Math 524: Linear Algebra

Notes #8 — Operators on Complex Vector Spaces

Peter Blomgren
⟨blomgren@sdsu.edu⟩

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

http://terminus.sdsu.edu/

Fall 2021
(Revised: December 7, 2021)

Outline

Student Learning Targets, and Objectives 1 of 2

Target Generalized Eigenvectors and Nilpotent Operators
- Objective: Be able to identify generalized eigenspaces $G(\lambda, T)$
- Objective: Be able to identify a Nilpotent Operator, N, by the dimension of its Generalized Eigenspace $G(0, N)$
- Objective: Be able to construct an orthonormal basis so that the matrix of a Nilpotent Operator is upper trianguler with respect to the basis

Target Decomposition of an Operator
- Objective: Be able to Decompose all operators on complex vector spaces V onto direct sums of invariant generalized eigenspaces
- Objective: Be able to identify a Block Diagonal Matrix

Student Learning Targets, and Objectives 2 of 2

Target Characteristic Polynomial and the Cayley–Hamilton Theorem
- Objective: Be able to state the properties of the Characteristic Polynomial and its relation to the Eigenvalues of an Operator
- Objective: Be able to state the properties of the Minimal Polynomial and its relation to the Eigenvalues of an Operator
- Objective: Be able to derive the Characteristic and Minimal Polynomials for an Operator.

Target Jordan Form
- Objective: Be able to identify the Jordan Chains, and use them to construct a Jordan Basis for an Operator
- Objective: Be able to identify the Jordan Normal Form for an Operator

Introduction

We return to the issue of describing an operator in terms of its eigenspaces. In particular, we address the issue of non-Diagonalizability.

Rewind (Sum of Eigenspaces is a Direct Sum [NOTES#5])

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum. Furthermore, $\dim(E(\lambda_1, T)) + \cdots + \dim(E(\lambda_m, T)) \leq \dim(V)$

Rewind (Conditions Equivalent to Diagonalizability [NOTES#5])

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:

(a) T is diagonalizable.

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

(c) $\exists$ 1-D subspaces $U_1, \dots, U_n$ of V, each invariant under T, such that $V = U_1 \oplus \cdots \oplus U_n$ — There may be more than one U_* per eigenvalue!

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

(e) $\dim(V) = \dim(E(\lambda_1, T)) + \cdots + \dim(E(\lambda_m, T))$

$\mathrm{null}(T^k)$, for $T \in \mathcal{L}(V)$ — "Building the Toolbox"

We (temporarily) "discard" our inner products, and return to the simplicity of Vector Spaces. We look at the behavior of powers of operators T^k; first we look at the associated null-spaces

$\rightsquigarrow \rightsquigarrow \rightsquigarrow$ (Generalized) Eigenspaces

Theorem (Sequence of Increasing Null Spaces)

Suppose $T \in \mathcal{L}(V)$, then

$$\{0\} = \mathrm{null}(T^0) \subset \mathrm{null}(T^1) \subset \cdots \subset \mathrm{null}(T^k) \subset \mathrm{null}(T^{k+1}) \subset \cdots$$

Proof (Sequence of Increasing Null Spaces)

Suppose $T \in \mathcal{L}(V)$, let $k \geq 0$ and $v \in \mathrm{null}(T^k)$. Then $T^k(v) = 0$, and $T^{k+1}(v) = T(T^k(v)) = T(0) = 0$, so that $v \in \mathrm{null}(T^{k+1})$; thus $\mathrm{null}(T^k) \subset \mathrm{null}(T^{k+1})$. $\surd$

Equality in the Sequence of Null Spaces

Theorem (Equality in the Sequence of Null Spaces)

Suppose $T \in \mathcal{L}(V)$, let $m \geq 0$ such that $\mathrm{null}(T^m) = \mathrm{null}(T^{m+1})$, then

$$\mathrm{null}(T^m) = \mathrm{null}(T^{m+1}) = \mathrm{null}(T^{m+2}) = \cdots$$

Proof (Equality in the Sequence of Null Spaces)

Let $m, k \geq 0$. From the previous result we already have $\mathrm{null}(T^{k+m}) \subset \mathrm{null}(T^{k+m+1})$, to show equality we need to show $\mathrm{null}(T^{k+m+1}) \subset \mathrm{null}(T^{k+m})$: Let $v \in \mathrm{null}(T^{k+m+1})$, then

$$T^{m+1}(T^k(v)) = T^{k+m+1}(v) = 0$$

thus $T^k(v) \in \mathrm{null}(T^{m+1}) = \mathrm{null}(T^m) \Rightarrow T^{m+k}(v) = T^m(T^k(v)) = 0$ $\Rightarrow v \in \mathrm{null}(T^{k+m})$. $\surd$

Null Spaces Stop Growing

Theorem (Null Spaces Stop Growing)

Suppose $T \in \mathcal{L}(V)$, let $n = \dim(V)$, then

$$\mathrm{null}(T^n) = \mathrm{null}(T^{n+1}) = \cdots$$

Proof (Null Spaces Stop Growing)

By Contradiction: If the theorem is `false`, then

$$\{0\} = \mathrm{null}(T^0) \subsetneq \mathrm{null}(T^1) \subsetneq \cdots \subsetneq \mathrm{null}(T^n) \subsetneq \mathrm{null}(T^{n+1})$$

The strict inclusions means

$$\begin{aligned} 0 &= \dim(\mathrm{null}(T^0)) < \dim(\mathrm{null}(T^1)) < \cdots \\ &< \dim(\mathrm{null}(T^n)) < \dim(\mathrm{null}(T^{n+1})) \end{aligned}$$

so that $\dim(\mathrm{null}(T^{n+1})) \geq (n+1)$. But $\dim(V) = n$. $\surd$

Direct Sums of null and range

It is generally `true` that $V \neq \text{null}(T) \oplus \text{range}(T)$; *e.g.* recall examples where $\text{null}(T) = \text{range}(T)$.

Theorem ($V = \text{null}(T^n) \oplus \text{range}(T^n)$; $n = \dim(V)$)

Suppose $T \in \mathcal{L}(V)$, $n = \dim(V)$, *then*

$$V = \text{null}(T^n) \oplus \text{range}(T^n)$$

Proof ($V = \text{null}(T^n) \oplus \text{range}(T^n)$; $n = \dim(V)$)

(1) We show $\text{null}(T^n) \cap \text{range}(T^n) = \{0\}$:

Let $v \in \text{null}(T^n) \cap \text{range}(T^n)$, then $T^n(v) = 0$, and $\exists u \in V : v = T^n(u)$. Then $0 = T^n(v) = T^{2n}(u)$, using the previous result $\text{null}(T^n) = \text{null}(T^{2n})$, we must have $T^n(u) = 0$, hence $v = 0$.

Direct Sums of null and range

Proof ($V = \text{null}(T^n) \oplus \text{range}(T^n)$; $n = \dim(V)$)

(2) Since $\text{null}(T^n) \cap \text{range}(T^n) = \{0\}$ by [Direct Sum of Two Subspaces (Notes#1)] $\text{null}(T^n) + \text{range}(T^n)$ is a direct sum; and

$$\dim(\text{null}(T^n) \oplus \text{range}(T^n)) \stackrel{1}{=} \dim(\text{null}(T^n)) + \dim(\text{range}(T^n)) \stackrel{2}{=} \dim(V)$$

$\stackrel{1}{=}$ [A Sum is a Direct Sum if and only if Dimensions Add Up (Notes#3.2)]

$\stackrel{2}{=}$ [Fundamental Theorem of Linear Maps (Notes#3.1)]

Therefore $V = \text{null}(T^n) \oplus \text{range}(T^n)$. $\surd$

Direct Sums of null and range

Example

Consider $T \in \mathcal{L}(\mathbb{F}^3)$ defined by

$$\begin{aligned}
T(z_1, z_2, z_3) &= (4z_2, 0, 5z_3) \\
\text{range}(T) &= \{(w_1, 0, w_2) : w_1, w_2 \in \mathbb{F}\} \\
\text{null}(T) &= \{(w, 0, 0) : w \in \mathbb{F}\} \\
\hline
T^2(z_1, z_2, z_3) &= (0, 0, 25z_3) \\
T^3(z_1, z_2, z_3) &= (0, 0, 125z_3) \\
\text{range}(T^{\{2,3\}}) &= \{(0, 0, w) : w \in \mathbb{F}\} \\
\text{null}(T^{\{2,3\}}) &= \{(w_1, w_2, 0) : w_1, w_2 \in \mathbb{F}\}
\end{aligned}$$

Clearly $\mathbb{F}^3 = \text{range}(T^{\{2,3\}}) \oplus \text{null}(T^{\{2,3\}})$ — the theorem guarantees the result for $n = 3$, but here it happens sooner ($n = 2$).

Generalized Eigenvectors

As we have seen, some operators do not have enough eigenvectors to lead to a good description (diagonalization). We now introduce a remedy — **generalized eigenvectors**, which will aid in the description of the structure of operators.

For $\mathbb{C}$:normal $T^*T = TT^*$, and $\mathbb{R}$:self-adjoint $T = T^*$ operators we are guaranteed eigenspace decompositions

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

thanks to the [$\mathbb{C}/\mathbb{R}$ Spectral Theorems (Notes#7.1)].

[Schur's Theorem (Notes#6)] allows for an upper triangular matrix $\mathcal{M}(T)$ for every operator; but does not give a direct-sum decomposition of the space.

Generalized Eigenvectors and Eigenspaces

Definition (Generalized Eigenvector)

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. A vector $v \in V$ is called a **generalized eigenvector** of T corresponding to λ if $v \neq 0$ and

$$(T - \lambda I)^k(v) = 0$$

for some $k \geq 1$.

Definition (Generalized Eigenspace, $G(\lambda, T)$)

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. The **generalized eigenspace** of T corresponding to λ, denoted $G(\lambda, T)$, is defined to be the set of all generalized eigenvectors of T corresponding to λ, along with the 0 vector.

Generalized Eigenvectors and Eigenspaces

Since we get that standard eigenspace when $k = 1$, it is always `true` that

$$E(\lambda, T) \subset G(\lambda, T)$$

that is *"eigenvectors are also generalized eigenvectors."*

The next result answers the question *"what value of k should we pick?"*

Theorem (Description of Generalized Eigenspace)

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$ is an eigenvalue of T. Then $G(\lambda, T) = \mathrm{null}((T - \lambda I)^{\dim(V)})$.

Generalized Eigenvectors and Eigenspaces

Proof (Description of Generalized Eigenspace)

(1) Suppose $v \in \mathrm{null}((T - \lambda I)^{\dim(V)})$, then by definition $v \in G(\lambda, T)$, so $\mathrm{null}((T - \lambda I)^{\dim(V)}) \subset G(\lambda, T)$.

(2) Suppose $v \in G(\lambda, T)$, then $\exists k \geq 0$:

$$v \in \mathrm{null}((T - \lambda I)^k)$$

Applying [Sequence of Increasing Null Spaces] and [Null Spaces Stop Growing] to $(T - \lambda I)$ shows $v \in \mathrm{null}((T - \lambda I)^{\dim(V)})$ so that $G(\lambda, T) \subset \mathrm{null}((T - \lambda I)^{\dim(V)})$

Generalized Eigenvectors and Eigenspaces

Example

We revisit the $T \in \mathcal{L}(\mathbb{F}^3)$ defined by

$$\begin{array}{rcll} T(z_1, z_2, z_3) & = & (4z_2, 0, 5z_3) & \\ \mathrm{null}(T) & = & \{(w, 0, 0) : w \in \mathbb{F}\} & = \mathbf{E(0,T)} \\ \mathrm{null}(T^3) & = & \{(w_1, w_2, 0) : w_1, w_2 \in \mathbb{F}\} & = \mathbf{G(0,T)} \end{array}$$

Since $\dim(\mathrm{null}(T)) > 0$, $\lambda = 0$ is an eigenvalue; the other eigenvalue is $\lambda = 5$:

$$\mathbf{E(0,T)} = \{(\mathbf{w}, \mathbf{0}, \mathbf{0}) : \mathbf{w} \in \mathbb{F}\}, \quad \mathbf{E(5,T)} = \{(\mathbf{0}, \mathbf{0}, \mathbf{w}) : \mathbf{w} \in \mathbb{F}\}$$

and $(T - 5I)(z_1, z_2, z_3) = (4z_2 - 5z_1, -5z_2, 0)$, so

$$\begin{array}{rcl} (T - 5I)^2(z_1, z_2, z_3) & = & (25z_1 - 40z_2, 25z_2, 0) \\ (T - 5I)^3(z_1, z_2, z_3) & = & (300z_2 - 125z_1, -125z_2, 0) \\ \mathrm{null}((T - 5I)^3) & = & \{(0, 0, w) : w \in \mathbb{F}\} \end{array}$$

$$\mathbf{G(0,T)} = \{(\mathbf{w_1}, \mathbf{w_2}, \mathbf{0}) : \mathbf{w_1}, \mathbf{w_2} \in \mathbb{F}\}, \quad \mathbf{G(5,T)} = \{(\mathbf{0}, \mathbf{0}, \mathbf{w}) : \mathbf{w} \in \mathbb{F}\}$$

$$\mathbb{F}^3 = \mathbf{G(0,T)} \oplus \mathbf{G(5,T)}$$

Linearly Independent Generalized Eigenvectors

Rewind (Linearly Independent Eigenvectors [Notes#5])

Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T, and $v_1, \ldots, v_m$ are the corresponding eigenvectors; then $v_1, \ldots, v_m$ is linearly independent.

Theorem (Linearly Independent Generalized Eigenvectors)

Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T, and $v_1, \ldots, v_m$ are the corresponding **generalized** *eigenvectors; then $v_1, \ldots, v_m$ is linearly independent.*

Linearly Independent Generalized Eigenvectors

Proof (Linearly Independent Generalized Eigenvectors)

Suppose $a_1, \ldots, a_m \in \mathbb{C}$, $v_i \in G(\lambda_i, T)$ such that

$$0 = a_1 v_1 + \cdots + a_m v_m \qquad \text{(i)}$$

Let k be the largest non-negative integer such that $(T - \lambda_1 I)^k v_1 \neq 0$, and let $w = (T - \lambda_1 I)^k v_1$:

$$(T - \lambda_1 I) w = (T - \lambda_1 I)^{k+1} v_1 = 0$$

so $T(w) = \lambda_1 w$. Now, $(T - \lambda I) w = (\lambda_1 - \lambda) w \ \forall \lambda \in \mathbb{F}$, and

$$(T - \lambda I)^n w = (\lambda_1 - \lambda)^n w \qquad \text{(ii)}$$

$\forall \lambda \in \mathbb{F}$, $n = \dim(V)$.

We apply the operator (Note: the terms commute)

$$(T - \lambda_1 I)^k (T - \lambda_2 I)^n \cdots (T - \lambda_m I)^n$$

to (i).

Linearly Independent Generalized Eigenvectors

Proof (Linearly Independent Generalized Eigenvectors)

We get

$$\begin{aligned} 0 &\overset{1}{=} a_1 (T - \lambda_1 I)^k (T - \lambda_2 I)^n \cdots (T - \lambda_m I)^n v_1 \\ &\overset{2}{=} a_1 (T - \lambda_2 I)^n \cdots (T - \lambda_m I)^n w \\ &\overset{3}{=} a_1 (\lambda_1 - \lambda_2)^n \cdots (\lambda_1 - \lambda_m)^n w \end{aligned}$$

$\overset{1}{=}$ [Description of Generalized Eigenspaces]

$\overset{2}{=}$ $w = (T - \lambda_1 I)^k v_1$

$\overset{3}{=}$ (ii)

This forces $a_1 = 0$. We can now repeat the argument and show that $a_2 = a_3 = \cdots = a_m = 0$, which shows that $v_1, \ldots, v_m$ is linearly independent. $\surd$

Nilpotent Operators

Definition (Nilpotent)

An operator $N \in \mathcal{L}(V)$ is called **nilpotent** if $N^k = 0$ for some k.

Example (Some Nilpotent Operators)

- Operators with $\mathrm{null}(N) = \mathrm{range}(N)$, *e.g.* $N \in \mathcal{L}(\mathbb{F}^4)$
$$\begin{aligned} N(z_1, z_2, z_3, z_4) &= (z_3, z_4, 0, 0) \\ N^2(z_1, z_2, z_3, z_4) &= (0, 0, 0, 0) \end{aligned}$$
- Shift operators, *e.g.*
$$\begin{aligned} N(z_1, \ldots, z_n) &= (0, z_1, \ldots, z_{n-1}) \\ N^2(z_1, \ldots, z_n) &= (0, 0, z_1, \ldots, z_{n-2}) \\ N^n(z_1, \ldots, z_n) &= (0, 0, \ldots, 0) \end{aligned}$$
- $D \in \mathcal{L}(\mathcal{P}_m(\mathbb{F}))$ defined by $D(p) = p'$, since $D^{m+1}(p) = 0$.

Nilpotent Operators

Theorem (Nilpotent Operator Raised to Dimension of Domain is 0)

Suppose $N \in \mathcal{L}(V)$ is nilpotent, then $N^{\dim(V)} = 0$.

Proof (Nilpotent Operator Raised to Dimension of Domain is 0)

Since N is nilpotent $G(0, N) = V$. [Description of Generalized Eigenspace] implies $N^{\dim(V)} = 0$. $\surd$

Comment (Why Are We Here???)

Given $T \in \mathcal{L}(V)$, we want to find a basis $\mathfrak{B}(V)$ of V such that $\mathcal{M}(T; \mathfrak{B}(V))$ is as simple as possible, meaning that the matrix contains many 0's. Nilpotent operators will help us in this pursuit.

Matrix of a Nilpotent Operator

Theorem (Matrix of a Nilpotent Operator)

Suppose $N \in \mathcal{L}(V)$ is nilpotent, then $\exists$ a basis $\mathfrak{B}(V)$ so that

$$\mathcal{M}(N; \mathfrak{B}(V)) = \begin{bmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

that is $\mathcal{M}(N; \mathfrak{B}(V))$ is strictly upper triangular

Matrix of a Nilpotent Operator

Proof (Matrix of a Nilpotent Operator)

(1) Let $\mathfrak{B}_1(\mathrm{null}(N))$ be a basis for $\mathrm{null}(N)$.

(2) Let $\mathfrak{B}_2(\mathrm{null}(N^2))$ be an extension of $\mathfrak{B}_1(\mathrm{null}(N))$ to a basis for $\mathrm{null}(N^2)$.

(k+1) Let $\mathfrak{B}_{k+1}(\mathrm{null}(N^{k+1}))$ be an extension of $\mathfrak{B}_k(\mathrm{null}(N^k))$ to a basis for $\mathrm{null}(N^{k+1})$.

STOP when $\mathfrak{B}_{k+1}(\mathrm{null}(N^{k+1}))$ is a basis for V [Nilpotent Operator Raised to Dimension of Domain is 0] guarantees this will happen.

Now, consider this basis $\mathfrak{B}(V) = v_1, \ldots, v_n$ and the matrix $\mathcal{M}(N; \mathfrak{B}(V))$:

(i) The first $\dim(\mathrm{null}(N))$ columns corresponding to $\mathfrak{B}_1(\mathrm{null}(N))$ are all zeros (since they are a basis for $\mathrm{null}(N)$.

Matrix of a Nilpotent Operator

Proof (Matrix of a Nilpotent Operator)

(ii) The next $\dim(\mathrm{null}(N^2)) - \dim(\mathrm{null}(N))$ columns correspond to the extension of $\mathfrak{B}_1(\mathrm{null}(N))$ to a basis for $\mathrm{null}(N^2)$; any of these vectors $v_\ell \in \mathrm{null}(N^2)$, so $N(v_\ell) \in \mathrm{null}(N)$; this means

$$N(v_\ell) = a_1 v_1 + \cdots + a_{\dim(\mathrm{null}(N))} v_{\dim(\mathrm{null}(N))}$$

since $\ell > \dim(\mathrm{null}(N))$ this means only entries strictly above the diagonal are non-zero.

(k) As we process the columns $\dim(\mathrm{null}(N^{k+1})) - \dim(\mathrm{null}(N^k))$ corresponding to the extension of $\mathfrak{B}_k(\mathrm{null}(N))$ to a basis for $\mathrm{null}(N^{k+1})$; any of vectors in that block $v_\ell \in \mathrm{null}(N^{k+1})$, so $N(v_\ell) \in \mathrm{null}(N^k)$; which like above forces all diagonal and sub-diagonal entries in $\mathcal{M}(N)$ to be zeros.

$\surd$

Matrix of a Nilpotent Operator — Examples

Example (Revisited from [Slide 20])

- Given $N(z_1, z_2, z_3, z_4) = (z_3, z_4, 0, 0)$, we have

$$\mathcal{M}(N; \mathfrak{B}_{\text{std.coord}}) = \left[\begin{array}{cc:cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hdashline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \surd$$

- $N(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1})$, we have

$$\mathcal{M}(N; \mathfrak{B}_{\text{std.coord}}) = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \quad \text{not quite what we want}$$

Matrix of a Nilpotent Operator — Examples

Example (Revisited from [Slide 20])

- $N(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1})$, we have

$$\mathcal{M}(N; \mathfrak{B}_{\text{std.coord.reverse.order}}) = \left[\begin{array}{c|ccccc} & e_n & e_{n-1} & e_{n-2} & \cdots & e_1 \\ \hline e_n & 0 & 1 & 0 & \cdots & 0 \\ e_{n-1} & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ e_2 & \vdots & & & \ddots & 1 \\ e_1 & 0 & \cdots & \cdots & \cdots & 0 \end{array}\right] \quad \surd$$

Matrix of a Nilpotent Operator — Examples

Example (Revisited from [Slide 20])

- $D \in \mathcal{L}(\mathcal{P}_m(\mathbb{F}))$ defined by $D(p) = p'$, we have

$$\mathcal{M}(N; \mathfrak{B}_{\text{std.poly}}) = \left[\begin{array}{c|ccccc} & 1 & x & x^2 & \cdots & x^n \\ \hline 1 & 0 & 1 & 0 & \cdots & 0 \\ x & \vdots & \ddots & 2 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ x^{n-1} & \vdots & & & \ddots & n \\ x^n & 0 & \cdots & \cdots & \cdots & 0 \end{array}\right] \quad \surd$$

⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 8A-{3, 4, **5**}

Live Math :: Covid-19 Version 8A-5

8A-5: Suppose $T \in \mathcal{L}(V)$, m is a positive integer, and $v \in V$ is such that $T^{m-1}(v) \neq 0$ but $T^m(v) = 0$. Prove that

$$v, T(v), T^2(v), \ldots, T^{m-1}(v)$$

is linearly independent.

⁂ Step "0" ⁂

Suppose $a_0, a_1, \ldots, a_{m-1} \in \mathbb{F}$ are such that

$$a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_{m-1} T^{m-1}(v) = 0. \qquad \text{(8A-5.i)}$$

Since $T^{m-1}(v) \neq 0$, applying T^{m-1} to (8A-5.i) we get $a_0 T^{m-1}(v) = 0$; this implies $a_0 = 0$.

Live Math :: Covid-19 Version 8A-5

⁂ Step "1" ⁂

We now have $a_1, \ldots, a_{m-1} \in \mathbb{F}$ such that

$$a_1 T(v) + a_2 T^2(v) + \cdots + a_{m-1} T^{m-1}(v) = 0. \qquad \text{(8A-5.ii)}$$

Applying T^{m-2} to (8A-5.ii) we get $a_1 T^{m-1}(v) = 0$; this implies $a_1 = 0$.

⁂ Step "k" ⁂

Keep turning "the crank," and we get $a_0 = a_1 = \cdots = a_{m-1} = 0$, which means that

$$v, T(v), T^2(v), \ldots, T^{m-1}(v)$$

is linearly independent.

Invariance of The Null Space and Range of $p(T)$

We now put our new pieces together and show that every operator on a finite-dimensional complex vector space has enough generalized eigenvectors to provide a decomposition.

We need some "glue" for the proof of the main result:

Theorem (The Null Space and Range of $p(T)$ are Invariant Under T)

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$, then $\mathrm{null}(p(T))$ and $\mathrm{range}(p(T))$ are invariant under T.

Invariance of The Null Space and Range of $p(T)$

Proof (The Null Space and Range of $p(T)$ are Invariant Under T)

Suppose $v \in \mathrm{null}(p(T))$, then $p(T)(v) = 0$, hence

$$(p(T)(T(v)) = T(p(T)(v)) = T(0) = 0$$

$\to T(v) \in \mathrm{null}(p(T))$. $\surd_1$

Suppose $v \in \mathrm{range}(p(T))$. Then $\exists u \in V$: $v = p(T)(u)$, hence

$$T(v) = T(p(T)(u)) = p(T)(T(u))$$

$\to T(v) \in \mathrm{range}(p(T))$. $\surd_2$

Description of Operators on Complex Vector Spaces

Theorem (Description of Operators on Complex Vector Spaces)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. Then

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

(b) *each $G(\lambda_k, T)$ is invariant under T*

(c) *each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent.*

If we "trade in" our eigenspaces $E(\lambda_k, T)$ for generalized eigenspaces $G(\lambda_k, T)$, we can decompose **all** operators on a direct sum of invariant subspaces!

This is a fairly big deal.

Description of Operators on Complex Vector Spaces

Proof (Description of Operators on Complex Vector Spaces)

Let $n = \dim(V)$.

(b,c) By [Description of Generalized Eigenspace] $G(\lambda_k, T) = \text{null}((T - \lambda_k I)^n)$, applying [The Null Space and Range of $p(T)$ are Invariant Under T] with $p(z) = (z - \lambda_k)^n$, we get (b)$^{\text{invariance}}$.
(c)$^{\text{nilpotency}}$ follows directly from the definitions.

(a) [Induction–Base] If $n = 1$, (a) is trivially `true`.
Let $n > 1$, and assume:
[Induction–Hypothesis] (a) holds for all W: $\dim(W) < n$.
Since V is a complex vector space, T has an eigenvalue [Existence of Eigenvalues (Notes#5)]. Applying [$V = \text{null}(T^n) \oplus \text{range}(T^n)$; $n = \dim(V)$] to $(T - \lambda_1 I)$ shows

$$V = G(\lambda_1, T) \oplus U \qquad \text{(a*)}$$

where $U = \text{range}((T - \lambda_1 I)^n)$.

Description of Operators on Complex Vector Spaces

Proof (Description of Operators on Complex Vector Spaces)

Using [The Null Space and Range of $p(T)$ are Invariant Under T] with $p(z) = (z - \lambda_1)^n$, we see that U is invariant under T

Since $G(\lambda_1, T) \neq \{0\}$, $\dim(U) < n$, and we can apply the inductive hypothesis to $T|_U$.

All generalized eigenvectors corresponding to λ_1 are in $G(\lambda_1, T)$, hence the eigenvalues of $T|_U \in \{\lambda_2, \dots, \lambda_m\} \not\ni \lambda_1$. We can now write
$U = G(\lambda_2, T|_U) \oplus \cdots \oplus G(\lambda_m, T|_U)$

Showing that $G(\lambda_k, T|_U) = G(\lambda_k, T)$ completes the proof.

For each $k \in \{2, \dots, m\}$, the inclusion $G(\lambda_k, T|_U) \subset G(\lambda_k, T)$ is clear; to show the other direction, let $v \in G(\lambda_k, T)$. By (a*) we can write $v = v_1 + u$ where $v_1 \in G(\lambda_1, T)$, and $u \in U$. Need: $v_1 = 0$, $u \in G(\lambda_k, T|_U)$

Description of Operators on Complex Vector Spaces

Proof (Description of Operators on Complex Vector Spaces)

By the inductive hypothesis, we can write $u = v_2 + \cdots + v_m$, where each $v_\ell \in G(\lambda_\ell, T|_U) \subset G(\lambda_\ell, T)$. We have

$$v = v_1 + v_2 + \cdots + v_m$$

Since generalized eigenvectors corresponding to distinct eigenvalues are linearly independent [Linearly Independent Generalized Eigenvectors], this expression forces $v_\ell = 0 \; \forall \ell \neq k$.

Since $k \in \{2, \dots, m\}$ we must have $v_1 = 0$; thus $v = u \in U$ and since $v \in U$ we can conclude $v \in G(\lambda_k, T|_U)$. $\surd$

A Basis of Generalized Eigenvectors

The following looks like an afterthought, but it completes the story (so far)...

Theorem (A Basis of Generalized Eigenvectors)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.

Proof (A Basis of Generalized Eigenvectors)

Choose a basis of each $G(\lambda_k, T)$ in [Description of Operators on Complex Vector Spaces]. Put all these bases together to form a basis of V consisting of generalized eigenvectors of T.

Multiplicity

Definition (Multiplicity)

- Suppose $T \in \mathcal{L}(V)$. The multiplicity of an eigenvalue λ of T is defined to be the dimension of the corresponding generalized eigenspace $G(\lambda, T)$
- the multiplicity of an eigenvalue λ of T equals $\dim\left(\text{null}\left((T-\lambda I)^{\dim(V)}\right)\right)$

Comment (Multiplicity Math 254 vs. Math 524)

Math 254	Math 524
"algebraic multiplicity"	$\dim(\text{null}((T-\lambda I)^{\dim(V)})) = \dim(G(\lambda, T))$
"geometric multiplicity"	$\dim(\text{null}(T-\lambda I)) = \dim(E(\lambda, T))$

Multiplicity

Example

Let $T \in \mathcal{L}(\mathbb{C}^3)$ be defined by

$$T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3)$$

with respect to the standard basis:

$$\mathcal{M}(T) = \begin{bmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{bmatrix}, \quad \lambda(T) = \{6, 7\}.$$

$$E(6,T) = \text{span}((1,0,0)), \quad E(7,T) = \text{span}((10,2,1)))$$

$$G(6,T) = \text{span}((1,0,0),(0,1,0)), \quad G(7,T) = \text{span}((10,2,1))$$

$$\mathbb{C}^3 = G(6,T) \oplus G(7,T)$$

$$\mathfrak{B} = \{(1,0,0),(0,1,0),(10,2,1)\}$$

is a basis of $\mathbb{C}^3$ consisting of generalized eigenvectors of T.

Sum of the Multiplicities Equals $\dim(V)$

Theorem (Sum of the Multiplicities Equals $\dim(V)$)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then the sum of the multiplicities of all the eigenvalues of T equals $\dim(V)$.

Proof (Sum of the Multiplicities Equals $\dim(V)$)

[Description of Operators on Complex Vector Spaces] and [A Sum is a Direct Sum if and only if Dimensions Add Up (Notes#3.2)]

Comment (Multiplicity Without Determinants)

It is worth noting that our definition of multiplicity does not require determinants.

Also, we do not need two "types" of multiplicity.

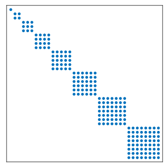
Block Diagonal Matrices

We introduce a bit more language for the discussion of matrix forms:

Definition (Block Diagonal Matrix)

A **Block Diagonal Matrix** is a square matrix of the form

$$A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$$



where $A_1, \ldots, A_m$ are square matrices along the diagonal; all other entries are 0.

Block Diagonal Matrix

Example (Block Diagonal Matrix)

Let

$$A_1 = \begin{bmatrix} 4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

then $A = \mathrm{diag}(A_1, A_2, A_3)$ is a block diagonal matrix:

$$A = \left[\begin{array}{c|cc|ccc} 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c|cc|ccc} 4 & & & & & \\ \hline & 2 & 1 & & & \\ & & 2 & & & \\ \hline & & & 1 & 1 & 0 \\ & & & & 1 & 1 \\ & & & & & 1 \end{array}\right]$$

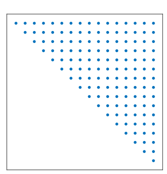
Block Diagonal Matrix with Upper-Triangular Blocks

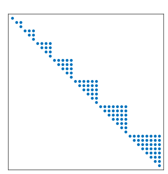
Theorem (Block Diagonal Matrix with Upper-Triangular Blocks)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$.

Then there is a basis of V with respect to which T has a block diagonal matrix of the form $A = \mathrm{diag}(A_1, \ldots, A_m)$, where each A_k is a $(d_k \times d_k)$ upper-triangular matrix of the form

$$A_k = \begin{bmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{bmatrix}$$





Upper-Triangular Matrix vs. Block Diagonal Matrix with Upper-Triangular Blocks

Rewind (Schur's Theorem [NOTES#6])

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some **orthonormal basis** of V.

Comment

Here, we are trading away the orthonormal basis; and we are getting more zeros in the matrix.

This is useful for theoretical purposes, but not always a good idea in practical computations.

For computational stability and accuracy [MATH 543], orthonormal bases are *very desirable.*

Block Diagonal Matrix with Upper-Triangular Blocks

Proof (Block Diagonal Matrix with Upper-Triangular Blocks)

Each $(T-\lambda_k I)|_{G(\lambda_k,T)}$ is nilpotent [Description of Operators on Complex Vector Spaces]. For each k, we select a basis $\mathfrak{B}_k$ of $G(\lambda_k,T)$ (which is a vector space with $\dim(G(\lambda_k,T)) = d_k$ such that

$$\mathcal{M}((T-\lambda_k I)|_{G(\lambda_k,T)}; \mathfrak{B}_k)$$

is strictly upper triangular; thus

$$\mathcal{M}(T|_{G(\lambda_k,T)}; \mathfrak{B}_k) = \mathcal{M}((T-\lambda_k I)|_{G(\lambda_k,T)} + \lambda_k I|_{G(\lambda_k,T)})$$

is upper triangular with λ_k repeated on the diagonal.

Collecting the bases $\mathfrak{B}_k(G(\lambda_k,T))$, $k=1,\dots,m$ gives a basis $\mathfrak{B}(V)$; and the $\mathcal{M}(T;\mathfrak{B}(V))$ has the desired structure.

Note, the example matrix on [Slide 42] is in this form. For an operator T on a 6-dimensional vectors space with $\mathcal{M}(T)$ as in the example, the eigenvalues are $\{4,2,1\}$ with corresponding multiplicities $\{1,2,3\}$.
Additionally, the matrices on [Slides 25–27] are in this form.

Generalized Eigenvectors and Nilpotent Operators
Decomposition of an Operator
Characteristic and Minimal Polynomials
Jordan Form
Description of Operators on Complex Vector Spaces
Multiplicity of an Eigenvalue
Block Diagonal Matrices
Square Roots

Block Diagonal Matrix with Upper-Triangular Blocks

Example (Revisited from [Slide 39])

Let $T \in \mathcal{L}(\mathbb{C}^3)$ be defined by $T(z_1,z_2,z_3) = (6z_1+3z_2+4z_3, 6z_2+2z_3, 7z_3)$ with respect to the standard basis the matrix is not in the desired form:

$$\mathcal{M}(T) = \begin{bmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{bmatrix}, \quad \lambda(T) = \{6,7\}.$$

However, $G(6,T) = \operatorname{span}((1,0,0),(0,1,0))$, $G(7,T) = \operatorname{span}((10,2,1))$, so that

$$\mathfrak{B} = \{(1,0,0),(0,1,0),(10,2,1)\},$$

and

$$\mathcal{M}(T;\mathfrak{B}) = \left[\begin{array}{cc|c} 6 & 3 & \\ & 6 & \\ \hline & & 7 \end{array}\right]$$

Generalized Eigenvectors and Nilpotent Operators
Decomposition of an Operator
Characteristic and Minimal Polynomials
Jordan Form
Description of Operators on Complex Vector Spaces
Multiplicity of an Eigenvalue
Block Diagonal Matrices
Square Roots

Square Roots

Some, but not all operators have square roots. At this point, we know that positive operators have positive square roots [Characterization of Positive Operators (Notes#7.2)]. To that we add:

Theorem (Identity Plus Nilpotent has a Square Root)

Suppose $N \in \mathcal{L}(V)$ is nilpotent, then $(I+N)$ has a square root.

Proof (Identity Plus Nilpotent has a Square Root)

We use the Taylor series for $\sqrt{1+x}$ as motiviation:

$$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots + \text{"} a_\infty x^\infty \text{"}$$

for our purpose, the values of the coefficients are not (yet) important; since $N \in \mathcal{L}(V)$ is nilpotent $N^m = 0$ for some value of m, we seek a square root of the form

$$\sqrt{I+N} = I + a_1 N + a_2 N^2 + \cdots + a_{m-1} N^{m-1}$$

Generalized Eigenvectors and Nilpotent Operators
Decomposition of an Operator
Characteristic and Minimal Polynomials
Jordan Form
Description of Operators on Complex Vector Spaces
Multiplicity of an Eigenvalue
Block Diagonal Matrices
Square Roots

Square Roots

Proof (Identity Plus Nilpotent has a Square Root)

We select the coefficients $a_1,\dots,a_{m-1}$ so that

$$I + N = (I + a_1 N + a_2 N^2 + \cdots + a_{m-1} N^{m-1})^2$$

Given enough patience, we can figure out what the coefficient values should be; but all we need is that they exist. $\surd$

We can now use this results to guarantee that all invertible operators (over $\mathbb{C}$) have square roots...

Square Roots

Note that this result does not hold over $\mathbb{R}$, *e.g.* $T(x) = -x$, $x \in \mathbb{R}$ does not have a square root.

Theorem (Over $\mathbb{C}$, Invertible Operators Have Square Roots)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a square root.

Proof (Over $\mathbb{C}$, Invertible Operators Have Square Roots)

Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. For each k $\exists$ a nilpotent $N_k \in \mathcal{L}(G(\lambda_k, T))$ such that such that $T|_{G(\lambda_k,T)} = \lambda_k I + N_k$ [Description of Operators on Complex Vector Spaces]. Since T is invertible $\lambda_k \neq 0$, we can write

$$T|_{G(\lambda_k,T)} = \lambda_k \left(I + \frac{N_k}{\lambda_k} \right), \quad k = 1, \dots, m$$

Square Roots

Proof (Over $\mathbb{C}$, Invertible Operators Have Square Roots)

The scaled operator N_k/λ_k are nilpotent, each $(I + N_k/\lambda_k)$ has a square root [Identity Plus Nilpotent has a Square Root].
$R_k = \sqrt{\lambda_k}\sqrt{I + N_k/\lambda_k}$ is the square root R_k of $T|_{G(\lambda_k,T)}$.

Any $v \in V$ can be uniquely written in the form

$$v = u_1 + \cdots + u_m, \;\; u_k \in G(\lambda_k, T)$$

[Description of Operators on Complex Vector Spaces]. Now define $R \in \mathcal{L}(V)$ by

$$R(v) = R_1(u_1) + \cdots + R_m(u_m),$$

since $\forall u_\ell \in G(\lambda_\ell, T)$ $R_\ell(u_\ell) \in G(\lambda_\ell, T)$

$$R^2(v) = R_1^2(u_1) + \cdots + R_m^2(u_m), = T|_{G(\lambda_1,T)}(u_1) + \cdots + T|_{G(\lambda_m,T)}(u_m) = T(v)$$

⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 8B-{3, **4**, 5}

Live Math :: Covid-19 Version 8B-4

8B-4: Suppose $T \in \mathcal{L}(V)$, $\dim(V) = n$, and $\mathrm{null}(T^{n-2}) \neq \mathrm{null}(T^{n-1})$. Show that T has at most two distinct eigenvalues.

Since $\mathrm{null}(T^{n-2}) \neq \mathrm{null}(T^{n-1})$

$$\{0\} = \mathrm{null}(T^0) \subsetneq \mathrm{null}(T^1) \subsetneq \cdots \subsetneq \mathrm{null}(T^{n-2}) \subsetneq \mathrm{null}(T^{n-1})$$

Therefore

$$\dim(\mathrm{null}(T^{n-1})) \geq (n-1) \Leftrightarrow \dim(G(0,T)) \geq (n-1)$$

Also, we know $(\lambda_j \neq 0)$

$$\underbrace{V}_{\dim(V)=n} = \underbrace{G(0,T)}_{\dim(G(0,T))\geq(n-1)} \oplus \underbrace{\left[\overbrace{G(\lambda_1,T) \oplus \cdots \oplus G(\lambda_m,T)}^{W} \right]}_{\dim(W)\leq 1}$$

Characteristic Polynomial

Keep in mind: All the polynomial action here is over $\mathbb{F} = \mathbb{C}$.

Definition (Characteristic Polynomial)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T, with multiplicities $d_1, \dots, d_m$. The polynomial

$$p_T(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$$

is called the **characteristic polynomial** of T.

Comment

Again, we have defined something familiar from [MATH 254] without the use of the determinant.

Characteristic Polynomial

Example (Characteristic Polynomials)

The characteristic polynomials associated with previous examples

- [SLIDE 39]: $p(z) = (z-6)^2(z-7)^1$
- [SLIDE 42]: $p(z) = (z-4)^1(z-2)^2(z-1)^3$

Theorem (Degree and Zeros of Characteristic Polynomial)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then

(a) *the characteristic polynomial, $p_T(z)$ of T has degree $\dim(V)$*

(b) *the zeros of $p_T(z)$ are the eigenvalues of T.*

Proof (Degree and Zeros of Characteristic Polynomial)

(a) follows from [SUM OF THE MULTIPLICITIES EQUALS $\dim(V)$], and
(b) from the definition of the characteristic polynomial.

Cayley–Hamilton Theorem

Theorem (Cayley–Hamilton Theorem)

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $p_T(z)$ denote the characteristic polynomial of T. Then $p_T(T) = 0$.

Comment (Cayley–Hamilton Theorem over $\mathbb{R}$)

The Cayley–Hamilton Theorem also holds for real vector spaces.

Comment (Importance of the Cayley–Hamilton Theorem)

The Cayley–Hamilton Theorem is one of the key structural theorems in linear algebra. For one thing it gives us the "license" to find eigenvalues using the characteristic polynomial.

Cayley–Hamilton Theorem

Proof (Cayley–Hamilton Theorem)

Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of the operator T, and let $d_1, \dots, d_m$ be the dimensions of the corresponding generalized eigenspaces $G(\lambda_1, T), \dots, G(\lambda_m, T)$.

For each $k = 1, \dots, m$, we know that $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent. Thus we have $(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0$ [NILPOTENT OPERATOR RAISED TO DIMENSION OF DOMAIN IS 0].

Every vector in V is a sum of vectors in $G(\lambda_1, T), \dots, G(\lambda_m, T)$ [DESCRIPTION OF OPERATORS ON COMPLEX VECTOR SPACES]; *i.e.* $\forall v \in V$, and $\exists v_\ell \in G(\lambda_\ell, T)$: $v = v_1 + \cdots + v_\ell$.

To prove that $p_T(T) = 0$ ($\Leftrightarrow p_T(T)v = 0 \ \forall v \in V$), we need only show that $p_T(T)|_{G(\lambda_k, T)} = 0$, $k = 1, \dots, m$.

Cayley–Hamilton Theorem

Proof (Cayley–Hamilton Theorem)

Fix $k \in \{1, \ldots, m\}$. We have

$$p_T(T) = (T - \lambda_1 I)^{d_1} \cdots (T - \lambda_m I)^{d_m}.$$

The operators on the right side of the equation above all commute, so we can move the factor $(T - \lambda_k I)^{d_k}$ to be the last term in the expression on the right.

Since $(T - \lambda_k I)^{d_k}|_{G(\lambda_k, T)} = 0$, we conclude that $p_T(T)|_{G(\lambda_k, T)} = 0$. $\surd$

Monic Polynomial

Here, we introduce an alternative polynomial which can be used to identify eigenvalues.

First, we need some language and notation (us usual!)

Definition (Monic Polynomial)

A monic polynomial is a polynomial whose highest-degree coefficient equals 1.

Example

- Monic — $p(z) = z^{407} - \pi z^{103} + \sqrt{7}$
- Not monic — $q(z) = (1 + \epsilon)z^2 + 1$, $\epsilon > 0$

Minimal Polynomial

Theorem (Minimal Polynomial)

Suppose $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial p of smallest degree such that $p(T) = 0$.

Proof (Minimal Polynomial)

Let $n = \dim(V)$, then the list (of length $(n^2 + 1)$)

$$I, T, T^2, \ldots, T^{n^2}$$

is not linearly independent in $\mathcal{L}(V)$, since $\dim(\mathcal{L}(V)) = n^2$. Let m be the smallest positive integer such that the list

$$I, T, T^2, \ldots, T^m \qquad \text{(i)}$$

is linearly dependent. [Linear Dependence (Notes#2)] implies that one of the operators in the list above is a linear combination of the previous ones.

Minimal Polynomial

Proof (Minimal Polynomial)

The choice of m means that T^m is a linear combination of $I, T, T^2, \ldots, T^{m-1}$; hence $\exists a_0, a_1, \ldots, a_{m-1} \in \mathbb{F}$ such that

$$a_0 I + a_1 T + a_2 T^2 + \ldots a_{m-1} T^{m-1} + T^m = 0 \qquad \text{(ii)}$$

We use the coefficients to define a monic polynomial $p \in \mathcal{P}(\mathbb{F})$ by:

$$p(z) = a_0 + a_1 z + a_2 z^2 + \ldots a_{m-1} z^{m-1} + z^m$$

By (ii) $p(T) = 0$. That takes care of existence.

To show uniqueness, note that the choice of m implies that no monic polynomial $q \in \mathcal{P}(\mathbb{F})$ with degree smaller than m can satisfy $q(T) = 0$.

Suppose $q \in \mathcal{P}(\mathbb{F})$ with degree m and $q(T) = 0$. Then $(p - q)(T) = 0$ and $\deg(p - q) < m$. The choice of m now implies that $(p - q)$ is the zero-polynomial $\Leftrightarrow q = p$, completing the proof.

Minimal Polynomial of an Operator T

Definition (Minimal Polynomial (of an operator T))

Suppose $T \in \mathcal{L}(V)$. Then the minimal polynomial of T is the unique monic polynomial p of smallest degree such that $p(T) = 0$.

The proof of the last theorem shows that the degree of the minimal polynomial of each operator on V is at most $(\dim(V))^2$. The [CAYLEY–HAMILTON THEOREM] tells us that if V is a complex vector space, then the minimal polynomial of each operator on V has degree at most $\dim(V)$.

This improvement $(\dim(V))^2 \to \dim(V)$ also holds on real vector spaces.

Finding the Minimal Polynomial — Take#1

"Guaranteed"* to Work, Labor Intensive:

Given the matrix $\mathcal{M}(T)$ (with respect to some basis) of an operator $T \in \mathcal{L}(V)$. The minimal polynomial of T can be identified as follows: Consider the system of ($\dim(V)^2$ — each matrix entry) linear equations**

$$a_0\mathcal{M}(I) + a_1\mathcal{M}(T) + \cdots + a_{m-1}\mathcal{M}(T)^{m-1} = -\mathcal{M}(T)^m \qquad \text{(i)}$$

for successive values of $m = 1, \dots, \dim(V)^2$; until there is a solution $a_1, \dots, a_{m-1}$; the minimal polynomial is then given by

$$p(z) = a_0 + a_1 z + \cdots + a_{m-1}z^{m-1} + z^m$$

* Requires an iPhone XIX with infinite precision processing capabilities.

** The linear systems are of the form $A\vec{x} = \vec{b}$, where $A \in \mathbb{F}^{\dim(V)^2 \times m}$, $\vec{b} \in \mathbb{F}^{\dim(V)^2}$, and the solution vector $\vec{x} \in \mathbb{F}^m = (a_0, a_1, \dots, a_{m-1})$.

Finding the Minimal Polynomial — Take#2

Works "Almost Always"*, Less Labor Intensive:

Given the matrix $\mathcal{M}(T)$. The minimal polynomial of T can with probability 1 be identified as follows: Pick a random vector $v \in \mathbb{F}^{\dim(V)}$, and consider the system of ($\dim(V)$ — each vector entry) linear equations

$$a_0\mathcal{M}(I)v + a_1\mathcal{M}(T)v + \cdots + a_{m-1}\mathcal{M}(T)^{m-1}v = -\mathcal{M}(T)^m v \qquad \text{(i)}$$

for successive values of $m = 1, \dots, \dim(V)$; until there is a solution $a_1, \dots, a_{m-1}$; the minimal polynomial is the given by

$$p(z) = a_0 + a_1 z + \cdots + a_{m-1}z^{m-1} + z^m$$

* The random $v \in \mathbb{F}^{\dim(V)}$ must be such that $v = u_1 + \cdots + u_m$, $u_\ell \neq 0$; where $\mathbb{F}^{\dim(V)} = G(\lambda_1, \mathcal{M}(T)) \oplus \cdots \oplus G(\lambda_m, \mathcal{M}(T))$, and $u_\ell \in G(\lambda_\ell, \mathcal{M}(T))$, and still requires an iPhone XIX with infinite precision processing capabilities. See [MATH 543] for discussion on finite precision computing.

Finding the Minimal Polynomial — Computation

Example

Let $T \in \mathcal{L}(\mathbb{C}^5)$, with $\mathcal{M}(T)$ wrt the standard basis:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Lets try the "random" vector $v = (1,1,1,1,1)$, we construct $A \in \mathbb{C}^{5\times 6}$ by letting the kth column a_k be $\mathcal{M}(T)^{k-1}v$:

$$A = \begin{bmatrix} 1 & -3 & -3 & -3 & -3 & -21 \\ 1 & 7 & 3 & 3 & 3 & 39 \\ 1 & 1 & 7 & 3 & 3 & 3 \\ 1 & 1 & 1 & 7 & 3 & 3 \\ 1 & 1 & 1 & 1 & 7 & 3 \end{bmatrix}$$

Finding the Minimal Polynomial — Computation

Example

We are looking for a solution to a linear system; we put [MATH 254] to good use, and compute

$$\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -3 \\ 0 & 1 & 0 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Since the 6th column is linearly dependent (it does not have a leading one); we identify the linear relation:

$$-3a_1 + 6a_2 - a_6 = 0$$

which yields the minimal polynomial $p(z) = 3 - 6z + z^5$

Finding the Minimal Polynomial — Computation

Example

Finally, "we"* compute $p(\mathcal{M}(T)) = 3I - 6\mathcal{M}(T) + \mathcal{M}(T)^5$:

$$3\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} - 6\begin{bmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} -3 & 0 & 0 & 0 & -18 \\ 6 & -3 & 0 & 0 & 36 \\ 0 & 6 & -3 & 0 & 0 \\ 0 & 0 & 6 & -3 & 0 \\ 0 & 0 & 0 & 6 & -3 \end{bmatrix} = 0$$

Which shows that we indeed have the minimal polynomial.

* We = I + my computer.

Multiple of the Minimal Polynomial

Theorem ($q(T) = 0 \Leftrightarrow q$ is a Multiple of the Minimal Polynomial)

Suppose $T \in \mathcal{L}(V)$ $q \in \mathcal{P}(\mathbb{F})$. Then $q(T) = 0$ if and only if q is a polynomial multiple of the minimal polynomial of T.

Proof ($q(T) = 0 \Leftrightarrow q$ is a Multiple of the Minimal Polynomial)

Let p be the minimal polynomial of T.

($\Leftarrow$) Suppose q is a polynomial multiple of p. Thus $\exists s \in \mathcal{P}(\mathbb{F})$ such that $q = ps$, and

$$q(T) = p(T)s(T) = 0s(T) = 0$$

Multiple of the Minimal Polynomial

Proof ($q(T) = 0 \Leftrightarrow q$ is a Multiple of the Minimal Polynomial)

($\Rightarrow$) Suppose $q(T) = 0$. By [DIVISION ALGORITHM FOR POLYNOMIALS (NOTES#4)], $\exists s, r \in \mathcal{P}(\mathbb{F})$ such that

$$q = ps + r \qquad \text{(i)}$$

and $\deg(r) < \deg(p)$, therefore

$$0 = q(T) = \underbrace{p(T)}_{\text{min.poly}} s(T) + r(T) = r(T)$$

hence, $r(T) = 0 \Rightarrow r \equiv 0 \Rightarrow q = ps$. $\surd$

Characteristic Polynomial and Minimal Polynomial

Theorem (Characteristic Polynomial is a Multiple of Minimal Polynomial)

Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.

Note: We have not yet defined the characteristic polynomial when $\mathbb{F} = \mathbb{R}$, but once we do, the above theorem will apply.

Proof (Characteristic Polynomial is a Multiple of Minimal Polynomial)

By [Cayley–Hamilton Theorem], $p_T^{\text{char}}(T) = 0$; and [$q(T) = 0 \Leftrightarrow q$ is a Multiple of the Minimal Polynomial] shows $p_T^{\text{char}}(T) = s(T)p_T^{\text{min}}(T)$.

The Minimal Polynomial → Eigenvalues

Theorem (Eigenvalues are the Zeros of the Minimal Polynomial)

Let $T \in \mathcal{L}(V)$. Then the zeros of the minimal polynomial of T are the eigenvalues of T.

Proof (Eigenvalues are the Zeros of the Minimal Polynomial)

Let

$$p(z) = a_0 + a_1 z + \cdots + a_{m-1}z^{m-1} + z^m$$

be the minimum polynomial of T.

Suppose $\lambda \in \mathbb{F}$ is a zero of p. Then p can be written in the form

$$p(z) = (z - \lambda)q(z)$$

where q is a monic polynomial with coefficients in $\mathbb{F}$ [Each Zero of a Polynomial Corresponds to a Degree-1 Factor (Notes#4)],

The Minimal Polynomial → Eigenvalues

Proof (Eigenvalues are the Zeros of the Minimal Polynomial)

Since $p(T) = 0$, we have

$$0 = (T - \lambda I)q(T)(v), \; \forall v \in V.$$

Since $\deg(q) < \deg(p)$ $\exists v \in V : q(T)(v) \neq 0$; therefore λ must be an eigenvalue of T.

Now, suppose $\lambda \in \mathbb{F}$ is an eigenvalue of T. $\exists v \in V : v \neq 0$, $T^j(v) = \lambda^j v$, $j = 1, \ldots$. Now,

$$\begin{aligned} 0 = p(T)(v) &= (a_0 I + a_1 T + \cdots + a_{m-1}T^{m-1} + T^m)(v) \\ &= (a_0 + a_1\lambda + \cdots + a_{m-1}\lambda^{m-1} + \lambda^m)v \\ &= p(\lambda)v \end{aligned}$$

$\Rightarrow p(\lambda) = 0.$ $\surd$

The Minimal Polynomial ↔ Eigenvalues

Example (Re-revisited [Slides 39, 46])

Let $T \in \mathcal{L}(\mathbb{C}^3)$ be defined by
$T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3)$ wrt the standard basis:

$$\mathcal{M}(T) = \begin{bmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{bmatrix}, \quad \lambda(T) = \{6, 7\}.$$

$$\begin{aligned} G(6,T) &= \operatorname{span}((1,0,0),(0,1,0)) & \dim(G(6,T)) &= 2 \\ G(7,T) &= \operatorname{span}((10,2,1)) & \dim(G(7,T)) &= 1 \end{aligned}$$

the characteristic polynomial is $\mathbf{p_T(z) = (z - 6)^2(z - 7)}$; the minimal polynomial is either $(z-6)^2(z-7)$ or $(z-6)(z-7)$. Since

$$(\mathcal{M}(T) - 6I)(\mathcal{M}(T) - 7I) = \begin{bmatrix} 0 & -3 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (\mathcal{M}(T) - 6I)^2(\mathcal{M}(T) - 7I) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

it follows that the minimal polynomial of T is $\mathbf{p_T^{\min}(z) = (z - 6)^2(z - 7)}$.

The Minimal Polynomial ↔ Eigenvalues

Example (Modified)

Let $T \in \mathcal{L}(\mathbb{C}^3)$ be defined by $T(z_1, z_2, z_3) = (6z_1, 6z_2, 7z_3)$ wrt the standard basis:

$$\mathcal{M}(T) = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \quad \lambda(T) = \{6, 7\}.$$

$$\begin{array}{ll} G(6,T) = \mathrm{span}((1,0,0),(0,1,0)) & \dim(G(6,T)) = 2 \\ G(7,T) = \mathrm{span}((0,0,1)) & \dim(G(7,T)) = 1 \end{array}$$

the characteristic polynomial is $\mathbf{p_T(z) = (z-6)^2(z-7)}$; the minimal polynomial is either $(z-6)^2(z-7)$ or $(z-6)(z-7)$. Since

$$(\mathcal{M}(T) - 6I)(\mathcal{M}(T) - 7I) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

it follows that the minimal polynomial of T is $\mathbf{p_T^{\min}(z) = (z-6)(z-7)}$.

⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 8C-{3, **4**, 5}

Live Math :: Covid-19 Version 8C-4

8C-4: Given an example $T \in \mathcal{L}(\mathbb{C}^4)$ whose characteristic polynomial is $p(z) = (z-1)(z-5)^3$, and minimal polynomial is $q(z) = (z-1)(z-5)^2$.

Any T with $\mathcal{M}(T) \in \mathbb{R}^{4\times 4}$, upper triangular, with diagonal $(1,5,5,5)$ will have $p(z) = (z-1)(z-5)^3$.

※ Diagonal $\mathcal{M}(T)$ ※

We try

$$\mathcal{M}(T) = \begin{bmatrix} 1 & & & \\ & 5 & & \\ & & 5 & \\ & & & 5 \end{bmatrix} \quad \Leftrightarrow \quad T(z_1, z_2, z_3, z_4) = (z_1, 5z_2, 5z_3, 5z_4)$$

$$q(z) \in \{\ (z-1)(z-5),\ (z-1)(z-5)^2,\ (z-1)(z-5)^3\ \}$$

Live Math :: Covid-19 Version 8C-4

$$(\mathcal{M}(T) - I_4)(\mathcal{M}(T) - 5I_4) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad q(z) = (z-1)(z-5).$$

※ Upper Triangular $\mathcal{M}(T)$ ※

$$\mathcal{M}(T) = \begin{bmatrix} 1 & & & \\ & 5 & 1 & \\ & & 5 & \\ & & & 5 \end{bmatrix} \quad \Leftrightarrow \quad T(z_1, z_2, z_3, z_4) = (z_1, 5z_2 + z_3, 5z_3, 5z_4)$$

$$(\mathcal{M}(T) - I_4)(\mathcal{M}(T) - 5I_4) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad q(z) \neq (z-1)(z-5).$$

Live Math :: Covid-19 Version

8C-4

$$(\mathcal{M}(T) - I_4)(\mathcal{M}(T) - 5I_4)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \boxed{q(z) = (z-1)(z-5)^2.}$$

Jordan "Normal" / "Canonical" Form

At this point we know that if V is a complex vector space, then $\forall T \in \mathcal{L}(V)$ there is a basis of V with respect to which T has [Block Diagonal Matrix with Upper-Triangular Blocks (Slide 43)].

Now, we a chasing more zeros: the goal is a basis of V wrt which the matrix of T contains 0's everywhere except possibly on

- the diagonal *(the eigenvalues)*, and
- the first super-diagonal *(we allow 1's or 0's)*.

We use nilpotent operators to get us there...

Examples

Example (Compare with [Slide 25–26])

Once again we consider a shift-operator in $\mathbb{F}^4$: $N(z_1, \ldots, z_4) = (0, z_1, \ldots, z_3)$; its action on $v = (1,0,0,0)$ generates a basis

$$\mathfrak{B}(\mathbb{F}^4) = \left\{N^3(v), N^2(v), N(v), v\right\} = \{e_4, e_3, e_2, e_1\}, \text{ and}$$

$$\mathcal{M}(N, \mathfrak{B}(\mathbb{F}^4)) = \left[\begin{array}{c|cccc} & e_4 & e_3 & e_2 & e_1 \\ \hline e_4 & 0 & 1 & 0 & 0 \\ e_3 & 0 & 0 & 1 & 0 \\ e_2 & 0 & 0 & 0 & 1 \\ e_1 & 0 & 0 & 0 & 0 \end{array}\right]$$

Definition (Jordan Chain — Generator / Lead Vector; adopted from [Wikipedia])

Given an eigenvalue λ, its corresponding Jordan block gives rise to a **Jordan chain**. The **generator**, or **lead vector**, v_r of the chain is a generalized eigenvector such that $(A - \lambda I)^r v_r = 0$, where r is the size of the Jordan block. The vector $v_1 = (A - \lambda I)^{r-1} v_r$ is an eigenvector corresponding to λ.

In general, $v_{i-1} = (A - \lambda I)v_i$. The lead vector generates the chain via repeated multiplication by $(A - \lambda I)$. $\mathfrak{B} = \{v_1, \cdots, v_r\}$ is a basis for the Jordan block.

Examples

Example

Let $N \in \mathcal{L}(\mathbb{F}^6)$: $N(z_1, \ldots, z_6) = (0, z_1, z_2, 0, z_4, 0)$.

Here thinking of a space isomorphic to $\mathbb{F}^6$ helps: $\mathbb{F}^6 \cong \mathbb{F}^3 \times \mathbb{F}^2 \times \mathbb{F}^1$.

On each space we have a right-shift operator: $N_{\mathbb{F}^3}(z_1, z_2, z_3) = (0, z_1, z_2)$, $N_{\mathbb{F}^2}(z_1, z_2) = (0, z_1)$, $N_{\mathbb{F}^1}(z_1) = (0)$, and we can define the linear map $N_\times : \mathbb{F}^3 \times \mathbb{F}^2 \times \mathbb{F}^1 \mapsto \mathbb{F}^3 \times \mathbb{F}^2 \times \mathbb{F}^1$ by

$$\begin{aligned} N_\times((z_1, z_2, z_3), (z_4, z_5), z_6) &= (N_{\mathbb{F}^3}(z_1, z_2, z_3), N_{\mathbb{F}^2}(z_4, z_5), N_{\mathbb{F}^1}(z_6)) \\ &= ((0, z_1, z_2), (0, z_4), (0)). \end{aligned}$$

By the previous example, the lead vectors $w_{\mathbb{F}^3,1} = (1,0,0)$, $w_{\mathbb{F}^2,1} = (1,0)$, and $w_{\mathbb{F}^1,1} = (1)$ will generate bases for $\mathfrak{B}(\mathbb{F}^3)$, $\mathfrak{B}(\mathbb{F}^2)$, and $\mathfrak{B}(\mathbb{F}^1)$ so that

$$\mathcal{M}\left(N_{\mathbb{F}^3}, \mathfrak{B}(\mathbb{F}^3)\right) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{M}\left(N_{\mathbb{F}^2}, \mathfrak{B}(\mathbb{F}^2)\right) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{M}\left(N_{\mathbb{F}^1}, \mathfrak{B}(\mathbb{F}^1)\right) = \begin{bmatrix} 0 \end{bmatrix}.$$

Examples

Example

If we "translate" all that back to $N \in \mathcal{L}(\mathbb{F}^6)$: $N(z_1, \ldots, z_6) = (0, z_1, z_2, 0, z_4, 0)$.
We have 3 lead vectors:

$$w_3 = (1,0,0,0,0,0), \quad w_2 = (0,0,0,1,0,0), \quad w_1 = (0,0,0,0,0,1).$$

$\mathfrak{B}(\mathbb{F}^6) = \left\{ N^2(w_3), N(w_3), w_3, N(w_2), w_2, w_1 \right\}$, so that

$$\mathfrak{B}(\mathbb{F}^6) = \left\{ \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix} \right\}$$

and

$$\mathcal{M}(N, \mathfrak{B}(\mathbb{F}^6)) = \left[\begin{array}{ccc|cc|c} \begin{bmatrix} 0 & 1 & 0 \\ & 0 & 1 \\ & & 0 \end{bmatrix} & & & & & \\ \hline & & & \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix} & & \\ \hline & & & & & [0] \end{array} \right]$$

Basis Corresponding to a Nilpotent Operator

This theorem formalizes what we have demonstrated in the examples:

Theorem (Basis Corresponding to a Nilpotent Operator)

Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then there exist vectors $v_1, \ldots, v_n \in V$ and nonnegative integers $m_1, \ldots, m_n$ such that

(a) $N^{m_1}(v_1), \ldots, N(v_1), v_1, \ldots, N^{m_n}(v_n), \ldots, N(v_n), v_n$ *is a basis of* V,

(b) $N^{m_1+1}(v_1) = \cdots = N^{m_n+1}(v_n) = 0$.

Proof (Basis Corresponding to a Nilpotent Operator)

[Induction–Base] If $\dim(V) = 1$, the only nilpotent operator is 0, let $v \neq 0$, and $m_1 = 0$.

[Induction–Hypothesis] Assume $n = \dim(V) > 1$ and the theorem holds on all spaces of smaller dimension.

Basis Corresponding to a Nilpotent Operator

Proof (Basis Corresponding to a Nilpotent Operator)

Because N is nilpotent, N is not injective $\Rightarrow$ not surjective [For $\mathcal{L}(V)$: Injectivity $\Leftrightarrow$ Surjectivity in Finite Dimensions (Notes#3.2)] and hence $\dim(\text{range}(N)) < \dim(V)$. Thus we can apply our inductive hypothesis to $N|_{\text{range}(N)} \in \mathcal{L}(\text{range}(N))$.

By [Induction–Hypothesis] applied to $N|_{\text{range}(N)}$ there exist vectors $v_1, \ldots, v_n \in \text{range}(N)$ nonnegative integers $m_1, \ldots, m_n$ such that

$$N^{m_1}(v_1), \ldots, N(v_1), v_1, \ldots, N^{m_n}(v_n), \ldots, N(v_n), v_n \qquad \text{(i)}$$

is a basis of $\text{range}(N)$, and $N^{m_1+1}(v_1) = \cdots = N^{m_n+1}(v_n) = 0$

Basis Corresponding to a Nilpotent Operator

Proof (Basis Corresponding to a Nilpotent Operator)

Since $(\forall \ell)\ v_\ell \in \text{range}(N)\ \exists u_\ell \in V : v_\ell = N(u_\ell)$; thus $N^{k+1}u_\ell = N^k v_\ell$.
We use this to rewrite and augment (i):

$$N^{m_1+1}(u_1), \ldots, N(u_1), u_1, \ldots, N^{m_n+1}(u_n), \ldots, N(u_n), u_n \qquad \text{(ii)}$$

We need to verify that this is a list of linearly independent vectors.

Assume some linear combination of the vectors in (ii) equals zero; apply N to that linear combination; this yields a linear combination of the vectors in (i) equal to zero. Since those vectors are linearly independent, the coefficients multiplying the vectors in the set (i) must be zero.

What remains is a linear combination of

$$\left\{ N^{m_1+1}(u_1), \ldots, N^{m_n+1}(u_n) \right\} = \left\{ N^{m_1}(v_1), \ldots, N^{m_n}(v_n) \right\}$$

which is a subset of (i), and hence those coefficients are also zero. $\Rightarrow$ the list in (ii) is linearly independent.

Basis Corresponding to a Nilpotent Operator

Proof (Basis Corresponding to a Nilpotent Operator)

Next, we extend (ii) to a basis of V [Linearly Independent List Extends to a Basis (Notes#2)]: (we need coverage for $\mathrm{null}(N) \cap \mathrm{range}(N)^\perp$)

$$N^{m_1+1}(u_1), \ldots, N(u_1), u_1, \ldots, N^{m_n+1}(u_n), \ldots, N(u_n), u_n, w_1, \ldots, w_p \qquad \text{(iii)}$$

Each $N(w_k) \in \mathrm{range}(N) \Rightarrow N(w_k) \in \mathrm{span}(i) = \mathrm{span}(N(ii))$

We can find $x_\ell \in \mathrm{span}(ii)$ so that $N(w_\ell) = N(x_\ell)$; let $u_{n+\ell} = w_\ell - x_\ell \neq 0$. By construction $N(u_{n+\ell}) = 0$, and

$$N^{m_1+1}(u_1), \ldots, N(u_1), u_1, \ldots, N^{m_n+1}(u_n), \ldots, N(u_n), u_n, u_{n+1}, \ldots, u_{n+p} \qquad \text{(iv)}$$

spans V, because its span contains each x_ℓ and each $u_{n+\ell}$ and hence each w_ℓ and (iii) spans V.

(iv) has the same length as (iii), so we have a basis with the desired properties.

Jordan Basis ⇝ Jordan Form

Definition (Jordan Basis)

Suppose $T \in \mathcal{L}(V)$. A basis of V is called a **Jordan basis**, $\mathfrak{J}(V)$ for T if wrt this basis T has a block diagonal matrix, where each block A_k is upper-triangular with diagonal entries λ_k, and first super-diagonal entries all ones:

$$\mathcal{M}(T; \mathcal{J}(V)) = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{bmatrix}, \quad A_k = \begin{bmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{bmatrix}$$

Theorem (Jordan Form)

Suppose V is a complex vector space. If $T \in \mathcal{L}(V)$, then there is a basis of V that is a Jordan basis for T.

Jordan Basis ⇝ Jordan Form

Proof (Jordan Form)

First consider a nilpotent operator $N \in \mathcal{L}(V)$, and the vectors $v_1, \ldots, v_n \in V$ given by [Basis Corresponding to a Nilpotent Operator].

For each k, N sends the first vector in the list $N^{m_k}(v_k), \ldots, N(v_k), v_k$ to 0, and "left-shifts" the other vectors in the list. That is, [Basis Corresponding to a Nilpotent Operator] gives a basis of V wrt which N has a block diagonal matrix, where each matrix on the diagonal has the form

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

Thus the theorem holds for nilpotent operators...

Jordan Basis ⇝ Jordan Form

Proof (Jordan Form)

Now suppose $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. We have the generalized eigenspace decomposition

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

where each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent [Description of Operators on Complex Vector Spaces]. Thus some basis of each $G(\lambda_k, T)$ is a Jordan basis for $(T - \lambda_k I)|_{G(\lambda_k, T)}$. Put these bases together to get a basis of V that is a Jordan basis for T. $\surd$

Jordan Basis $\rightsquigarrow$ Jordan Form

Example (Jordan Form)

Consider

$$A = \begin{bmatrix} 177 & 548 & 271 & -548 & -356 \\ 19 & 63 & 14 & -79 & -23 \\ 8 & 24 & 17 & -20 & -20 \\ 42 & 132 & 55 & -141 & -76 \\ 56 & 176 & 80 & -184 & -105 \end{bmatrix}$$

We try to find the minimal polynomial; we "randomly" select $v = (1, 0, 0, 0, 0)$, and form $B = \left[v\ Av\ A^2v\ A^3v\ A^4v\ A^5v\right]$:

$$B = \begin{bmatrix} 1 & 177 & 957 & 4245 & 16761 & 62457 \\ 0 & 19 & 66 & 273 & 996 & 3567 \\ 0 & 8 & 48 & 216 & 864 & 3240 \\ 0 & 42 & 204 & 894 & 3480 & 12882 \\ 0 & 56 & 288 & 1272 & 4992 & 18552 \end{bmatrix}$$

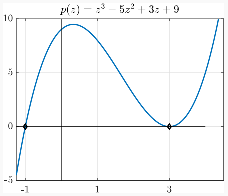
We need to find the first column which is linearly dependent on the previous; hence, we row-reduce:

Jordan Basis $\rightsquigarrow$ Jordan Form

Example (Jordan Form)

$$\mathrm{rref}(B) = \begin{bmatrix} 1 & 0 & 0 & -9 & -45 & -198 \\ 0 & 1 & 0 & -3 & -24 & -111 \\ 0 & 0 & 1 & 5 & 22 & 86 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence, $-9v - 3Av + 5A^2v = A^3v$; and we have a candidate for the minimal polynomial: $p(z) = z^3 - 5z^2 + 3z + 9 = (z+1)(z-3)^2$.



The way we have done it – using a "random" vector to start the problem, we are NOT guaranteed that this is the minimal polynomial. However, applying the polynomial to the full original matrix will give us the answer; in this case, indeed $A^3 - 5A^2 + 3A + 9I_5 = 0$.

If the test had failed: $p(z)$ would have been one of the factors in the polynomial (so the work would not have been completely wasted). Another "random" guess (not a linear combination of the vectors in B) would be needed to identify more factors.

Jordan Form

Example (Jordan Form)

First we compute the eigenspaces

$$\begin{aligned} E(-1, A) &= \mathrm{null}(A + 1I_5) = \mathrm{span}((2,1,0,1,1)), \\ E(3, A) &= \mathrm{null}(A - 3I_5) = \mathrm{span}\left((-19, 14, -6, 5, 0)), ((24, -7, 4, 0, 4)\right). \end{aligned}$$

Clearly, these 3 vector cannot span $\mathbb{C}^5$, we need genereralized eigenspaces...

$$\begin{aligned} \mathrm{null}\left((A + 1I_5)^2\right) &= \mathrm{null}(A + 1I_5) \quad \Rightarrow G(-1, A) = E(-1, A) \\ \mathrm{null}\left((A - 3I_5)^2\right) &= \mathrm{span}\left(\begin{bmatrix}-8\\3\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}11\\0\\0\\3\\0\end{bmatrix}, \begin{bmatrix}4\\0\\0\\0\\3\end{bmatrix}\right) \end{aligned}$$

Whereas, technically $G(3, A) = \mathrm{null}\left((A - 3I_5)^5\right)$, there can be no growth beyond this point (including $G(-1, A)$ we already have 5 vectors).

Jordan Form

Example (Jordan Form)

The null-space dimension of nilpotent "$\mathcal{M}(((T - \lambda_\ell I)^k)|_{G(\lambda_\ell, T)})$" matrix-blocks equal to k; hence the differences

- $\dim(\mathrm{null}\left((A - 3I_5)^2\right)) - \dim(\mathrm{null}\left((A - 3I_5)^1\right)) = 4 - 2 = 2$ tells us that we have 2 blocks of size 2 or larger; and
- $\dim(\mathrm{null}\left((A - 3I_5)^3\right)) - \dim(\mathrm{null}\left((A - 3I_5)^2\right)) = 4 - 4 = 0$ tells us that we have 0 blocks of size 3 or larger.

At this point we know the Jordan Form of A:

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

What remains is figuring out the basis $\mathfrak{B}(V)$ which gets us there.

Jordan Form

Example (Jordan Form)

We apply $(A-3I_5)^k$, $k=0,\ldots$ to each vector in $G(3,A)$ to form 4 Jordan Chains:

$$\left\{\begin{bmatrix}252\\28\\8\\60\\80\end{bmatrix},\begin{bmatrix}-8\\3\\0\\0\\0\end{bmatrix}\right\},\left\{\begin{bmatrix}271\\14\\14\\55\\80\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}\right\},\left\{\begin{bmatrix}270\\-28\\28\\30\\64\end{bmatrix},\begin{bmatrix}11\\0\\0\\3\\0\end{bmatrix}\right\},\left\{\begin{bmatrix}372\\7\\-28\\-60\\-100\end{bmatrix},\begin{bmatrix}4\\0\\0\\0\\3\end{bmatrix}\right\}.$$

We form a basis using the vectors from 2 of the chains, and $E(-1,A)$, *e.g*

$$\mathfrak{B}(V)=\operatorname{span}\left(\begin{bmatrix}2\\1\\0\\1\\1\end{bmatrix},\begin{bmatrix}252\\28\\8\\60\\80\end{bmatrix},\begin{bmatrix}-8\\3\\0\\0\\0\end{bmatrix},\begin{bmatrix}271\\14\\14\\55\\80\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}\right)$$

Jordan Form

Example (Jordan Form)

... and we arrive:

$$A=\begin{bmatrix}177&548&271&-548&-356\\19&63&14&-79&-23\\8&24&17&-20&-20\\42&132&55&-141&-76\\56&176&80&-184&-105\end{bmatrix}$$

$$\mathfrak{B}(V)=\operatorname{span}\left(\begin{bmatrix}2\\1\\0\\1\\1\end{bmatrix},\begin{bmatrix}252\\28\\8\\60\\80\end{bmatrix},\begin{bmatrix}-8\\3\\0\\0\\0\end{bmatrix},\begin{bmatrix}271\\14\\14\\55\\80\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}\right)$$

$$\mathcal{M}(A,\mathfrak{B}(V))=\left[\begin{array}{c|cc|cc}-1&&&&\\\hline&3&1&&\\&&3&&\\\hline&&&3&1\\&&&&3\end{array}\right]$$

⟨⟨⟨ Live Math ⟩⟩⟩

e.g. 8D-{3, 4, 5}

Suggested Problems

8.A—1, 2, 3, 4, 5

8.B—1, 2, 3, 4, 5

8.C—1, 2, 3, 4, 5

8.D—1, 2, 3, 4, 5

Assigned Homework — HW#8, Not Due

Strongly Suggested Problems

8.A—1, 2

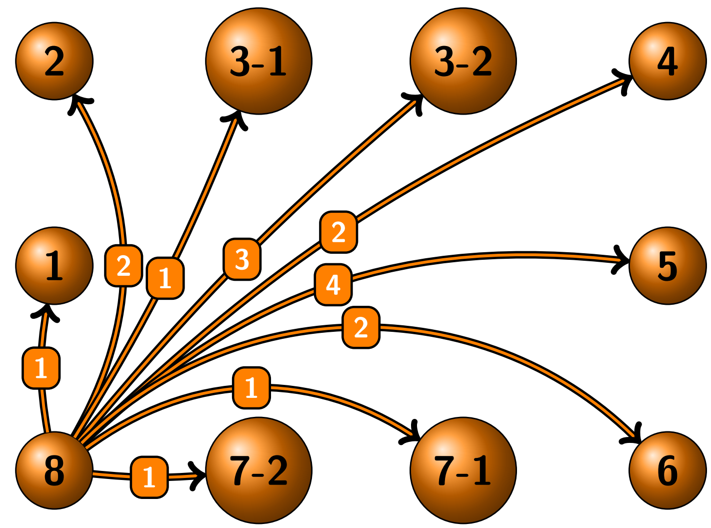
8.B—1, 2

8.C—1, 2

8.D—1, 2

Expect variants on the take-home and in-class finals.

Explicit References to Previous Theorems or Definitions (with count)



Explicit References to Previous Theorems or Definitions