18.700 - LINEAR ALGEBRA, DAY 21
GENERALIZED EIGENVECTORS
GENERALIZED EIGENSPACE DECOMPOSITION

SAM SCHIAVONE

CONTENTS

I. PRE-CLASS PLANNING

I.1. **Goals for lesson.**

(1) Students will learn the definition of generalized eigenvector.
(2) Students will learn the definition of a generalized eigenspace.
(3) Students will learn the general eigenspace decomposition theorem.
(4) Students will learn the Cayley-Hamilton theorem.

I.2. **Methods of assessment.**

(1) Student responses to questions posed during lecture
(2) Student responses to worksheet

I.3. **Materials to bring.** (1) Laptop + adapter (2) Worksheets (3) Chalk

(0:00)

II. Lesson Plan

II.1. **Last time.**

- Defined Singular Value Decomposition.
- Learned how to compute an SVD of a matrix.

II.2. **8A: Generalized Eigenvectors, cont.**

II.2.1. *Generalized eigenvectors.* Recall that T is diagonalizable iff

$$V = V_1 \oplus \cdots \oplus V_n \tag{1}$$

where $V_i = \operatorname{span}(v_i)$ is a 1-dimensional T-invariant subspace for each $i = 1, \ldots, n$. But we know that not every linear operator T is diagonalizable: e.g., $V = \mathbb{F}^2$ and $T = L_A$ for

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

But what if we allow for T-invariant subspaces of larger dimension in (1)? This leads to the following notion.

Definition 2. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. A vector $v \in V$ is a *generalized eigenvector* of T associated to λ if $v \neq 0$ and

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

for some $k \in \mathbb{Z}_{>0}$.

Theorem 3. *Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.*

Proof. Let $n := \dim(V)$. By strong induction on n.

Base case: $n = 1$. Then every nonzero vector is an eigenvector of T.

Inductive step: Suppose $n \geq 2$ and the result holds for all $k < n$. Since $\mathbb{F} = \mathbb{C}$, then there exists an eigenvalue λ of T. Recall then that

$$V = \ker(T - \lambda I)^n \oplus \operatorname{img}(T - \lambda I)^n .$$

Case 1: $\ker(T - \lambda I)^n = V$. Then every nonzero vector in V is a generalized eigenvector of T, so the result holds.

Case 2: $\ker(T - \lambda I)^n \neq V$. Then $\operatorname{img}(T - \lambda I)^n \neq \{0\}$. Since λ is an eigenvalue of T, then $\dim(\ker(T - \lambda I)^n) \geq 1$. Thus

$$0 < \dim(\operatorname{img}(T - \lambda I)^n) < n .$$

Let $U := \operatorname{img}(T - \lambda I)^n$. Then U is a nonzero T-invariant subspace of dimension $< n$. By the inductive hypothesis applied to the restriction $T|_U$, there is a basis of U consisting of generalized eigenvectors of $T|_U$. Adjoin this basis to a basis of $\ker(T - \lambda I)^n$. Since

$$V = \ker(T - \lambda I)^n \oplus \operatorname{img}(T - \lambda I)^n ,$$

then the result is a basis of V consisting of generalized eigenvectors of T. □

Lemma 4. *Suppose $T \in \mathcal{L}(V)$. Then each generalized eigenvector of T corresponds to a unique eigenvalue of T.*

Proof. Exercise. □

Proposition 5. *Suppose $T \in \mathcal{L}(V)$. Then every list of generalized eigenvectors of T corresponding to distinct eigenvalues is linearly independent.*

Proof. Exercise. Similar to the proof for eigenvectors. □

II.2.2. *Nilpotent operators.*

Definition 6. An operator $T \in \mathcal{L}(V)$ is *nilpotent* if $T^m = 0$ for some $m \in \mathbb{Z}_{\geq 0}$.

Example 7. Let $V = \mathbb{F}^2$. The operator $T := L_A$ with

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is nilpotent.

Proposition 8. *Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then* $\operatorname{minpoly}(T) = z^m$ *for some* $m \leq \dim(V)$.

Proof. Since T is nilpotent, then $T^k = 0$ for some $k \in \mathbb{Z}_{\geq 0}$. Then $\operatorname{minpoly}(T)$ divides z^k so

$$z^k = \operatorname{minpoly}(T) f(z)$$

for some $f \in \mathcal{P}(\mathbb{F})$. By unique factorization, then $\operatorname{minpoly}(T) = z^m$ for some $m \in \mathbb{Z}_{\geq 0}$. Moreover, we know that $\deg(\operatorname{minpoly}(T)) \leq \dim(V)$, so $m \leq \dim(V)$. □

Proposition 9. *Let $n := \dim(V)$ and suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $T^n = 0$.*

Proof. By the above, $\operatorname{minpoly}(T) = z^m$ for some $m \leq n$. Then

$$T^n = T^{n-m} T^m = T^{n-m} \circ 0 = 0\,.$$

□

Proposition 10. *Suppose $T \in \mathcal{L}(V)$.*

(a) If T is nilpotent, then 0 is an eigenvalue of T and T has no other eigenvalues.
(b) If $\mathbb{F} = \mathbb{C}$ and 0 is the only eigenvalue of T, then T is nilpotent.

Proof. (a) By the previous proposition, $\operatorname{minpoly}(T) = T^m$ for some m. The eigenvalues of T are exactly the roots of $\operatorname{minpoly}(T)$.

(b) Since $\mathbb{F} = \mathbb{C}$, then $\operatorname{minpoly}(T)$ splits into degree 1 factors. Then $\operatorname{minpoly}(T) = z^m$ for some $m \in \mathbb{Z}_{\geq 0}$, so $T^m = 0$.

□

Proposition 11. *Suppose $T \in \mathcal{L}(V)$. TFAE.*

(a) T is nilpotent.
(b) $\operatorname{minpoly}(T) = z^m$ *for some* $m \in \mathbb{Z}_{\geq 0}$.
(c) There is a basis $\mathcal{B}$ of V such that

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}$$

where all entries on and below the diagonal are 0.

Proof. (a) $\implies$ (b): Already done.

(b) $\implies$ (c): Since $\operatorname{minpoly}(T) = z^m$ splits into degree 1 factors, then there is a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is upper triangular. Its diagonal entries are exactly the eigenvalues of T, namely 0, so we obtain a matrix of the desired form.

(c) $\implies$ (a): A direct calculation shows that $([T]_{\mathcal{B}})^n = 0$. □

II.3. **Generalized eigenspace decomposition.**

Definition 12. Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. The *generalized eigenspace* of T corresponding to λ, denoted $G_\lambda(T)$ or G_λ, is

$$G_\lambda(T) := \left\{v \in V : (T - \lambda I)^k(v) = 0 \text{ for some } k \in \mathbb{Z}_{>0}\right\}.$$

In other words, G_λ is the set of generalized eigenvectors of T corresponding to λ, together with the 0 vector.

Proposition 13. *Let $n := \dim(V)$. Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then $G_\lambda(T) = \ker(T - \lambda I)^n$.*

Proof. ($\supseteq$): Given $v \in \ker(T - \lambda I)^n$, then $v \in G_\lambda(T)$. (Just take $k = n$ in the definition.)

($\subseteq$): Given $v \in G_\lambda(T)$, then $v \in \ker(T - \lambda I)^k$ for some k, so $(T - \lambda I)^k(v) = 0$.

<u>Case 1</u>: $k \leq n$. Then

$$(T - \lambda I)^n(v) = (T - \lambda I)^{n-k}(T - \lambda I)^k(v) = (T - \lambda I)^{n-k}(0) = 0\,,$$

so $v \in \ker(T - \lambda I)^n$.

<u>Case 2</u>: $k > n$. By a previous result, we have

$$\ker(T - \lambda I)^n = \ker(T - \lambda I)^{n+1} = \ker(T - \lambda I)^{n+2} = \cdots = \ker(T - \lambda I)^k = \cdots.$$

Thus $v \in \ker(T - \lambda I)^k = \ker(T - \lambda I)^k$. □

Proposition 14 (Generalized eigenspace decomposition)**.** *Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then*

(a) $G_{\lambda_k}(T)$ is T-invariant for all $k = 1, \ldots, m$;
(b) $(T - \lambda_k I)|_{G_{\lambda_k}}$ is nilpotent for all $k = 1, \ldots, m$;
(c) $V = G_{\lambda_1} \oplus \cdots \oplus G_{\lambda_m}$.

Proof. Let $n = \dim(V)$.

(a) Fix $k \in \{1, \ldots, m\}$. By a previous result, then

$$G_{\lambda_k}(T) = \ker((T - \lambda_k I)^n)\,.$$

This is the kernel of a polynomial evaluated at T, hence is T-invariant.

(b) Fix $k \in \{1, \ldots, m\}$. Given $v \in G_{\lambda_k}(T) = \ker((T - \lambda_k I)^n)$, then

$$(T - \lambda_k I)^n(v) = 0\,.$$

Then

$$(T - \lambda_k I)^n|_{G_{\lambda_k}} = 0$$

so $(T - \lambda_k I)|_{G_{\lambda_k}}$ is nilpotent.

(c) By a previous result, there is a basis of V consisting of generalized eigenvectors of T. Thus every vector in V can be written as a linear combination of generalized eigenvectors, so

$$G_{\lambda_1} + \cdots + G_{\lambda_m} = V\,.$$

We now show the sum is direct. Suppose that

$$v_1 + \cdots + v_m = 0$$

with $v_k \in G_{\lambda_k}$ for each $k = 1, \ldots, m$. Since generalized eigenvectors corresponding to distinct eigenvalues are linearly independent, then we have $v_1 = \cdots = v_m = 0$. Thus the sum is direct.

□

Definition 15. Let $T \in \mathcal{L}(V)$.

- The *geometric multiplicity* of an eigenvalue λ of T is
$$\dim(E_\lambda(T)) = \dim(\ker(T - \lambda I)\,.$$
- The *generalized multiplicity* (or *algebraic multiplicity*) of an eigenvalue λ of T is
$$\dim(G_\lambda(T)) = \dim(\ker(T - \lambda I)^{\dim(V)}\,.$$

Example 16. Let

$$A := \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix}\,.$$

Then

$$A - 2I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -5 \end{pmatrix}\,,$$

which has 2 pivots, so $\dim(E_2) = 1$. Now

$$(A - 2I)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -125 \end{pmatrix}\,,$$

so $\dim(G_2) = 2$. Thus 2 has geometric multiplicity 1, and algebraic multiplicity 2.

Definition 17. Suppose $\mathbb{F} = \mathbb{C}$ and $T = \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with algebraic multiplicities $d_1, \ldots, d_m$. The *characteristic polynomial of T* is

$$\operatorname{charpoly}(T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}\,.$$

Remark 18. We will later give a formula for $\operatorname{charpoly}(T)$ that doesn't require knowing the eigenvalues of T. (It will involv determinants.)

Proposition 19. *Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then*

(a) $\operatorname{charpoly}(T)$ *has degree* $\dim(V)$*; and*
(b) the zeroes of $\operatorname{charpoly}(T)$ *are exactly the eigenvalues of T.*

Proof. (a) Recall that the algebraic multiplicity of λ_k is $\dim(G_{\lambda_k}(T))$. Since

$$V = G_{\lambda_1} \oplus \cdots \oplus G_{\lambda_m},$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T, then

$$\dim(V) = \dim(G_{\lambda_1}) + \cdots + \dim(G_{\lambda_m}).$$

(b) Immediate from the definition.

□

Theorem 20 (Cayley-Hamilton)**.** *Suppose* $\mathbb{F} = \mathbb{C}$. *Suppose* $T \in \mathcal{L}(V)$ *and let* $q = \text{charpoly}(T)$. *Then* $q(T) = 0$ *(i.e., the zero linear map).*

Proof. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, and let $d_k := \dim(G_{\lambda_K})$ be the algebraic multiplicity of λ_k for $k = 1, \ldots, m$. For each k, we have seen that $(T - \lambda_k I)|_{G_{\lambda_k}}$ is nilpotent, so

$$(T - \lambda_k I)^{d_k}|_{G_{\lambda_k}}.$$

By the generalized eigenspace decomposition, each vector $v \in V$ can be written as $v = v_1 + \cdots + v_m$ with $v_k \in G_{\lambda_k}$ for each k. Thus to show that $q(T) = 0$, it suffices to show $q(T)|_{G_{\lambda_k}} = 0$ for each k.

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

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

Recall that polynomials in T commute, so we can change the order of the factors above. Thus

$$\begin{aligned} q(T)|_{G_k} &= (T - \lambda_1 I)^{d_1} \cdots (T - \lambda_{k-1})^{d_{k-1}} (T - \lambda_{k+1})^{d_{k+1}} \cdots (T - \lambda_m I)^{d_m}|_{G_k} \underbrace{(T - \lambda_k I)^{d_k}|_{G_k}}_{0} \\ &= 0. \end{aligned}$$

□

Proposition 21. *Suppose* $\mathbb{F} = \mathbb{C}$ *and* $T \in \mathcal{L}(V)$. *Then* $\text{minpoly}(T)$ *divides* $\text{charpoly}(T)$, *i.e.,*

$$\text{charpoly}(T) = \text{minpoly}(T)\, f(z)$$

for some $f(z) \in \mathcal{P}(\mathbb{F})$.

Proof. Letting $q := \text{charpoly}(T)$, then $q(T) = 0$. By a previous result, then $\text{minpoly}(T)$ must divide q. □

Proposition 22. *Suppose* $\mathbb{F} = \mathbb{C}$ *and* $T \in \mathcal{L}(V)$. *Let* $\mathcal{B}$ *be a basis of* V *such that* $[T]_{\mathcal{B}}$ *is upper triangular. For each eigenvalue* λ *of* T, *then number of times that* λ *appears on the diagonal of* $[T]_{\mathcal{B}}$ *is equal to the algebraic multiplicity of* λ.

Proof. Let $A := [T]_{\mathcal{B}}$. [Write out A with its entries.] Then for each k we have

$$T(v_k) = \overbrace{c_1 v_1 + \cdots + c_{k-1} v_{k-1}}^{u_k} + \lambda_k v_k,$$

where $u_k \in \text{span}(v_1, \ldots, v_{k-1})$. Thus if $\lambda_k \neq 0$, then $T(v_k)$ is not a linear combination of $T(v_1), \ldots, T(v_{k-1}) \in \text{span}(v_1, \ldots, v_{k-1})$. By the Linear Dependence Lemma, then the collection of $T(v_k)$ such that $\lambda_k \neq 0$ is linearly independent.

Let d be the number of indices $k \in \{1, \ldots, n\}$ such that $\lambda_k = 0$. By the above, then

$$n - d \leq \dim(\operatorname{img}(T)) = \dim(V) - \dim(\ker(T)) = n - \dim(\ker(T))$$

by Rank-Nullity. Then $\dim(\ker(T)) \leq d$.

Now, note that $[T^n]_{\mathcal{B}} = [T]_{\mathcal{B}}^n = A^n$. Moreover, the diagonal entries of A^n are $\lambda_1^n, \ldots, \lambda_n^n$. Since $\lambda_k^n = 0$ iff $\lambda_k = 0$, then 0 appears on the diagonal of A^n d times, too. Thus the reasoning above applies just as well to T^n, so we have

$$\dim(\ker(T^n)) \leq d\,. \tag{23}$$

For each eigenvalue λ of T, let m_λ denote the algebraic multiplicity of λ, and let d_λ be the number of times λ appears on the diagonal of A. Replacing T with $T - \lambda I$ in (23), then

$$m_\lambda \leq d_\lambda \tag{24}$$

for each eigenvalue λ of T. Summing over all eigenvalues λ, we have [start in middle]

$$n = \dim(V) = \sum_\lambda m_\lambda \leq \sum_\lambda d_\lambda = n$$

where the second equality follows from the generalized eigenspace decomposition, and the last equality from the fact that the diagonal of A consists of n entries.

Thus the inequality in (24) must in fact be an equality for all eigenvalues λ. □

Definition 25. A *block diagonal matrix* is a square matrix of the form

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

where $A_1, \ldots, A_m$ are square matrices (of possibly different sizes) lying on the diagonal, and all other entres are 0.

Example 26 (Give example.)**.**

Proposition 27. *Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with algebraic multiplicities $d_1, \ldots, d_m$. Then there is a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is block diagonal*

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

where each A_k is a $d_k \times d_k$ upper triangular matrix of the form

$$A_k := \begin{pmatrix} \lambda_k & * & \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}.$$

Proof. By a previous result, $(T - \lambda_k I)|_{G_{\lambda_k}}$ is nilpotent for each k. Thus for each k we can choose a basis $\mathcal{B}_k$ such that $[(T - \lambda_k I)|_{G_{\lambda_k}}]_{\mathcal{B}_k}$ is strictly upper triangular. [Draw picture.] Now,

$$T|_{G_{\lambda_k}} = (T - \lambda_k I)|_{G_{\lambda_k}} + \lambda_k I|_{G_{\lambda_k}}$$

so

$$[T|_{G_{\lambda_k}}]_{\mathcal{B}} = [(T - \lambda_k I)|_{G_{\lambda_k}}]_{\mathcal{B}} + [\lambda_k I|_{G_{\lambda_k}}]_{\mathcal{B}}$$

[draw picture below].

This deals with a single block. Now concatenate the bases $\mathcal{B}_1, \ldots, \mathcal{B}_m$ to form a basis $\mathcal{B}$ of V. Then $[T]_{\mathcal{B}}$ is of the desired form. □