$, $\gamma = \{w_1, w_2, \dots, w_m\}$ ordered bases for V and W , respectively. Then the matrix representation of T is defined as:

$$\boxed{[T]_{\gamma}^B = \left(\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ [T(v_1)]_{\gamma} & [T(v_2)]_{\gamma} & \cdots & [T(v_n)]_{\gamma} \\ \downarrow & \downarrow & \downarrow \end{array} \right) \in M_{m \times n}}$$

(*) Notation: If T is $V \rightarrow V$ (and B an ordered basis for V), may write $[T]_B = [T]_B^B$.

(*) Infinite-Dimensional Vector Spaces

4/4/24

(*) Linear Independence of Infinite Sets

Discussion 1

Def: Let V a vector space, $S \subseteq V$ with $\dim S = \infty$. We say that S is linearly independent if every finite subset $A \subseteq S$ is linearly independent; otherwise, say that S is linearly dependent.

(*) Note: In general, in infinite-dimensional vector spaces, certain theorems (e.g. rank-nullity) may no longer apply.

(*) Fields: A field can either have:

(i) $|F| = \infty$, or (ii) $|F| < \infty$ [finite field] $\Rightarrow |F| = p^k$ for some prime, $k \in \mathbb{Z}^+$

(*) Infinite-Dimensional Vector Spaces

Ex: (i) $P(\mathbb{R})$: vector space of all polynomials with real coefficients

(ii) $\mathbb{R}^\infty = \{(a_i)_{i=0}^\infty : a_i \in \mathbb{R}\}$: vector space of all [infinite] real sequences

note: all polynomials $p(x) \in P(\mathbb{R})$ are finite-dimensional [have "final" nonzero coefficient]
 $\rightarrow \dim(\mathbb{R}^\infty) > \dim(P(\mathbb{R}))$

Infinite dimensions: (i) The existence of an injective (but not surjective) f'n $f: A \rightarrow B$ does not imply $\dim(A) < \dim(B)$; ex: $f: \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$, $(1, 2, \dots) \mapsto (0, 1, 0, 2, \dots)$

(ii) $|\mathbb{Q}| > |\mathbb{R}|$

For any set S & field F , the set $L(S, F)$ of linear functions $S \rightarrow F$ is a vector space

Invertibility & Isomorphism (Review)

4/5/24

Lecture 3

Theorem: Let V, W fin. dim. w/ ordered bases B, Y and let $T_1, T_2: V \rightarrow W$ linear; then:

$$(i) T_1 = T_2 \Leftrightarrow [T_1]_B^Y = [T_2]_B^Y$$

$$(ii) [T_1 + T_2]_B^Y = [T_1]_B^Y + [T_2]_B^Y$$

$$(iii) [cT_1]_B^Y = c[T_1]_B^Y \quad \forall c \in F$$

Composition of Linear Transformations (Review)

Thm: Given vector spaces V, W, Z with ordered bases α, B , and Y , respectively; and

$T_1: V \rightarrow W, T_2: W \rightarrow Z$ linear, then their composition $T_2 \circ T_1: V \rightarrow Z$ is linear and:

$$[T_2 \circ T_1]_\alpha^Y = [T_2]_B^Y [T_1]_\alpha^B$$

Thm: Let V, W fin. dim. v.s.'s with ordered bases B and Y , and $T: V \rightarrow W$ linear; then:

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

Invertibility & Isomorphisms (Review)

Def: Let V, W v.s.'s and $T: V \rightarrow W$ linear; then a function $U: W \rightarrow V$ is called the inverse of T

& $T \circ U = \text{id}_W$ and $U \circ T = \text{id}_V$ [Identity function $V \rightarrow V$ and $W \rightarrow W$].

If T has an inverse, say T is invertible.

Facts: (i) If T is invertible, then its inverse is unique [denoted T^{-1}].

(ii) T is invertible $\Leftrightarrow T$ is bijective $\Leftrightarrow T$ surjective and injective

(*) $T: V \rightarrow W$ linear $\rightarrow T$ surjective iff $\text{im}(T) = W$; T injective iff $\text{kern}(T) = \{0\}$

(**) $T: V \rightarrow W, \dim V = \dim W \rightarrow T$ injective $\Leftrightarrow T$ surjective

Dual Spaces

4/5/24

Lecture 3

+ Lecture 4

Facts (cont.): If T_1, T_2 composable, linear, and invertible, then:

(a) $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$ exists [$T_2 \circ T_1$ invertible]

(b) (T_1^{-1}) is invertible with $(T_1^{-1})^{-1} = T_1$

Thm: Let V, W fin. dim. v.s.'s with ordered bases B, γ and $T: V \rightarrow W$ linear; then T is invertible $\Leftrightarrow [T]_{\beta}^{\gamma}$ invertible; furthermore, $[T^{-1}]_{\beta}^{\gamma} = ([T]_{\beta}^{\gamma})^{-1}$.

Def: Let V, W be vector spaces; we say that V is isomorphic to W if \exists an isomorphism [invertible linear transformation] from V to W .

Dual Spaces

Def: Let V a v.s. over F . A linear functional on V is a linear transformation from V to F .

(*) Ex: $A \mapsto \text{tr}A$ is a linear functional $M_{n \times n}(F) \rightarrow F$

Def: Let V a v.s. w/ ordered basis $B = \{v_1, \dots, v_n\}$. The i^{th} coordinate function w.r.t. B is the function $f_i(v) = a_i$, where $v = a_1v_1 + a_2v_2 + \dots + a_nv_n$ ($[v]_B = (a_1, a_2, \dots, a_n)$).

(*) Notation: Also denoted as $f_i(v_j) = \delta_{ij}$ [$= 1$ if $i=j$; 0 otherwise]

Def: Given a vector space V over field F , define the dual space of V as $V^* = L(V, F)$ [the set of linear transformations $V \rightarrow F$].

Note: Let V a v.s. over F ; then V^* is itself a vector space over F .

18/24

Dual Spaces (cont.)

Lecture 4

Lecture 5

Thm: Let V a v.s. with ordered basis $\beta = \{v_1, \dots, v_n\}$. We define the dual basis of β to be the set $\beta^* = \{f_1, \dots, f_n\}$, where f_i is the i^{th} coordinate function w.r.t. β .

Prop: Let V a v.s. with ordered basis $\beta = \{v_1, \dots, v_n\}$; then the dual basis β^* is a basis for V^* . In particular, for any $f \in V^*$:

$$\boxed{f = \sum_{i=1}^n f(v_i) f_i \quad \forall f: V \rightarrow F \in V^*}$$

(*) Proof: Let $v = a_1 v_1 + a_2 v_2 + \dots + a_n v_n [v \in V]$.

$$\begin{aligned} \text{Then } \left[\sum_{i=1}^n f(v_i) f_i \right] (v) &= \sum_{i=1}^n [f(v_i) f_i](v) \\ &= \sum_{i=1}^n f(v_i) \cdot f_i(v) \\ &= \sum_{i=1}^n f(v_i) \cdot a_i \\ &= \sum_{i=1}^n a_i f(v_i) = f\left(\sum_{i=1}^n a_i v_i\right) = f(v). \end{aligned}$$

□

Def: Let V, W f.d.v.s.'s with ordered bases β and γ , respectively. Let $T: V \rightarrow W$ linear; we define the transpose of T to be the function $T^*: W^* \rightarrow V^*$ defined by:

$$\boxed{T^*(g) = g \circ T \quad \forall g \in W^*}$$

Theorem: Let V, W, β, γ as above; let $T: V \rightarrow W$ linear. Then T^* is a linear transformation from V to W ; in particular:

$$\boxed{[T^*]_{\beta^*}^{\gamma^*} = ([T]_{\beta}^{\gamma})^*}$$

Dual Spaces (cont.)

4/10/24

Lecture 5

(*) Proof: Need to check 3 claims:

$$(i) T^*(g) \in V^* \text{ [is linear]} \quad \forall g \in W^*: (T^*(g))(0) = g \circ T(0) = g(0) = 0$$

$$(T^*(g))(cx+y) = g(T(cx+y)) = g(cxT(x)+T(y)) = cg(T(x))+g(T(y)) = c(T^*(g))(x) + (T^*(g))(y),$$

$$(ii) T^* \text{ is linear: } (T^*(0))(v) = 0 \quad [\text{trivial}]$$

$$(T^*(cg+h))(v) = (cg+h)(T(v)) = cg(T(v)) + h(T(v)) = (cT^*(g) + T^*(h))(v)$$

$$(iii) [T^*]_{\gamma^*}^{B^*} = ([T]_B^\gamma)^+: \text{ Let } B = \{x_1, \dots, x_n\} \text{ & } \gamma = \{y_1, \dots, y_n\}; B^* = \{f_1, \dots, f_n\}, \gamma^* = \{g_1, \dots, g_m\}.$$

Let $A = [T]_{\gamma}^B$, i.e.:

$$A = \begin{pmatrix} [T(x_1)]_\gamma & \dots & [T(x_n)]_\gamma \end{pmatrix} = (A_{ij})$$

Looking at $[T]_{\gamma^*}^{B^*} = B$:

$$[T^*]_{\gamma^*}^{B^*} = \left([T^*(g_1)]_{B^*} \dots [T^*(g_m)]_{B^*} \right) = B$$

$$\text{Recall that } T^*(g_i) = g_i \circ T = \sum_{j=1}^m (g_i \circ T)(x_j) \cdot f_j$$

$$\rightarrow B_{ij} = (g_i \circ T)(x_j) = g_i \left(\sum_{k=1}^n A_{kj} y_k \right) = \sum_{k=1}^n A_{kj} g_i(y_k) = A_{ji}$$

Rec: The vector spaces $L(V, W)$ [$\dim V = n, \dim W = m$] and $M_{mn}(F)$ are isomorphic through the isomorphism $\Phi_B^* : L(V, W) \rightarrow M_{mn}(F)$ defined by $T \mapsto [T]_B^\gamma$.

→ Cor: The linear transformation T^* is the unique linear map with $[T^*]_{\gamma^*}^{B^*} = ([T]_B^\gamma)^+$.

$$(*) \dim(L(V, W)) = \dim(M_{mn}(F)) = m \cdot n.$$

4/11/24

(*) Dual Spaces

Discussion 2

(*) Dual Spaces

Let V a fin.-dim. v.s. over F w/ basis $\beta = \{v_1, \dots, v_n\}$

$\rightarrow V^* = L(V, F) := \{\text{linear transformations } V \rightarrow F\}$ is a vector space; called dual space of V
 with dual basis $\{f_i : v_j \mapsto \delta_{ij}\}$ [$\rightarrow f \in V^* \Rightarrow f = \sum_{i=1}^n f(v_i) f_i$]

Can take double dual V^{**} of V , dual space of V^* : $V^{**} = L(V^*, F)$

Define $e_j(f_i) = f_i(v_j) \rightsquigarrow$ form basis for V^{**} : $e(f \in V^*) = (\sum_{i=1}^n e_i(f_i))(f) = f(\sum_{i=1}^n e_i v_i)$
 Evaluating f_i at v_j

Observe: $c = c_1 e_1 + \dots + c_n e_n \Rightarrow e(f) = f(c_1 v_1 + \dots + c_n v_n)$, evaluates f at $v = c_1 v_1 + \dots + c_n v_n$

In particular can find an isomorphism $V \rightarrow V^{**}$, $v \mapsto e_v$ [$e_v : V^* \rightarrow F$, $f \mapsto f(v)$]

* isomorphism is basis-agnostic; i. does not depend on choice of basis (or having a basis)
 (alt.: is a canonical isomorphism)

Consequence: Taking infinite duals of V gives just two vector spaces: V [odd duals], V^* [even duals]

(*) Ex: V^{***} isomorphic to V^*

Let V, W vector spaces; then any $T: V \rightarrow W$ linear induces transpose $T^*: W^* \rightarrow V^*$ defined

by $T^*(f \in W^*) = f \circ T$

- For fixed bases β, γ for V, W + duals β^*, γ^* for V^*, W^* : $[T^*]_{\gamma^*}^{\beta^*} = ([T]^{\gamma}_{\beta})^T$

Dual Spaces (cont.)

4/12/24

Lecture 6

Isomorphisms

Thm: Let V, W v.s.'s over field F ; then V isomorphic to $W \Leftrightarrow \dim(V) = \dim(W)$.

Dual Spaces (cont.)

Let $\dim V < \infty$; may define $V^{***} := (V^*)^*$.

\rightarrow Observe: $\dim(V) = \dim(V^*) = \dim(V^{**}) \Rightarrow V, V^{***}$ are isomorphic

Def: For $x \in V$, define $\hat{x}: V^* \rightarrow F$ by $\hat{x}(f) = f(x) \quad \forall f \in V^*$. $[\hat{x} \in V^{**}]$

(*) Proof $[\hat{x} \in V^{**}; \hat{x}$ linear]: $\hat{x}(0) = 0$, trivial

$$\hat{x}(cf + g) = (cf + g)(x) = cf(x) + g(x) = c\hat{x}(f) + \hat{x}(g).$$



Thm: Let $\dim V < \infty$; then the function $\Psi: x \mapsto \hat{x}$ is an isomorphism $V \rightarrow V^{**}$.

$$\boxed{\begin{array}{l} x \in V \\ \Psi(\) \in V^{**} \\ \hat{x} \in V^{***} \end{array}}$$

(*) Proof

Lemma: If $\hat{x}(f) = 0 \quad \forall f \in V^*$, then $x = 0_V$.

(*) Proof: $x = a_1v_1 + \dots + a_nv_n$; $\hat{x}(f_i) = 0 \quad \forall i \Rightarrow a_1 = a_2 = \dots = a_n = 0$.

$$= \hat{(cx+y)}(f)$$

(i) Ψ is linear: $x, y \in V; c \in F \rightarrow (\Psi(cx+y))(f) = f(cx+y) = cf(x) + f(y) = (c\Psi(x) + \Psi(y))(f)$.

(ii) Ψ is bijective: $N(\Psi) = \{0\} \Leftrightarrow \Psi$ injective.

By lemma: since " $\hat{x}=0$ " $\Rightarrow x=0_V$, then $N(\Psi) = \{0_V\} \Rightarrow \Psi$ injective.

$\wedge \Psi$ linear

Since $\dim V = \dim V^{***}$, then Ψ injective $\Leftrightarrow \Psi$ bijective.



4/12/24

Lecture 6

Eigenvalues & Eigenvectors

Eigenvalues & Eigenvectors

Def. Let T a linear operator on V [i.e. $T: V \rightarrow V$]; then a vector $v \in V$, $v \neq 0_v$ is called an eigenvector of T if $T(v) = \lambda v$ for some scalar $\lambda \in F$ (called an eigenvalue of T).

Thm: Let $\dim V < \infty$, T a linear operator on V . Then for any two bases B, γ for V , have that $\det([T]_B) = \det([T]_\gamma)$, i.e. determinant does not depend on basis. Can denote this unique value $\det(T)$.

Prop: Let T a linear operator on V ; then:

(i) T is bijective $\Leftrightarrow \det(T) \neq 0$.

(ii) If T is bijective, then $\det(T^{-1}) = \frac{1}{\det(T)}$.

(iii) Let U another linear operator on V ; then $\det(T \circ U) = \det(T) \cdot \det(U)$.

Thm: Let $A \in M_{n \times n}(F)$; then a scalar $\lambda \in F$ is an eigenvalue of A iff:

$$\boxed{\det(A - \lambda I_n) = 0}$$

→ Def: Define the characteristic polynomial of A as the polynomial $f(t) = \det(A - tI_n)$.

↪ (*) Note: (i) The characteristic polynomial of $A \in M_{n \times n}(F)$ has degree n & leading coeff. $(-1)^n$

(ii) A matrix $A \in M_{n \times n}(F)$ has at most n distinct eigenvalues.

Invariant Subspaces

4/15/24

Lecture 7

Invariant Subspaces

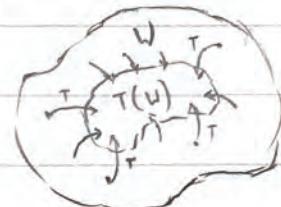
Def: Let T a linear operator on V . We call a subspace $W \subseteq V$ T -invariant if $T(W) \subseteq W$, i.e. $T(w) \in W \forall w \in W$.

Def: Let T a lin. op. on V , $W \subseteq V$ T -invariant. We define the restriction of T to W as the function $T_W: W \rightarrow W$ defined by $T_W(w) = T(w) \forall w \in W$.

linear operator

(*) Notable T -invariant subspaces: $\{0\}$, V , $\text{im}(T)$, $\text{null}(T)$

$\dim V < \infty$



Thm: Let T a linear operator on V , and let $W \subseteq V$ T -invariant. Then the characteristic polynomial of T_W divides the characteristic polynomial of T .

(*) Proof: Let $\dim W = k \leq \dim V = n$. Let $\gamma = \{v_1, \dots, v_k\}$ be an ordered basis for W , and let $\beta = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ be its extension to a basis for V . Let $A = [T]_{\beta}$, $B = [T_W]_{\gamma}$; have that

$$A = \begin{pmatrix} [T(v_1)]_{\beta} & \dots & [T(v_n)]_{\beta} \end{pmatrix}; \quad B = \begin{pmatrix} [T_W(v_1)]_{\gamma} & \dots & [T_W(v_k)]_{\gamma} \end{pmatrix}$$

Observe that since W is T -invariant, then $T(v_i)$ $[1 \leq i \leq k]$ only maps to v_1, \dots, v_k ; namely:

$$[T_W(v_i)]_{\gamma} = \langle a_1, \dots, a_k \rangle \Rightarrow [T(v_i)]_{\beta} = (a_1, \dots, a_k, 0, \dots, 0) = \langle [T_W(v_i)]_{\gamma}, 0, \dots, 0 \rangle$$

Thus:

$$A = \begin{pmatrix} [T(v_1)]_{\beta} & \dots & [T(v_k)]_{\beta} & [T(v_{k+1})]_{\beta} & \dots & [T(v_n)]_{\beta} \end{pmatrix} = \begin{pmatrix} B & | & C \\ \hline 0 & \dots & 0 & D & I_{n-k} \end{pmatrix}$$

$$\rightarrow \det(A - tI_n) = \det \begin{pmatrix} B - tI_k & | & C \\ \hline 0 & \dots & 0 & D - tI_{n-k} \end{pmatrix} = \det(B - tI_k) \det(D - tI_{n-k}).$$

char. poly. of T_W

The Cayley-Hamilton Theorem

4/17/24

Lecture 8

The Cayley-Hamilton Theorem

Def: Let T be a linear operator on V , and let $v \in V$, $v \neq 0_v$. Then we call the subspace $W = \text{span}\{v, T(v), T^2(v), \dots\}$, the T -cyclic subspace generated by v .

From last time: saw that the characteristic polynomial of T_w provides info about T ; want to find W with easy-to-find char. poly. [Observe: W T -cyclic $\Rightarrow W$ T -invariant]

Thm. Let T linear operator on V , and let W be the T -cyclic subspace generated by $v \in V$ [$v \neq 0_v$]; let $\dim(W) = k$. Then:

(i) The set $\{v, T(v), \dots, T^{k-1}(v)\}$ is a basis for W

(ii) Let $a_0, \dots, a_{k-1} \in F$ s.t. $a_0 v + a_1 T(v) + \dots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$; then the characteristic polynomial of T_w is:

$$f(t) = (-1)^k (a_0 + a_1 t + \dots + a_{k-1} t^{k-1} + t^k)$$

Corollary (Cayley-Hamilton): Let T be a linear operator on V [$\dim V < \infty$], and let $f(t)$ be the characteristic polynomial of T . Then $f(T)$ is the zero transformation. [$(f(T))(v) = 0 \forall v \in V$]

Alternatively, let $A \in M_{n \times n}(F)$ with characteristic polynomial $f(t)$; then:

$$f(A) = 0_{n \times n} \quad [\text{n} \times n \text{ matrix of all zeroes}]$$

(*) Proof: Let $v \in V$ nonzero; want to show that $(f(T))(v) = 0$. Let W be the T -cyclic subspace generated by v , with char. poly $g(t)$; by the theorem, have that $g(t) = (-1)^k (a_0 + a_1 t + \dots + a_{k-1} t^{k-1} + t^k)$ for some polynomial h ($a_i \in F$). $(g(T))(v) = (-1)^k (a_0 + a_1 T + \dots + a_{k-1} T^{k-1})(v) = 0$. Know that W T -invariant $\Rightarrow g(t)$ divides $f(t)$; then $(f(T))(v) = (g(T)(v))h(T)(v) = 0$. \square

Inner Products

4/19/24

Lecture 9

Def: Let V v.s. over field \mathbb{C} . A complex inner product on V is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ that satisfies $\forall x, y, z \in V$ and $c \in \mathbb{C}$:

(i) Linearity: $\langle cx + z, y \rangle = c\langle x, y \rangle + \langle z, y \rangle$

(ii) Conjugate symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle}$

(iii) Positive-definite: If $x \neq 0_V$, then $\langle x, x \rangle > 0$ and is real.

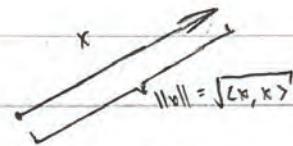
Properties: (i) $\langle 0_V, x \rangle = \langle x, 0_V \rangle = 0$

(ii) $\langle x, x \rangle$ is real and non-negative

(iii) $\langle x, x \rangle = 0$ iff $x = 0_V$

(iv) $\langle x, ay + bz \rangle = \bar{a}\langle x, y \rangle + \bar{b}\langle x, z \rangle$ [conjugate linearity]

(v) $\langle x+y, x+y \rangle = \langle x, x \rangle + 2\operatorname{Re} \langle x, y \rangle + \langle y, y \rangle$



(*) Ex: Define the standard inner product on \mathbb{C}^n as $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$, $\mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$

Def: Let V v.s. over field \mathbb{R} . A real inner product on V is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$

that satisfies $\forall x, y, z \in V$ and $c \in \mathbb{R}$:

(i) Linearity: $\langle cx + z, y \rangle = c\langle x, y \rangle + \langle z, y \rangle$

(ii) Symmetry: $\langle x, y \rangle = \langle y, x \rangle$

(iii) Positive-definite: If $x \neq 0_V$, then $\langle x, x \rangle$ is positive [and real].

Properties: Some as above, but applied to real numbers: (iv) $\bar{a}, \bar{b} [\mathbb{C}] \rightarrow a, b [\mathbb{R}]$

(v) $\operatorname{Re} \langle x, y \rangle [\mathbb{C}] \rightarrow \langle x, y \rangle [\mathbb{R}]$

(*) Ex: Define the standard inner product on \mathbb{R}^n as $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$, $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$

Inner Product Spaces

4/19/24

Lecture 9
(cont.)

(* Note: Define an inner product space as a vector space equipped with an inner product

Inner Product Spaces

Def: Let $A \in M_{n \times n}(F)$, where $F = \mathbb{R}$ or \mathbb{C} . We define the conjugate transpose of A to be the matrix $A^* \in M_{n \times n}(F)$ such that $(A^*)_{ij} = \overline{A}_{ji}$.

Def: Let V a real or complex inner product space. For $x \in V$, define the norm of x by:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Prop. Let V real or complex IPS; if $\langle x, y \rangle = \langle x, z \rangle \forall x \in V$, then $y = z$.

Theorem: Let V be an IPS over F , where $F = \mathbb{R}$ or \mathbb{C} . Then $\forall x, y \in V$ and $c \in F$:

(i) $\|cx\| = |c| \cdot \|x\|$

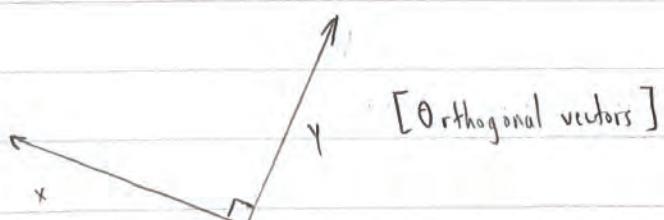
(ii) $\|x\| \geq 0 \forall x \in V$. In particular, $\|x\| = 0$ if $x = 0_V$.

(iii) Cauchy-Schwarz inequality: $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$.

(iv) Triangle inequality: $\|x+y\| \leq \|x\| + \|y\|$.

Def: Let V real/complex IPS; Say that two vectors $x, y \in V$ are orthogonal if $\langle x, y \rangle = 0$.

A subset $S \subseteq V$ is called orthogonal if any 2 distinct vectors in S are orthogonal, and called orthonormal if it is orthogonal and consists only of vectors with norm 1.



Orthonormal Bases

4/19/24

Lecture 9.1

+ Lecture 10

Orthonormal Bases

Def: Let V be an IPS. A subset $B \subseteq V$ is an orthonormal basis for V if it is an ordered basis for V and orthonormal i.e. for any $v, w \in B$:

$$\langle v, w \rangle = \begin{cases} 1 & v=w \\ 0 & v \neq w \end{cases}$$

Theorem: Let V be a nonzero finite-dimensional inner product space; then V admits an orthonormal basis B . Furthermore, if $B = \{v_1, \dots, v_n\}$ and $x \in V$, then:

$$x = \sum_{i=1}^n \langle x, v_i \rangle v_i$$

Adjoints of Linear Operators

Def: Let V be a fin.-dim. IPS, and let T be a linear operator on V . Then we define the adjoint of T to be the [unique] function $T^*: V \rightarrow V$ that satisfies $\forall x, y \in V$:

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle$$

Prop. The adjoint T^* always exists for any lin. op. T . Furthermore, T^* is linear.

(*) Gram-Schmidt (Review)

Theorem: Let $S = \{v_1, v_2, \dots, v_k\}$ orthornormal set in a vector space V ($\dim V = n$). Then $k \leq n$ and S can be extended to an orthonormal basis for V . In particular, if $k = n$, then S is an orthonormal basis for V .

Adjoints of Linear Operators

4/24/24

Lecture 10

(cont.)

Adjoints of Linear Operators

Theorem: Let V IPS and let T, U linear operators on V . Then:

- (i) $T+U$ has adjoint $(T+U)^* = T^* + U^*$
- (ii) For any $c \in F$, cT has adjoint $(cT)^* = \bar{c}T^*$.
- (iii) $T \circ U$ has adjoint $(TU)^* = U^* \circ T^*$.
- (iv) T^* has adjoint $T^{*\star} = T$.
- (v) The identity mapping has adjoint $I^* = I$.

Theorem: Let V be a fin-dim. IPS, and let B be an orthonormal basis for V . Let T be a linear operator on V ; then:

$$[T^*]_B = [T]_B^* \quad [\text{conjugate transpose}]$$

Self-Adjoint Operators

Def: Let T linear operator on IPS V . We say that T is self-adjoint if $T = T^*$. Similarly, say that an $n \times n$ [real or complex] matrix A is self-adjoint if $A = A^*$.

Unitary & Orthogonal Operators

Def: Let T linear operator on fin.-dim. IPS V (over field $F = \mathbb{R}$ or \mathbb{C}). If $\|T(x)\| = \|x\| \forall x \in V$, we say that:

- (a) T is a unitary operator if $F = \mathbb{C}$
- (b) T is an orthogonal operator if $F = \mathbb{R}$

(* Ex: Rotation [rotation matrices], reflection are unitary operators

Unitary & Orthogonal Operators

4/24/24

Lemma: Let U a self-adjoint operator on an IPS V , and suppose that $\langle x, U(x) \rangle = 0 \forall x \in V$. Lecture 10

Then U is the zero operator. + Lecture 11

(*) Proof: For any $x \in V$, have:

$$\begin{aligned} 0 &= \langle x + U(x), U(x + U(x)) \rangle \\ &= \langle x + U(x), U(x) + U^2(x) \rangle \\ &= 0 + \langle x, U^2(x) \rangle + \langle U(x), U(x) \rangle + 0 \\ &= \langle U^*(x), U(x) \rangle + \langle U(x), U(x) \rangle = 2\|U(x)\|^2 = 0 \quad \forall x \in V. \quad \square \end{aligned}$$

Theorem: Let T linear operator on fin-dim. IPS V over \mathbb{R} or \mathbb{C} ; then TFAE:

(i) $T \circ T^* = I$

(ii) $T^* \circ T = I$

(iii) $\langle T(x), T(y) \rangle = \langle x, y \rangle \quad \forall x, y \in V$ [T preserves the inner product]

(iv) If B is an orthonormal basis for V , then $T(B)$ is an orthonormal basis for V

(v) \exists orthonormal basis B for V s.t. $T(B)$ an orthonormal basis for V

(vi) $\|T(x)\| = \|x\| \quad \forall x \in V$, i.e. T unitary/orthogonal [T preserves the norm]

(*) Proof: (i) \Rightarrow (ii) / (ii) \Rightarrow (i): Follows from the fact that $[T][T^*]_B = I \Rightarrow [T^*]_B = [T]^{-1}_B$.

(ii) \Rightarrow (iii): $T^* \circ T = I \Rightarrow \langle x, y \rangle = \langle T^* \circ T(x), y \rangle = \langle T(x), T(y) \rangle$

(iii) \Rightarrow (iv): $T(B)$ orthonormal per (iii) [$\langle x, y \rangle = 0 \Rightarrow \langle T(x), T(y) \rangle = 0$], has $\dim V$ -many elements

(iv) \Rightarrow (v): \exists orthonormal basis B for V ; then $T(B)$ also an orthonormal basis for V

(v) \Rightarrow (vi): Let $B = \{v_1, \dots, v_n\}$, $x = a_1 v_1 + \dots + a_n v_n \in V$; then:

$$\|T(x)\|^2 = \langle T(x), T(x) \rangle = \left\langle \sum_{i=1}^n a_i T(v_i), \sum_{i=1}^n a_i T(v_i) \right\rangle = \sum_{i=1}^n a_i \bar{a}_i = \sum_{i=1}^n |a_i|^2 = \left\langle \sum_{i=1}^n a_i v_i, \sum_{i=1}^n a_i v_i \right\rangle = \|x\|^2.$$

(vi) \Rightarrow (ii): $\langle x, y \rangle = \langle T(x), T(y) \rangle = \langle x, T^* \circ T(y) \rangle \Rightarrow \langle x, (I - T^* \circ T)(y) \rangle = 0$

$U = I - T^* \circ T$ self-adjoint [$\langle (I - T^* \circ T)(x), y \rangle = \langle x, (I - T^* \circ T)(y) \rangle \Rightarrow I - T^* \circ T = 0$ transformation]. □

4/20/24

Unitary & Orthogonal Matrices

Lecture 11

+ Lecture 12

Def: A square matrix $A \in M_{n \times n}(F)$ is called: (i) orthogonal if $A^T A = A A^T = I_n$.
(ii) unitary if $A^* A = A A^* = I_n$.

Prop. The following are equivalent:

- (i) $A \in M_{n \times n}(\mathbb{R})$ is an orthogonal matrix
- (ii) The rows of A form an orthonormal basis for \mathbb{R}^n
- (iii) The columns of A form an orthonormal basis for \mathbb{R}^n

(*) Can write similar for
 $A \in M_{n \times n}(\mathbb{C})$ unitary,
rows/columns orthonormal basis
for \mathbb{C}^n

(*) Proof: (i) \Rightarrow (ii): Let A orthogonal:

$$\rightarrow s_{ij} = (A A^T)_{ij} = \sum_{k=1}^n A_{ik} (A^T)_{kj} = \sum_{k=1}^n A_{ik} A_{jk} \Rightarrow \text{rows of } A \text{ are orthonormal}$$

A has $n = \dim(\mathbb{R}^n)$ rows, therefore rows of A comprise an orthonormal basis.

(ii) \Rightarrow (i): Via opposite calc. as above, rows of A orthogonal $\Rightarrow (A A^T)_{ij} = \delta_{ij} \Rightarrow A A^T = I_n$.

(*) (i) \Rightarrow (iii), (iii) \Rightarrow (i) via similar arguments [using $A^* A$]. □

Def: A real or complex matrix $A \in M_{n \times n}(F)$ is called normal if A commutes with A^* , i.e.:

$$A^* A = A A^*$$

(*) Note: If A real symmetric [$A = A^T$] or Hermitian [$A = A^*$], then A is normal

(*) Theorem: A matrix $A \in M_{n \times n}(F)$ is normal iff \exists an orthonormal basis for F^n consisting of eigenvectors of A .

Unitary Equivalence

4/29/24

Lecture 12

Def: Let $A \in M_{n \times n}(F)$ real or complex. Then A is called orthogonally [unitarily] equivalent to a matrix $D \in M_{n \times n}(F)$ if \exists an orthogonal [unitary] matrix P such that:

$$D = P^* A P$$

Theorem: (i) Let $A \in M_{n \times n}(\mathbb{C})$; then A is normal iff A is unitarily equivalent to a diagonal matrix.

(ii) Let $A \in M_{n \times n}(\mathbb{R})$; then A is symmetric iff A is orthogonally equivalent to a real diagonal matrix.

(*) Proof [just (i)]:

(\Rightarrow) Let A normal [$A^* A = A A^*$]. Per prev. theorem, A normal $\Rightarrow \exists$ orthonormal basis for \mathbb{C}^n comprised of eigenvectors of A . Let $B = \{v_1, \dots, v_n\}$ orthonormal basis, where v_1, \dots, v_n are eigenvectors w/ eigenvalues $\lambda_1, \dots, \lambda_n$

$$\Rightarrow A \text{ unitarily equivalent to } D = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix}$$

(\Leftarrow) Let $A = P^* D P$ [D diagonal, P unitary]; then:

$$A^* A = (P^* D P)^* (P^* D P) = P^* D^* P P^* D P = P^* D^* \underbrace{D P}_{P \text{ unitary} \Rightarrow P P^* = I} = P^* D^* P = A A^*$$

D diagonal $\Rightarrow D = D^*$

(*) Note: To orthogonally/unitarily diagonalize a matrix:

(i) Find the eigenvalues & corresponding eigenvectors [forming a basis for V]

(ii) Orthogonalize the set of eigenvectors:

(a) Within each eigenspace, orthogonalize its basis of eigenvectors (via Gram-Schmidt, e.g.)

(b) Normalize all eigenvectors, i.e. divide each eigenvector by its norm

(iii) Set D, P as given in above proof

Rigid Motions

5/3/24

Lecture 13

Recall: Let $V = \mathbb{R}^2$ and $\theta \in \mathbb{R}$. Have the following matrices:

(i) Rotation by θ about the origin:
 [counter-clockwise] $[T]_B = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ [Rotation matrix]
 $\det(T) = 1$

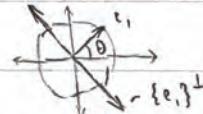
(ii) Reflection about a line θ away from
 the positive x-axis: $[T]_B = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$ [Reflection matrix]
 $\det(T) = -1$

Theorem: Let T an orthogonal operator on \mathbb{R}^2 and $A = [T]_B$, w/ B standard basis for \mathbb{R}^2 . Then exactly one of the following is true:

- (i) Either T is a rotation and $\det(A) = 1$, or
- (ii) T is a reflection [about a line through the origin] and $\det(A) = -1$.

(*) Proof: Since T orthogonal, \exists orthonormal basis $\{T(e_1), T(e_2)\}$ for \mathbb{R}^2 [by past thm.]

Since $\|T(e_1)\| = 1$, then $T(e_1) = (\cos \theta, \sin \theta)$ for some $\theta \in [0, 2\pi)$.



Since $T(e_1) \perp T(e_2)$, have 2 possibilities for $T(e_2)$:

$$\text{(i)} \quad T(e_2) = (-\sin \theta, \cos \theta) \quad \text{(ii)} \quad T(e_2) = (\cos(\theta + \pi/2), \sin(\theta + \pi/2))$$

(i) $T(e_2) = (-\sin \theta, \cos \theta) \Rightarrow T$ a rotation, or

(ii) $T(e_2) = (\cos(\theta + \pi/2), \sin(\theta + \pi/2)) \Rightarrow T$ a reflection.

◻

Def: Let V real IPS. A function $f: V \rightarrow V$ is called a rigid motion if, $\forall x, y \in V$:

$$\|f(x) - f(y)\| = \|x - y\|$$

(*) Note: Intuitively, the norm of a difference of two vectors represents the distance between them;
 thus a rigid motion is a function that preserves distances between elements

(*) Note: Any orthogonal operator on V is a rigid motion.

Rigid Motions (cont.)

5/3/24

Lecture 13
(cont.)

Theorem: Let $f: V \rightarrow V$ be a rigid motion on a fin. dim. real IPS V . Then \exists unique orthogonal operator T on V and unique translation g on V s.t. $f = g \circ T$.

Recall: A translation on V is a function $g: V \mapsto V + v_0$ for some $v_0 \in V$ fixed. Any translation is a rigid motion.

(*) Proof (Theorem: Existence)

Let $T: V \rightarrow V$ defined by $T(x) = f(x) - f(0) \quad \forall x \in V$, and let $g: V \rightarrow V$ defined by $g(x) = x + f(0)$.

Clearly, $f = g \circ T$ and g is a translation; remains to show that T is orthogonal. [Note: no assumption of T linear.]

, & a rigid motion

$$\begin{aligned} \text{For any } x \in V, \text{ have: } \|T(x)\| &= \|f(x) - f(0)\| = \|x\| \\ \Rightarrow \|T(x) - T(y)\|^2 &= \langle T(x) - T(y), T(x) - T(y) \rangle \\ &= \langle T(x), T(x) \rangle - 2\langle T(x), T(y) \rangle + \langle T(y), T(y) \rangle \\ &= \|T(x)\|^2 - 2\langle T(x), T(y) \rangle + \|T(y)\|^2. \end{aligned}$$

$$\begin{aligned} \text{Know: (i)} \|x-y\|^2 &= \langle x-y, x-y \rangle = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2, \quad (\text{ii}) \|x-y\|^2 = \|T(x-y)\|^2 = \|T(x)-T(y)\|^2 \\ \Rightarrow \langle x, y \rangle &= \langle T(x), T(y) \rangle \quad [\text{T preserves inner product}] \end{aligned}$$

$$\begin{aligned} \text{Remains to show T linear: } \|T(x+cy) - T(x) - cT(y)\|^2 &= \|\{T(x+cy) - T(x)\} - cT(y)\|^2 \\ \rightarrow &= \|T(cx+y) - T(x)\|^2 - 2c\langle T(cx+y) - T(x), T(y) \rangle + c^2\|T(y)\|^2 \\ &= \end{aligned}$$

Direct Sums

5/6/24

Lecture 14

Def: A vector space V is called the direct sum of W_1, W_2, \dots, W_k [denoted $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$]

if W_1, W_2, \dots, W_k are subspaces of V s.t.:

$$(i) W_1 + W_2 + \dots + W_k = V$$

$$(ii) W_1, W_2, \dots, W_k \text{ are disjoint, i.e. } W_i \cap \left[\sum_{j \neq i} W_j \right] = \{0_V\} \quad \forall 1 \leq i \leq k.$$

(*) Ex: Let $W_1 = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ and $W_2 = \text{span}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$; then $V = \mathbb{R}^2 = W_1 \oplus W_2$

Prop. Let V a v.s. and W_1, W_2, \dots, W_k subspaces of V . Then $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ iff
for every $v \in V$, \exists unique $w_i \in W_i$ for $1 \leq i \leq k$ s.t. $v = w_1 + w_2 + \dots + w_k$.

(*) Corollary: If V fin. dim. and $V = W_1 \oplus \dots \oplus W_k$, then $\dim(V) = \sum_{i=1}^k \dim(W_i)$.

(*) Proof (Prop): Proof for $k=2$:

(\Rightarrow) Let $V = W_1 \oplus W_2$, then $V = W_1 + W_2$; hence, for any $v \in V$, $v = w_1 + w_2$ for some $w_1 \in W_1, w_2 \in W_2$.

For uniqueness: suppose $v = \tilde{w}_1 + \tilde{w}_2$ for some $\tilde{w}_1 \in W_1, \tilde{w}_2 \in W_2$.

$$\text{Then } w_1 + w_2 = \tilde{w}_1 + \tilde{w}_2 \Rightarrow w_1 - \tilde{w}_1 = \tilde{w}_2 - w_2.$$

$w_1, \tilde{w}_1 \in W_1 \Rightarrow w_1 - \tilde{w}_1 \in W_1$, and similar for $\tilde{w}_2 - w_2 \in W_2$; then have that

$$w_1 - \tilde{w}_1 = \tilde{w}_2 - w_2 \in W_1 \cap W_2.$$

By defn. of direct sum, $W_1 \cap W_2 = \{0_V\}$; then $w_1 - \tilde{w}_1 = \tilde{w}_2 - w_2 = 0_V$

$$\Rightarrow w_1 = \tilde{w}_1, w_2 = \tilde{w}_2.$$

(\Leftarrow) Obviously, $V \subseteq W_1 + W_2$. Since $W_1, W_2 \subseteq V$, then $W_1 + W_2 \subseteq V \Rightarrow W_1 + W_2 = V$.

Assume, for contradiction, that $W_1 \cap W_2 \neq \{0_V\}$. Then $\exists w \neq 0_V$ s.t. $w \in W_1, w \in W_2$.

Then: $0_V + 0_V = w + (-w) = 0_V \in V$, violating the uniqueness condition. \blacksquare

Projections & Orthogonal Complement

5/6/24

Lecture 14

+ Lecture 15

Def: Let V a v.s., and W_1, W_2 subspaces of V s.t. $V = W_1 \oplus W_2$. Then the function $T: V \rightarrow V$ defined by $T(v) = w_1$, where $w_1 \in W_1$ is the [unique] vector s.t. $v = w_1 + w_2$ for some $w_2 \in W_2$, is called the projection of V on W_1 along W_2 .

(*) Note: The projection of V on W_1 along W_2 is linear.

Def: Let V an IPS, and let $S \subseteq V$ be a nonempty subset of V . Then we define the orthogonal complement of S to be the set S^\perp defined by:

$$S^\perp = \{x \in V : \langle x, y \rangle = 0 \forall y \in S\}$$

Prop: Properties of the Orthogonal Complement

(i) If S is a subset of V , then its orthogonal complement S^\perp is a subspace of V

$$(ii) \{0_v\}^\perp = V; V^\perp = \{0_v\}$$

(iii) If W is a subspace of V , then $W \cap W^\perp = \{0_v\}$

(iv) If W is a fin. dim. subspace of V , then $(W^\perp)^\perp = W$.

(*) Proof: (i) Trivially, $0_v \in S^\perp$. Let $u, v \in S^\perp$ and $c \in F$; then $\forall y \in V$:

$$\langle cu+v, y \rangle = \langle cu, y \rangle + \langle cv, y \rangle = c\langle u, y \rangle + \langle v, y \rangle = 0 + 0 = 0 \Rightarrow cu+v \in S^\perp.$$

(ii) Clearly, $\langle 0_v, y \rangle = 0 \forall y \in V$. From prev. lectures: $\langle x, y \rangle = 0 \forall y \in V \Leftrightarrow x = 0_v$.

(iii) Clearly, $0_v \in W \cap W^\perp$. Let $x \in W \cap W^\perp$; then $\|x\|^2 = \langle x, x \rangle = 0 \Rightarrow x = 0_v \Rightarrow W \cap W^\perp = \{0_v\}$.

(iv) Let $w \in W$; then $\langle w, y \rangle = 0 \forall y \in W^\perp \Rightarrow w \in (W^\perp)^\perp \Rightarrow w \in (W^\perp)^\perp$.

Orthogonal Projections

5/8/24

Lecture 15

(cont.)

Prop. Let V an IPS, W a finite-dimensional subspace of V . Then we have that:

$$V = W \oplus W^\perp$$

(*) Proof: Want to show $V = W + W^\perp$. Let $\{v_1, \dots, v_k\}$ orthonormal basis for W , and let $v \in V$.

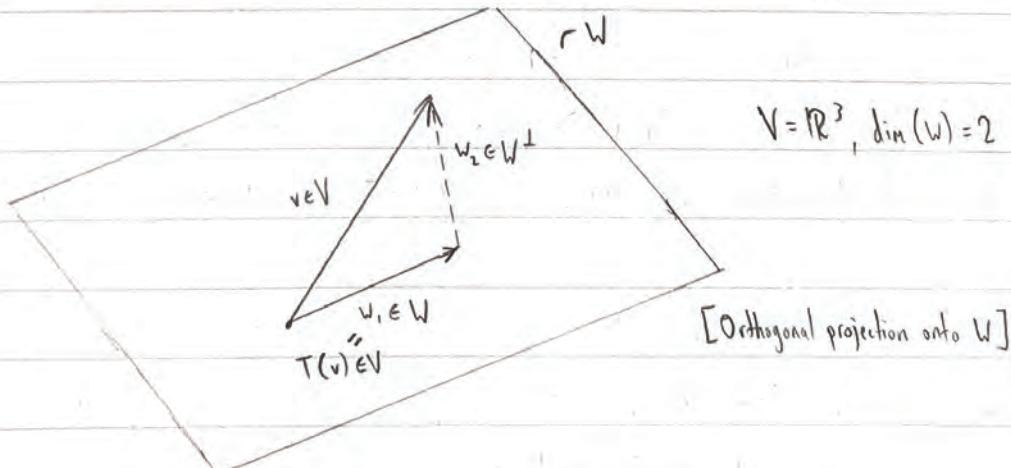
$$\text{Then have: } v = \underbrace{\langle v, v_1 \rangle v_1 + \dots + \langle v, v_k \rangle v_k}_{:= w_1 \in W} + \underbrace{v - \langle v, v_1 \rangle v_1 - \dots - \langle v, v_k \rangle v_k}_{:= w_2 \in V^\perp}$$

$$\begin{aligned} \text{Then for each } v_i, 1 \leq i \leq k: \langle w_2, v_i \rangle &= \langle v - \sum_{j=1}^k \langle v, v_j \rangle v_j, v_i \rangle = \langle v, v_i \rangle - \sum_{j=1}^k \langle v, v_j \rangle \langle v_j, v_i \rangle \\ &\rightarrow \langle v, v_i \rangle - \langle v, v_i \rangle = 0 \Rightarrow w_2 \in W^\perp \Rightarrow V = W + W^\perp. \text{ Since } W \cap W^\perp = \{0_V\}, \text{ then } V = W \oplus W^\perp. \quad \square \end{aligned}$$

Def: Let V an IPS, W fin-dim subspace of V . Then we define the orthogonal projection of V onto W [along W^\perp] to be the function $T: V \rightarrow V$ defined by $T(v) = w_1$, where $w_1 \in W$ is the unique vector $\in W$ s.t. $v = w_1 + w_2$ for some $w_2 \in W^\perp$.

(*) Note: Since T is a projection, then T is linear.

(*) Ex: Let $V = \mathbb{R}^2$, $W = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$ [$\Rightarrow W^\perp = \text{span}\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$]; then $T\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} a \\ 0 \end{pmatrix}$ for any $\begin{pmatrix} a \\ b \end{pmatrix} \in V$.



(*) Ex: Let V IPS, and let $w \in V, w \neq 0$; define $W = \text{span}\{w\}$ [1D subspace]

$$\rightarrow \text{for any } v \in V: v = \frac{\langle v, w \rangle}{\|w\|^2} w + \left(v - \frac{\langle v, w \rangle}{\|w\|^2} w\right) \rightarrow T(v) = \frac{\langle v, w \rangle}{\|w\|^2} w \text{ orthogonal projection on } W$$

Orthogonal Projections (cont.)

5/10/24

Lecture 16

Prop. Let V an IPS, W fin.-dim. subspace of V . Let T be the orthogonal projection of V onto W along W^\perp ; then:

- (i) $v - T(v) \in W^\perp \forall v \in V$
- (ii) $T(w) = w$ iff $w \in W$
- (iii) $T(u) = 0_v$ iff $u \in W^\perp$
- (iv) $R(T) = W$ and $N(T) = W^\perp$
- (v) $T^2 = T$

(vi) Let v_1, w_2, \dots, v_k orthonormal basis for W ; then:

$$T(v) = \langle v, v_1 \rangle w_1 + \langle v, v_2 \rangle w_2 + \dots + \langle v, v_k \rangle w_k$$

(*) Proof: (i) Let $v \in V$; then v has unique decomposition $v = w_1 + w_2$ w/ $w_1 \in W, w_2 \in W^\perp$.

Then $T(v) = w_1$, [by defn.] $\Rightarrow v - T(v) = v - w_1 = w_2 \in W^\perp$.

(ii) For any $w \in W$, have $w = \overset{W}{w} + \overset{W^\perp}{0}_v \Rightarrow T(w) = w$. Conversely, if $w \in V$ has $T(w) = w$, then $T(w) \in W \Rightarrow w \in W$.

(iii) For any $u \in W^\perp$, have $u = \overset{W}{0}_v + \overset{W^\perp}{u} \Rightarrow T(u) = 0_v$. Conversely, if $u \in V$ has $T(u) = 0_v$, then per (i): $u - T(u) = u - 0_v \in W^\perp \Rightarrow u \in W^\perp$.

(iv) Clearly, $R(T) \subseteq W$. Let $w \in W$; then $T(w + 0_v) = w \Rightarrow W \subseteq R(T) \Rightarrow R(T) = W$.

Per (iii), for any $u \in V$, $T(u) = 0_v$ iff $u \in W^\perp$; hence $W^\perp = N(T)$.

(v) Let $v \in V$ w/ $v = \overset{W}{w} + \overset{W^\perp}{v}$; then $T^2(v) = T(w)$; $w \in W \xrightarrow{(iv)} T^2(v) = w = T(v) \Rightarrow T^2 = T$.

(vi) Proven previously □

Prop. Let V fin.-dim. IPS, and let T be a projection. Then T is an orthogonal projection on some subspace W iff $R(T)^\perp = N(T)$.

(*) Proof [\Rightarrow]: T orthogonal projection $\Rightarrow R(T) = W, N(T) = W^\perp \Rightarrow R(T)^\perp = N(T)$. □

Orthogonal Projections (cont.)

5/10/24

Lecture 16

Thm: Let V fin.-dim. IDS and T linear operator on V . Then T is an orthogonal projection on some subspace W iff $T^2 = T = T^*$.

(*) Proof: (\Rightarrow) Let T an orthogonal projection. From before, have that $T^2 = T$; remains to show that $T = T^*$.

Let $x, y \in V$; can write $x = x_1 + x_2$, $y = y_1 + y_2$ for some $x_1, y_1 \in U$; $x_2, y_2 \in U^\perp$.

Hence:

$$(i) \quad \langle x, T(y) \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_1 \rangle = \langle x_1, y_1 \rangle$$

$$(ii) \quad \langle T(x), y \rangle = \langle x_1 + x_2, y_1 + y_2 \rangle = \langle x_1, y_1 \rangle + \underset{\parallel}{\langle x_2, y_1 \rangle} + \langle x_2, y_2 \rangle = \langle x_1, y_1 \rangle$$

$$\rightarrow \langle x, T(y) \rangle = \langle T(x), y \rangle \quad \forall x, y \in V \Rightarrow T^* = T.$$

(\Leftarrow) Lemma: Let T linear $V \rightarrow V$; then $T^2 = T$ iff T is a projection on $U_1 = \{y : T(y) = y\}$ along $N(T)$. [proof in HW]

Assume $T^2 = T = T^*$; by the lemma, T is a projection on $U_1 = \{y : T(y) = y\}$ along $N(T)$.

To show T an orthogonal projection, want to show that $R(T)^\perp = N(T)$.

First, want to show that $R(T) = U_1$. Obviously, $U_1 \subseteq R(T)$. Let $y \in R(T)$; then $\exists x \in V$ s.t. $T(x) = y$. Since $T^2 = T$: $T(y) = T(T(x)) = T(x) = y \Rightarrow y \in U_1$, therefore $R(T) \supseteq U_1$.

Since $U_1 = R(T)$, then T is a projection on $R(T)$ along $N(T)$. Let $x \in R(T)$, $y \in N(T)$; using that $T = T^*$, $x - T(x)$: $\langle x, y \rangle = \langle T(x), y \rangle = \langle x, T(y) \rangle = \langle x, 0 \rangle = 0 \Rightarrow y \in R(T)^\perp \Rightarrow N(T) \subseteq R(T)^\perp$.

Let $y \in R(T)^\perp$. Let $y_1 \in R(T)$, $y_2 \in N(T)$ s.t. $y = y_1 + y_2$. Then $\forall x \in R(T)$:

$$\rightarrow \langle x, y \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle = \langle x, T(y_1) \rangle + \langle x, T(y_2) \rangle = \langle x, y_1 \rangle. \text{ Since } y \in R(T)^\perp,$$

then $\langle x, y \rangle = \langle x, y_1 \rangle = 0 \quad \forall x \in R(T)$; hence $y_1 = 0_V \Rightarrow y = y_2 \Rightarrow y \in N(T)$, hence $R(T)^\perp \subseteq N(T)$. \square

The Spectral Theorem

5/13/24

Lecture 17

The Spectral Theorem/Spectral Decomposition

Theorem: Let T be a linear operator on a fin.-dim. IPS V over F with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k \in F$. Additionally, assume that T is normal [$F = \mathbb{C}$] / self-adjoint [$F = \mathbb{R}$].

For each $1 \leq i \leq k$, let W_i be the eigenspace corresponding to eigenvalue λ_i , and let T_i be the orthogonal projection of V onto W_i . Then:

$$(i) V = W_1 \oplus W_2 \oplus \dots \oplus W_k$$

(ii) For each W_i , $W_i^\perp = \bigoplus_{j \neq i} W_j$ [the direct sum of the subspaces $W_{j \neq i}$]

$$(iii) T_i T_j = \delta_{ij} T_i \text{ for } 1 \leq i, j \leq k$$

$$(iv) I = T_1 + T_2 + \dots + T_k$$

$$(v) \boxed{T = \lambda_1 T_1 + \lambda_2 T_2 + \dots + \lambda_k T_k} \quad [\text{Spectral decomposition}]$$

Recall: Given a linear operator T with eigenvalue λ , the eigenspace corresponding to λ is the subspace $E_\lambda = \{v \in V : T(v) = \lambda v\}$ of all eigenvectors with associated eigenvalue λ .

(*) Proof (The Spectral Theorem)

(i) Trivially, $W_i \cap \bigcup_{j \neq i} W_j = \{0\}$ and $W_1 + W_2 + \dots + W_k \subseteq V$. Since T is normal, then \exists an orthonormal basis for V consisting of eigenvectors of T , hence $W_1 + W_2 + \dots + W_k \subseteq V$.

(ii) Fix $1 \leq i, j \leq k$, $i \neq j$. Let $x_i \in W_i$, $x_j \in W_j$. $[T(x_i) = \lambda_i x_i ; T(x_j) = \lambda_j x_j]$. Then:
 $\lambda_i \langle x_i, x_j \rangle = \langle T(x_i), x_j \rangle = \langle x_i, T^*(x_j) \rangle$
 $\rightarrow = \langle x_i, \overline{T(x_j)} \rangle$ [T self-adjoint] $= \lambda_j \langle x_i, x_j \rangle$

Since $\lambda_i \neq \lambda_j$, then $\langle x_i, x_j \rangle = 0$; holds for any $j \neq i$

~ By linearity: $\langle x_i, y \rangle = 0 \quad \forall y \in \bigoplus_{j \neq i} W_j$.

(iii) For any $1 \leq i, j \leq k$: if $i \neq j$, then $W_i \perp W_j \Rightarrow T_i \circ T_j = T_0$. Otherwise, $T_i^2 = T_i$ [orthogonal projection].

5/13/24

The Spectral Theorem (cont.)

Lecture 17 (4) Proof (Spectral Theorem, cont.)

(cont) (iv) Per (ii): since T_i orthogonal projection, have that $N(T_i) = W_i^\perp = \bigoplus_{j \neq i} W_j$ for each $1 \leq i \leq k$.Per (i): since $W_1 \oplus W_2 \oplus \dots \oplus W_k = V$, then for every $v \in V \exists$ unique vectors $v_i \in W_i$, $v_1 \in W_1, \dots, v_k \in W_k$ s.t. $v = v_1 + v_2 + \dots + v_k$.For each $1 \leq i \leq k$:

$$\begin{aligned} & \text{For each } 1 \leq i \leq k: \\ & \quad \text{Since } W_i \perp \text{span}(W_1, \dots, \hat{W}_i, \dots, W_k) \Rightarrow T_i(v) = v. \end{aligned}$$

Hence:

$$\begin{aligned} (T_1 + T_2 + \dots + T_k)(v) &= (T_1 + T_2 + \dots + T_k)(v_1 + v_2 + \dots + v_k) \\ &= v_1 + v_2 + \dots + v_k = v \quad [= I(v)] \end{aligned}$$

(v) Let $v \in V$. As in (iv), have that \exists unique $v_1 \in W_1, v_2 \in W_2, \dots, v_k \in W_k$ such that $v = v_1 + v_2 + \dots + v_k$. [Recall: $v_i \in W_i \Rightarrow T(v_i) = \lambda_i v_i$]

Then:

$$\begin{aligned} T(v) &= T(v_1 + v_2 + \dots + v_k) \\ &= T(v_1) + T(v_2) + \dots + T(v_k) \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k \\ &= \lambda_1 T_1(v) + \lambda_2 T_2(v) + \dots + \lambda_k T_k(v) \\ &= (\lambda_1 T_1 + \lambda_2 T_2 + \dots + \lambda_k T_k)(v) \end{aligned}$$

$$\rightarrow T = \lambda_1 T_1 + \lambda_2 T_2 + \dots + \lambda_k T_k.$$



(*) Note: The spectral decomposition allows for decomposing any normal/self-adjoint operator into [much simpler] orthogonal projections \Rightarrow many uses & applications.

The Spectral Theorem, Corollaries

5/15/24

Lecture 18

Corollary: Let T be a normal $[F = \mathbb{C}]$ / self-adjoint operator, and let g be a polynomial.

Let $T = \lambda_1 T_1 + \lambda_2 T_2 + \dots + \lambda_k T_k$ be the spectral decomposition of T ; then:

$$g(T) = g(\lambda_1)T_1 + g(\lambda_2)T_2 + \dots + g(\lambda_k)T_k$$

(*) Ex: Let $A \in M_{n \times n}(\mathbb{R})$ symmetric \Rightarrow orthogonally equivalent to a real diagonal matrix]. Can compute its powers A^p [$p \in \mathbb{N}$] using its spectral decomposition.

Since A symmetric, then A self-adjoint $\Rightarrow \exists$ orthonormal basis for \mathbb{R}^n consisting of eigenvectors of A . Let W_1, \dots, W_k eigenspaces of A corresponding to [distinct] eigenvalues $\lambda_1, \dots, \lambda_k$, and let $m_i = \dim(W_i)$ for $1 \leq i \leq k$. Let B be an orthonormal basis for \mathbb{R}^n of eigenvectors.

$$\rightarrow [T]_B = \left[\sum_{i=1}^k \lambda_i T_i \right]_B = \sum_{i=1}^k \lambda_i [T_i]_B$$

For each T_i , have: $[T_i]_B = \begin{pmatrix} 0 & & & \\ & \ddots & & 0 \\ & & I_{m_i} & \\ & & & \ddots & 0 \end{pmatrix}$

$$\Rightarrow \sum_{i=1}^k \lambda_i [T_i]_B = \boxed{[T]_B = \begin{pmatrix} \lambda_1 I_{m_1} & & & \\ & \lambda_2 I_{m_2} & & \\ & & \ddots & \\ & & & \lambda_k I_{m_k} \end{pmatrix}} \Rightarrow [T^p]_B = \boxed{\begin{pmatrix} \lambda_1^p I_{m_1} & & & \\ & \lambda_2^p I_{m_2} & & \\ & & \ddots & \\ & & & \lambda_k^p I_{m_k} \end{pmatrix}}$$

Corollary: If $F = \mathbb{C}$, then T is normal iff, for some polynomial g :

$$T^* = g(T)$$

5/15/24

lecture 18

The Spectral Theorem, Corollaries (cont.)

Corollary (restated): If $F = \mathbb{C}$, then T is normal iff $T^* = g(T)$ for some polynomial g .

(*) Proof:

(\Rightarrow) Suppose T is normal, i.e. $T^*T = TT^*$. Let $T = \sum_{i=1}^k \lambda_i T_i$ spectral decomposition for T ; then we have that:

$$T^* = \sum_{i=1}^k (\lambda_i T_i)^* = \sum_{i=1}^k \bar{\lambda}_i T_i \quad \text{(*) } T, \text{orthogonal proj.} \Rightarrow T_i \text{ self-adjoint}$$

Can find polynomial g s.t. $g(\lambda_i) = \bar{\lambda}_i \quad \forall 1 \leq i \leq k$; e.g. define:

$$g_i(x) = \frac{(x-\lambda_1)(x-\lambda_2)\dots(x-\lambda_{i-1})(x-\lambda_{i+1})\dots(x-\lambda_k)}{(\lambda_i-\lambda_1)(\lambda_i-\lambda_2)\dots(\lambda_i-\lambda_{i-1})(\lambda_{i+1}-\lambda_{i+1})\dots(\lambda_i-\lambda_k)} \quad [g_i(\lambda_j) = \delta_{ij}]$$

$$g(x) = \sum_{i=1}^k \bar{\lambda}_i g_i(x)$$

$$\Rightarrow g(T) = \sum_{i=1}^k g(\lambda_i) T_i = \sum_{i=1}^k \bar{\lambda}_i T_i = T^*$$

(\Leftarrow) Suppose $T^* = g(T)$ for some polynomial g ; then:

$$TT^* = Tg(T) = g(T)T = T^*T \Rightarrow T \text{ normal.} \quad \square$$

Corollary: If $F = \mathbb{C}$, then T is self-adjoint iff T is normal and all eigenvalues λ_i of T are real-valued [$\lambda_i \in \mathbb{R}$].

Corollary: Let T be a linear operator with spectral decomposition $T = \sum_{i=1}^k \lambda_i T_i$. Then $\forall 1 \leq i \leq k$, have that $T_i = g_i(T)$ for some polynomial g_i .

(*) Proof ($T_i = g_i(T)$): Take g_i as above, s.t. $g_i(\lambda_j) = \delta_{ij}$. Then $\forall 1 \leq i \leq k$:

$$g_i(T) = \sum_{j=1}^k g_i(\lambda_j) T_j = T_i$$

The Geometry of Orthogonal Operators

5/20/24

Can show that any orthogonal operator on a fin.-dim. real IPS is describable in terms of rotations and reflections, starting with $\dim V = 1$ & $\dim V = 2$:

Lecture 19

One-Dimensional Vector Spaces

Def: A linear operator T on a 1D real IPS is called:

- (i) A rotation if $T(x) = x \quad \forall x \in V$, and
- (ii) A reflection if $T(x) = -x \quad \forall x \in V$.

Prop: Any orthogonal operator on a real IPS V is either a reflection or a rotation.

Two-Dimensional Vector Spaces

Recall: Let T be a linear operator on a 2D real IPS. Then we say that T :

- (i) ... is a rotation if \exists orthonormal basis B for V and $\theta \in \mathbb{R}$ such that:

$$[T]_B = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad [\det(T) = 1]$$

- (ii) ... is a reflection if \exists orthonormal basis B for V and $\theta \in \mathbb{R}$ such that:

$$[T]_B = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \quad [\det(T) = -1]$$

Prop: A linear operator T on 2D real IPS is a reflection iff \exists a one-dimensional subspace

W of V such that: (i) $T(x) = -x \quad \forall x \in W$

(ii) $T(v) = v \quad \forall v \notin W$

5/20/24

Lecture 19
(cont.)

The Geometry of Orthogonal Operators (cont.)

Theorem: Let T be an orthogonal operator on a 2D real IPS V . Then either (i) T is a rotation and $\det(T) = 1$, or (ii) T is a reflection and $\det(T) = -1$.

Looking at vector spaces of arbitrary dimension:

Lemma: Let T be a linear operator on nonzero fin.-dim. real IPS V . Then \exists a T -invariant subspace W of V s.t. $1 \leq \dim W \leq 2$.



Theorem: Let T be a linear operator on a nonzero fin.-dim. real IPS V . Then \exists pairwise orthogonal T -invariant subspaces W_1, W_2, \dots, W_m of V such that:

(i) $1 \leq \dim W_i \leq 2$ for each $i=1, 2, \dots, m$

(ii) $V = W_1 \oplus W_2 \oplus \dots \oplus W_m$

(*) Proof of Thm.: Via induction on $\dim(V)$. Base case: $\dim(V)=1 \Rightarrow$ trivial.

Suppose claim holds $\forall \dim(V) \leq n-1$. Let $\dim(V) = n$. By the lemma, \exists a T -invariant subspace W_1 of V s.t. $1 \leq \dim W_1 \leq V$. Since W_1 is a T -invariant subspace, then $W_1 \oplus W_1^\perp = V$ and W_1^\perp is T -invariant [using that T is orthogonal]; furthermore, the restriction $T_{W_1^\perp}$ of T to W_1^\perp is an orthogonal operator on W_1^\perp . Then per the inductive assumption, \exists pairwise orthogonal T -invariant subspaces W_2, W_3, \dots, W_m of $W_1^\perp \subseteq V$ s.t. $W_2 \oplus \dots \oplus W_m = W_1^\perp$ and $\dim(W_i) = 1$ or 2 for each $i=2, 3, \dots, m$.

Since $W_2, \dots, W_m \subseteq W_1^\perp$, then W_1 and W_i are pairwise orthogonal for each $i=2, 3, \dots, m$; hence W_1, W_2, \dots, W_m are pairwise orthogonal T -invariant subspaces satisfying (i) $[W_1 \oplus W_1^\perp = W_1 \oplus W_2 \oplus \dots \oplus W_m = V]$ and (ii). \blacksquare

Remark: We have that (i) any orthogonal operator on a 1D/2D real IPS is either a rotation or a reflection, and (ii) that any fin.-dim. real IPS can be decomposed into 1D/2D pairwise orthogonal T -invariant subspaces W_i . Since the restriction of T to W_i is either a rotation or a reflection [by (i)], we find that T itself is composed of rotations & reflections [on the subspaces W_i].

The Jordan Canonical Form

5/22/24

Lecture 20

Recall: Let T be a linear operator over a fin-dim V.S. V with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$. Then T is diagonalizable iff, for every $i=1, 2, \dots, k$, the multiplicity of λ_i is equal to $\dim(E\lambda_i)$.

→ Q: What if T is not diagonalizable?

(i) Currently, have Schur's lemma: for any $A \in M_{n \times n}(\mathbb{C})$, \exists Q unitary, U upper-triangular $[Q, U \in M_{n \times n}(\mathbb{C})]$ s.t. $A = QUQ^{-1}$.

(ii) Even better: Jordan canonical form:

$$A \sim \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ 0 & & \ddots & & J_k \end{pmatrix}, \quad J_i = \begin{pmatrix} \lambda_i & & 0 & & \\ 0 & \lambda_i & & & \\ & & \ddots & & \\ 0 & & & \lambda_i & \\ & & & 0 & \lambda_i \end{pmatrix} \text{ called a } \underline{\text{Jordan block}}$$

Formally:

Theorem: Let $A \in M_{n \times n}(F)$ s.t. the characteristic polynomial of A splits [into linear factors].

Then A is similar to the block diagonal matrix

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots & & J_k \end{pmatrix} \in M_{n \times n}(F)$$

where each block J_i is of the form

$$J_i = \begin{pmatrix} \lambda_i & 0 & \dots & 0 \\ 0 & \lambda_i & \dots & 0 \\ & & \ddots & \\ 0 & & & \lambda_i \end{pmatrix} \text{ [upper-triangular]}$$

and $\lambda_1, \lambda_2, \dots, \lambda_n$ are eigenvalues of A [not necessarily distinct].

Recall: Say that a characteristic polynomial $f(t)$ splits over a field F if \exists scalars $c, \lambda_1, \lambda_2, \dots, \lambda_n \in F$ [not necessarily distinct] s.t. $f(t) = c(t-\lambda_1)(t-\lambda_2) \dots (t-\lambda_n)$. In particular, every char. poly. for $A \in M_{n \times n}(\mathbb{C})$ splits.

5/22/24

The Jordan Canonical Form (cont)

Lecture 20

Prop: Let $B \in M_{k \times k}(\mathbb{C})$ be a Jordan block with diagonal element λ . Then:

- (i) B has only one eigenvalue λ with multiplicity k .
- (ii) $\dim(E_\lambda) = 1$, [note: implies B not diagonalizable unless $k=1$]
- (iii) $(B - \lambda I_k)^p \neq 0_{k \times k}$ if $1 \leq p < k$, but $(B - \lambda I_k)^p = 0_{k \times k} \forall p \geq k$.

(*) Proof of Prop:

(i) Since B is of the form

$$B = \begin{pmatrix} \lambda & & & \\ & \ddots & & \\ & & \lambda & \\ & & & \ddots & \lambda \end{pmatrix} \quad [\text{Upper-triangular}]$$

then B has characteristic polynomial $f(t) = \det(B - tI_k) = (-1)^k(t - \lambda)^k \Rightarrow \lambda$ is the only eigenvalue.

(ii) Notice that:

$$B - \lambda I_k = \begin{pmatrix} 0 & 1 & & \\ 0 & 0 & 1 & \\ & \ddots & \ddots & 0 \\ & & 0 & 0 \end{pmatrix}$$

By definition: $E_\lambda = N(B - \lambda I_k) \Rightarrow E_\lambda = \text{span}\{(1, 0, 0, \dots, 0)\}, \dim(E_\lambda) = 1$.

(iii) Looking at the form of $(B - \lambda I_k)$ from (ii), see that each instance of $(B - \lambda I_k)$ "shifts" each ↴ up one row [alt: right one column]

→ before $p=k$, at least one 1 will remain.

After $p=k$, all 1s will be gone \Rightarrow zero matrix.

The Jordan Canonical Form (cont.)

5/24/24

Lecture 21

Previously, had that some matrices were not diagonalizable because their eigenspaces were too small. To get around this, can "extend" the notion of an eigenspace - generalized eigenspaces.

Def. Let $A \in M_{n \times n}(\mathbb{C})$, and let $\lambda \in \mathbb{C}$ be an eigenvalue of A . Then the generalized eigenspace of A corresponding to eigenvalue λ is defined as:

$$K_\lambda = \left\{ x \in \mathbb{C}^n : (A - \lambda I_n)^p(x) = 0 \text{ for some } p \in \mathbb{N} \right\}$$

A vector $x \in \mathbb{C}^n$ is called a generalized eigenvector of A corresponding to eigenvalue λ if $x \neq 0$, and $x \in K_\lambda$. [x is contained in the generalized eigenspace]

Observation: For every eigenvalue λ of A , K_λ is a subspace of \mathbb{C}^n with $E_\lambda \subseteq K_\lambda$.

Def: Let $A \in M_{n \times n}(\mathbb{C})$. Let $x \in \mathbb{C}^n$ be a generalized eigenvector of A corresponding to λ [$x \in K_\lambda$]. Suppose $p \geq 1$ is the smallest integer > 0 s.t. $(A - \lambda I_n)^p(x) = 0$. Then the ordered set

$$\gamma = \left\{ (A - \lambda I_n)^{p-1}(x), (A - \lambda I_n)^{p-2}(x), \dots, (A - \lambda I_n)(x), x \right\}$$

is called a cycle of generalized eigenvectors [corresponding to eigenvalue λ] of length p .

Prop: For any such cycle $\{(A - \lambda I_n)^{p-1}(x), \dots, (A - \lambda I_n)(x), x\}$, the set is linearly independent.

(*) Proof: By definition, all vectors $v \in \gamma$ are nonzero. Suppose $\exists c_0, c_1, c_2, \dots, c_p \in \mathbb{C}$ such that $\sum_{i=0}^{p-1} c_i (A - \lambda I_n)^i(x) = 0$.

$$\text{Applying } (A - \lambda I_n)^{p-1} \text{ to both sides: } \sum_{i=0}^{p-1} c_i (A - \lambda I_n)^{i+p-1}(x) = c_0 (A - \lambda I_n)^{p-1}(x) = 0 \rightarrow c_0 = 0.$$

By a similar argument, can find that c_1, c_2, c_3, \dots are all 0 [induct up to $p-1$].



5/24/24

Lecture 21

+ Discussion 8

The Jordan Decomposition Theorem

Theorem (Jordan decomposition): Let $A \in M_{n \times n}(\mathbb{C})$ with distinct eigenvalues $\lambda_1, \dots, \lambda_k$ with corresponding multiplicities m_1, \dots, m_k . Then:

$$(i) \dim(K_{\lambda_i}) = m_i \text{ for each } i=1, 2, \dots, k$$

$$(ii) \mathbb{C}^n = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \dots \oplus K_{\lambda_k}$$

(iii) Each K_{λ_i} has an ordered basis consisting of a union of cycles of generalized eigenvectors corresponding to λ_i , for each $i=1, 2, \dots, k$.

(*) Note: To find the union of cycles for some eigenvalue λ , do the following:

(i) Let $\dim(V_{\lambda}) = r$. Can compute $\text{null}((A-\lambda I)^i) = r_i$ for each $i=1, 2, \dots$ until eventually $r_{N+1} = r_N \Rightarrow r_N = r$. [To show that N exists: $r_1 \geq r_2 \geq \dots \geq r_N \leq r$]

(ii) The value N [from (i)] corresponds to the length of the longest cycle. Define $s_i = r_i$, $s_2 = r_2 - r_1$, $s_3 = r_3 - r_2, \dots, s_N = r_N - r_{N-1}$. Then for each $k=1, 2, \dots, N$, the no. of cycles of length at least k is given by s_k .

(iii) Once the no. of cycles & corresponding lengths have been determined, can solve for the generalized eigenvectors:

v a [gen.] eigenvector \rightarrow solve for $v' \in V$ s.t. $(A - \lambda I_n)v' = v$

(*) Note: Let $\beta = \{v_1, \dots, v_m\}$ cycle of generalized eigenvectors

$$\rightarrow \beta = \left\{ \underbrace{v_1}_{T}, \underbrace{v_2}_{T}, \dots, \underbrace{v_{m-1}}_{T}, \underbrace{v_m}_{T} \right\}, T(v_i) = \begin{cases} \lambda v_i + v_{i-1} & i \neq 1 \\ \lambda v_i & i=1 \end{cases}$$

The Jordan Decomposition Theorem (cont.)

5/29/24

Lecture 22

+ Lecture 23

The Jordan Canonical Form

Let $A \in M_{n \times n}(\mathbb{C})$. Let $\beta = \bigcup_{i=1}^k \beta_i$, where each $\beta_i = \{v_1^{(i)}, \dots, v_{m_i}^{(i)}\}$ is the union of cycles for K_{λ_i} ; call β the Jordan basis. Then we have that:

$$A = QJQ^{-1} \quad [\text{Jordan form}]$$

w/ Q, J defined by:

$$Q = \begin{pmatrix} & & \\ \uparrow & \uparrow & \uparrow \\ v_1^{(1)} & v_2^{(1)} & \dots & v_{m_1}^{(1)} \\ & \downarrow & & \downarrow \\ & & & \end{pmatrix} \quad \text{and} \quad J = \begin{pmatrix} J_{\lambda_1} & & \\ & J_{\lambda_2} & \dots \\ & & J_{\lambda_k} \end{pmatrix} \quad \left[J_{\lambda_i} = \begin{pmatrix} \lambda_i & & & \\ & \ddots & & \\ & & \lambda_i & \\ & & & \lambda_i \end{pmatrix} \in M_{m_i \times m_i}(\mathbb{C}) \right]$$

To prove Jordan decomposition theorem, first want to prove 2 lemmas:

Lemma (1): Let $A \in M_{n \times n}(\mathbb{C})$, λ eigenvalue of A w/ corresponding gen. eigenspace K_λ . Then:

- (i) K_λ is T -invariant under the operator $T = L_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$.
- (ii) For any scalar $\mu \neq \lambda$, the operator $L_{A-\mu I_n}: K_\lambda \rightarrow K_\lambda$ is injective.
- (iii) For any eigenvalue μ of A such that $\mu \neq \lambda$, have that:

$$K_\mu \cap K_\lambda = \{0_v\}$$

(*) Proof of Lemma:

- (i) Sufficient to show that $L_A(K_\lambda) \subseteq K_\lambda$. Let $x \in K_\lambda$; then by definition of K_λ , $\exists p \geq 1, p \in \mathbb{N}$ s.t. $(A - \lambda I)^p(x) = 0_v$. Then:

$$(A - \lambda I)^p(L_A(x)) = (A - \lambda I)(A(x)) = A(A - \lambda I)^p(x) = A(0_v) = 0_v$$

Hence $x \in K_\lambda$ by definition of K_λ .

5/31/24

The Jordan Decomposition Theorem (cont)

Lecture 23

(i) Proof of Lemma (cont.)

+ Lecture 24

(ii) Assume, for the sake of contradiction, that $L_A - \mu I_n$ is not injective. Then $\exists x \in K_\lambda, x \neq 0$, s.t.

$(A - \mu I_n)(x) = 0$. Since $x \in K_\lambda$, \exists smallest integer $p \geq 1$ s.t. $(A - \lambda I_n)^p(x) = 0$. Let $y = (A - \lambda I_n)^{p-1}(x)$; then $y \neq 0$, but $y \in N(A - \lambda I_n)$. We also find that $y \in N(A - \mu I_n)$:

$$(A - \mu I_n)(y) = (A - \mu I_n)(A - \lambda I_n)^{p-1}(x) = (A - \lambda I_n)^{p-1}(A - \mu I_n)(x) = 0$$

Then $Ay = \lambda y = \mu y$, but $y \neq 0$ and we assumed $\lambda \neq \mu$. [contradiction]

(iii) Via similar argument as in (ii). ◻

Lemma (2): Let $A \in M_{n \times n}(\mathbb{C})$, λ eigenvalue of A w/ multiplicity m . Then:

$$(i) \dim(K_\lambda) \leq m$$

$$(ii) K_\lambda = N((A - \lambda I_n)^m).$$

(i) Proof of Lemma:

(i) Let $W = K_\lambda$; by Lemma (1), W is T -invariant for $T = L_A$. Define $\delta = \dim(W)$. We notice that the operator $T_W: W \rightarrow W$ has eigenvalue λ : $T_W(x) = T(x) = \lambda x$ \forall eigenvectors $x \in \mathbb{C}^n$ w/ eigenvalue λ . In particular, by Lemma (1), this eigenvalue is the only eigenvalue for T_W ; then T_W has char. poly.:

$$h(t) = (-1)^\delta (t - \lambda)^\delta \quad [\text{characteristic polynomial for } T_W]$$

Know that $h(t)$ divides the characteristic polynomial $f(t)$ of T ; since λ has multiplicity m in $f(t)$, therefore $\dim(W) = \delta \leq m$.

(ii) Let $x \in N((A - \lambda I_n)^m)$; clearly, $x \in K_\lambda$ [$N((A - \lambda I_n)^m) \subseteq K_\lambda$]. Conversely, let $x \in K_\lambda = W$; by Cayley-Hamilton, $h(T_W) = (-1)^\delta (T_W - \lambda I)^\delta = 0$. In particular, $(T_W - \lambda I_n)^\delta(x) = 0$; since $m \geq \delta$, therefore $(T_W - \lambda I_n)^m(x) = 0$ [$K_\lambda \subseteq N((A - \lambda I_n)^m)$]. ◻

The Jordan Decomposition Theorem (cont.)

6/3/24

Lecture 24

(cont.)

Theorem (Jordan decomposition, restated): Let $A \in M_{n \times n}(\mathbb{C})$ and $\lambda_1, \dots, \lambda_k$ distinct eigenvalues of A with multiplicities m_1, \dots, m_k . Then:

$$(i) \dim(K_{\lambda_i}) = m_i \text{ for each } i=1, \dots, k$$

$$(ii) \mathbb{C}^n = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \dots \oplus K_{\lambda_k}$$

(iii) Each K_{λ_i} has an ordered basis consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ_i , for each $i=1, \dots, k$.

(*) Proof (Jordan decomposition)

(ii) Via induction on # eigenvalues k .

(a) Base case: $k=1 \Rightarrow A$ has characteristic polynomial $(-1)^n(t-\lambda_1)^n$. By Cayley-Hamilton, $(A-\lambda_1)^n = 0_{n \times n} \Rightarrow$ by Lemma 2: $K_{\lambda_1} = N((A-\lambda_1 I_n)^n) = N(0_{n \times n}) = \mathbb{C}^n$.

(b) Inductive step: Assume result holds for matrices with $k-1$ distinct eigenvalues, $k \geq 2$.

Let $A \in M_{n \times n}(\mathbb{C})$ with characteristic polynomial $f(t) = (-1)^n(t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$.

Let $W = R(A-\lambda_k I_n)^{m_k}$; claim that W invariant under $T = L_A$. Let $x \in \mathbb{C}^n$, $y = (A-\lambda_k I_n)^{m_k} x \in W$; then:

$$Ay = A(A-\lambda_k I_n)^{m_k} x = (A-\lambda_k I_n)^{m_k}(Ax) \in W$$

Then by Lemma 1, for each $1 \leq i \leq k$, $L_{A-\lambda_i I}: K_{\lambda_i} \rightarrow K_{\lambda_i}$ is injective; by rank-nullity, $L_{A-\lambda_i I}$ is bijective.

Let $W = R(A-\lambda_k I_n)^{m_k}$; see that $K_{\lambda_i} \subseteq W \forall 1 \leq i \leq k-1$. Additionally, notice:

$T_w: W \rightarrow W$ has eigenvalues $\lambda_1, \dots, \lambda_{k-1}$, but not λ_k . By the inductive hypothesis, $\forall x \in \mathbb{C}^n$,

\exists unique $w_i \in K_{\lambda_i} \forall 1 \leq i \leq k-1$ s.t. $(A-\lambda_k I)^{m_k} x = w_1 + w_2 + \dots + w_{k-1}$.

Since $L_{A-\lambda_k I}^{m_k}: K_{\lambda_k} \rightarrow K_{\lambda_k}$ bijective, \exists unique $v_i \in V_{\lambda_i}$ s.t. $(A-\lambda_k I)^{m_k} v_i = w_i$, $\forall 1 \leq i \leq k-1$.

$$\xrightarrow{\substack{\text{Since } L_{A-\lambda_k I}^{m_k} \text{ is bijective, } \exists \text{ unique } v_i \in V_{\lambda_i} \text{ s.t. } (A-\lambda_k I)^{m_k} v_i = w_i \\ \text{and } (A-\lambda_k I)^{m_k} (\underbrace{v_1 + v_2 + \dots + v_{k-1}}_{=v_k \in V_{\lambda_k}}) = 0}} x = v_1 + v_2 + \dots + v_{k-1} + v_k.$$

6/3/24

The Jordan Decomposition Theorem (cont.)

Lecture 24 (*) Proof (Jordan decomposition, cont.)

(cont.)

(i) Let B_i an ordered basis for K_{λ_i} for $i=1, \dots, k$. By Lemma 1, have that $B_i \cap B_j = \emptyset \forall i \neq j$; by part (ii), have that $B = B_1 \cup \dots \cup B_k$ spans \mathbb{C}^n .

Let $q = |B|$; then $q \geq n$. However, by Lemma 2, $\dim(K_{\lambda_i}) \leq m_i$ for each $i=1, \dots, k$

$$\rightarrow q = \sum_{i=1}^k \dim(K_{\lambda_i}) \leq \sum_{i=1}^k m_i = n$$

Hence $q = n$ [B basis for \mathbb{C}^n] and $\dim(K_{\lambda_i}) = m_i$ for each $i=1, \dots, k$.

(iii) Observe: for any cycle γ of generalized eigenvectors of T , $W = \text{span}(\gamma)$ is T -invariant.

Let $\gamma_1, \dots, \gamma_q$ cycles of generalized eigenvectors of T corresponding to an eigenvalue λ s.t. the initial vectors of each γ_i form a linearly-independent set. Clearly, the γ_i 's are disjoint; can find that $\gamma = \bigcup_{i=1}^q \gamma_i$ is linearly independent via induction on $|\gamma|$.

To show that we can find such a basis γ for every gen. eigenspace K_λ of T , can prove via induction on $\dim(K_\lambda)$ [using that $R[(T-\lambda I)|_{K_\lambda}] \subseteq K_\lambda$ is a gen. eigenspace of the restriction of T to K_λ] that can find γ [lin. indep. [from above] with $\dim(\gamma) = \dim(K_\lambda)$].

Clearly, the set $\beta = \bigcup \gamma_\lambda$ consisting of the union of bases of generalized eigenspaces forms a linearly independent set. Importantly, can find that β is a basis for \mathbb{C}^n iff the characteristic polynomial of T splits into linear factors $(t - \lambda_i I_n)$; guaranteed for \mathbb{C} [not \mathbb{R}].

Assume the characteristic polynomial $f(t)$ of T splits, i.e. that $f(t)$ has form

$f(t) = (t - \lambda_1 I)^{m_1} \dots (t - \lambda_k I)^{m_k}$ for some eigenvalues $\lambda_1, \dots, \lambda_k$. By Cayley-Hamilton,

$$f(A) = (A - \lambda_1 I)^{m_1} \dots (A - \lambda_k I)^{m_k} = 0 \Rightarrow f(A)v = (A - \lambda_1 I)^{m_1} \dots (A - \lambda_k I)^{m_k} v = 0, \forall v \in \mathbb{C}^n,$$

therefore the generalized eigenspaces $K_{\lambda_1}, \dots, K_{\lambda_k}$ span $\mathbb{C}^n \Rightarrow \beta = \bigcup \gamma_\lambda$ is a basis for \mathbb{C}^n .

The Minimal Polynomial

6/5/24

Lecture 25

Recall: A monic polynomial is a polynomial with highest-degree term having coefficient 1.

Def: Let $A \in M_{n \times n}(F)$. The minimal polynomial $p(t)$ of A is defined as the monic polynomial of least positive degree for which $p(A) = 0_{n \times n}$.

Theorem (Existence & Uniqueness): Let $A \in M_{n \times n}(F)$. Then:

- (i) There is a monic polynomial $p(t) \in P(F)$ of least degree s.t. $p(A) = 0_{n \times n}$.
- (ii) For any polynomial $g(t)$: if $g(A) = 0_{n \times n}$, then $p(t)$ divides $g(t)$. In particular, $p(t)$ divides the characteristic polynomial of A .
- (iii) The minimal polynomial is unique.

(*) Proof of Theorem

(i) By Cayley-Hamilton, the characteristic polynomial $f(t)$ of A is a monic polynomial s.t. $f(A) = 0_{n \times n}$, hence $\{p(t) \in P(F) : p(t) \text{ monic, } p(A) = 0_{n \times n}\}$ nonempty \Rightarrow has a member w/ least degree.

(ii) Let $p(t)$ be a [the] minimal polynomial, and let $g(t) \in P(F)$ monic s.t. $g(A) = 0_{n \times n}$. By definition, $\deg(g(t)) \geq \deg(p(t)) \Rightarrow$ can perform polynomial long division to find:

$$g(t) = q(t)p(t) + r(t)$$

$$= 0_{n \times n}$$

for some $q(t), r(t)$; $\deg(r(t)) < \deg(p(t))$. Then $g(A) = q(A)p(A) + r(A) = 0_{n \times n}$, thus $r(A) = 0_{n \times n}$. Since $p(t)$ was a minimal polynomial of A and $\deg(p(t)) > \deg(r(t))$, thus $r(t) \equiv 0$.

(iii) Let p_1, p_2 be two minimal polynomials of A , $\deg(p_1) = \deg(p_2)$. Per (ii), p_1 divides p_2 ; then $p_2 = cp_1$ for some $c \in F$. Since p_1, p_2 are both monic, then $c = 1 \Rightarrow p_1 = p_2$. \square

The Minimal Polynomial (cont.)

6/5/24

Lecture 25 Theorem: Let $A \in M_{n \times n}(F)$, $p(t)$ minimal polynomial of A . Then a scalar $\lambda \in F$ is an eigenvalue of A iff $p(\lambda) = 0$.

+ Lecture 26

Corollary: The minimal and characteristic polynomials share the same Os.

(+) Proof of Theorem

(\Leftarrow) Since $p(t)$ divides the characteristic polynomial, then for any $\lambda \in F$:

$$p(\lambda) = 0 \Rightarrow f(\lambda) = q(\lambda)p(\lambda) = 0 \Rightarrow \lambda \text{ is an eigenvalue of } A.$$

(\Rightarrow) Let $\lambda \in F$ eigenvalue of A , $x \in V$ eigenvector of A corresponding to λ . Then

$$Ax = \lambda x \Rightarrow p(A)x = p(\lambda)x. \text{ Since } p(A)x = 0_{n \times n} \cdot x = 0_v \text{ but } x \neq 0_v \text{ [by definition]},$$

then it must be the case that $p(\lambda) = 0$. ◻

Applications of Minimal Polynomials

Prop (1). Let $A \in M_{n \times n}(F)$ and let $\lambda_1, \dots, \lambda_k$ be the eigenvalues [distinct] of A . Then A is diagonalizable iff it has minimal polynomial $p(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_k)$.

Prop (2). Let $A \in M_{n \times n}(F)$ and let $\lambda_1, \lambda_2, \dots, \lambda_k$ be the [distinct] eigenvalues of A . Then for each $i=1, \dots, k$, the size of the largest Jordan block associated with λ_i is equal to the exponent of $(t - \lambda_i)$ in the minimal polynomial.

Consequently: can use characteristic/minimal polynomials to verify Jordan decomposition is in some cases (not all), may be sufficient to determine Jordan canonical form of A outright

The Minimal Polynomial (cont.)

6/7/24

Lecture 26

(cont.)

FINAL

Def: Let T be a linear operator on a fin.-dim. v.s.. Then a polynomial $p(t)$ is called the minimal polynomial of T if $p(t)$ is the monic polynomial of least positive degree s.t. $p(T) = 0_T$.

Theorem: The minimal polynomial of T on n -dimensional vector space V is equal to the minimal polynomial of each matrix representation of T .

(*) Note: Existence & uniqueness of the minimal polynomial of T follows from same argument as previously.