12.2 Generalized Eigenvectors

March 30, 2020

We saw last time in Section 12.1 that a simple linear operator $A \in M_n(\mathbb{C})$ has the *spectral decomposition*

$$A = \sum_{i=1}^{n} \lambda_i P_i$$

where $\lambda_1, \ldots, \lambda_n$ are the distinct eigenvalues of A and $P_i \in \mathscr{L}(\mathbb{C}^n)$ is the eigenprojection onto the eigenspace

$$\mathscr{N}(\lambda_i I - A) = \mathscr{R}(P_i).$$

Something similar holds for semisimple A.

When A is not semisimple, there are not enough eigenvectors to form an eigenbasis; we must look for *generalized* eigenspaces that contains the eigenspaces in order to find something like the spectral decomposition of A.

Throughout we assume that V is a finite dimensional vector space over $\mathbb{F}$, which we know means that V is isomorphic to $\mathbb{F}^n$ for $n = \dim(V)$.

When we speak of a linear operator A on V we will mean a linear operator on $\mathbb{F}^n$, i.e., $A \in M_n(\mathbb{F})$.

Recall from Exercise 2.8 that for a linear operator B on any vector space (this includes infinite dimensional) we have the increasing sequence or *ascending chain* of subspaces

$$\mathscr{N}(B) \subset \mathscr{N}(B^2) \subset \cdots \subset \mathscr{N}(B^k) \subset \cdots .$$

When V is finite dimensional, the ascending chain stabilizes, i.e., there exists $K \in \mathbb{N}$ such that for all $k \geq K$ there holds $\mathscr{N}(B^k) = \mathscr{N}(B^{k+1})$, because the the nondecreasing sequence of dimensions $(\dim(\mathscr{N}(B^l)))_{l=0}^{\infty}$ is bounded above by $\dim(V)$ (proof of this upper bound is HW Exercise 12.6), where we understand $B^0 = I$.

Definition 12.2.1. The index of $B \in M_n(\mathbb{F})$, denoted by $\operatorname{ind}(B)$, is the smallest $k \in \{0, 1, 2, 3, \ldots\}$ such that

$$\mathcal{N}(B^k) = \mathcal{N}(B^{k+1}).$$

Example 12.2.2. If $B \in M_n(\mathbb{F})$ is invertible, i.e., $\det(B) \neq 0$, then $\mathcal{N}(B^l) = \{0\}$ for all $l = 0, 1, 2, 3, \ldots$.

Thus for invertible B we have $\operatorname{ind}(B) = 0$. To get a positive index requires that B is not invertible.

Theorem 12.2.3. If $\operatorname{ind}(B) = k$, then for all $l \geq k$ there holds

$$\mathcal{N}(B^l) = \mathcal{N}(B^{l+1}),$$

and each of the inclusions

$$\mathcal{N}(B^l) \subset \mathcal{N}(B^{l+1})$$

is proper for all $l = 0, \ldots, k-1$.

Proof. The finite dimensionality of $\mathbb{F}^n$ implies that only finite many of the inclusions in the ascending chain

$$\mathscr{N}(B) \subset \mathscr{N}(B^2) \subset \cdots \subset \mathscr{N}(B^k) \subset \cdots$$

can be proper.

You showed in Exercise 2.12, that if $\mathscr{N}(B^l) = \mathscr{N}(B^{l+1})$ for some $l = 0, 1, 2, 3, \ldots$, then $\mathscr{N}(B^j) = \mathscr{N}(B^{j+1})$ for all $j \geq l$.

Thus with $k = \mathrm{ind}(B)$ being the smallest value for which $\mathscr{N}(B^k) = \mathscr{N}(B^{k+1})$, we obtain for all $l \geq k$ that

$$\mathscr{N}(B^l) = \mathscr{N}(B^{l+1}),$$

and for all $l = 0, \ldots, k-1$ that the inclusions

$$\mathscr{N}(B^l) \subset \mathscr{N}(B^{l+1})$$

are proper. $\square$

What questions do you have?

Example (in lieu of 12.2.4). For the matrix

$$B = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we have

$$\mathscr{N}(B) = \mathrm{span}\{\mathrm{e}_2\} \text{ and } \mathscr{R}(B) = \mathrm{span}\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}.$$

Since

$$B^2 = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we have

$$\mathscr{N}(B^2) = \mathrm{span}\{\mathrm{e}_2, \mathrm{e}_3\} \text{ and } \mathscr{R}(B^2) = \mathrm{span}\{\mathrm{e}_1, \mathrm{e}_2\}.$$

Since

$$B^3 = \begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 27 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we have

$$\mathscr{N}(B^3) = \mathrm{span}\{\mathrm{e}_2, \mathrm{e}_3, \mathrm{e}_4\} \text{ and } \mathscr{R}(B^3) = \{\mathrm{e}_1\}.$$

Since for all $l \geq 3$ we have

$$B^l = \begin{bmatrix} 3^l & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

for $l \geq 3$ we have

$$\mathscr{N}(B^l) = \mathrm{span}\{\mathrm{e}_2, \mathrm{e}_3, \mathrm{e}_4\} \text{ and } \mathscr{R}(B^l) = \mathrm{span}\{\mathrm{e}_1\}.$$

This gives $\mathrm{ind}(B) = 3$.

Summarizing, for

$$B = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we have

$$\mathscr{N}(B) = \mathrm{span}\{\mathrm{e}_2\} \text{ and } \mathscr{R}(B) = \mathrm{span}\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\},$$

$$\mathscr{N}(B^2) = \mathrm{span}\{\mathrm{e}_2, \mathrm{e}_3\} \text{ and } \mathscr{R}(B^2) = \mathrm{span}\{\mathrm{e}_1, \mathrm{e}_2\},$$

and for $l \geq 3$,

$$\mathscr{N}(B^l) = \mathrm{span}\{\mathrm{e}_2, \mathrm{e}_3, \mathrm{e}_4\} \text{ and } \mathscr{R}(B^l) = \mathrm{span}\{\mathrm{e}_1\}.$$

Notice also that $\mathscr{N}(B^l)$ and $\mathscr{R}(B^l)$ intersect nontrivially when $l = 1, 2$, but that these subspaces intersect trivially when $l \geq 3$.

This is not a coincidence.

What questions do you have?

Theorem 12.2.5. For $B \in M_n(\mathbb{C})$, if $k \geq \operatorname{ind}(B)$, then

$$\mathbb{C}^n = \mathscr{R}(B^k) \oplus \mathscr{N}(B^k).$$

We present an important observation in the finite dimensional case about the vectors obtained by repeated powers of a linear operator acting on a given vector.

Proposition 12.2.7. For $B \in M_n(\mathbb{C})$ and $\mathrm{x} \in \mathbb{C}^n$, if there exists $m \in \mathbb{N}$ such that $B^m\mathrm{x} = 0$ and $B^{m-1}\mathrm{x} \neq 0$, then the set

$$\{\mathrm{x}, B\mathrm{x}, \ldots, B^{m-1}\mathrm{x}\}$$

is linearly independent.

We illustrate Proposition 12.2.7 with the matrix the 4×4 we saw before.

Example. For the matrix

$$B = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and the vector $\mathrm{x} = \mathrm{e}_4$ we have

$$B\mathrm{x} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \mathrm{e}_3,$$

$$B^2\mathrm{x} = B\mathrm{e}_3 = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathrm{e}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mathrm{e}_2,$$

and $B^3\mathrm{x} = B\mathrm{e}_2 = 0$; the vectors $\{\mathrm{x}, B\mathrm{x}, B^2\mathrm{x}\} = \{\mathrm{e}_4, \mathrm{e}_3, \mathrm{e}_2\}$ are linearly independent.

Recall that the eigenspace of a linear operator $A \in M_n(\mathbb{C})$ associated to one of its eigenvalues λ is the subspace

$$\Sigma_\lambda = \mathscr{N}(\lambda I - A),$$

where the dimension of this subspace is the geometric multiplicity of λ.

If $A \in M_n(\mathbb{C})$ is semisimple (which includes the simple case) with spectrum $\sigma(A) = \{\lambda_1, \ldots, \lambda_r\}$ (the distinct eigenvalues of A), then there holds

$$\mathbb{C}^n = \mathscr{N}(\lambda_1 I - A) \oplus \mathscr{N}(\lambda_2 I - A) \oplus \cdots \oplus \mathscr{N}(\lambda_r I - A),$$

where the geometric multiplicity of each eigenspace equals the algebraic multiplicity of the corresponding eigenvalue.

Using the union of the bases for the eigenspaces of a semisimple operator A results in a diagonal matrix where the diagonal entries are the eigenvalues of A appearing according to their multiplicity.

When A is not diagonalizable, we do not have an eigenbasis for $\mathbb{C}^n$.

But for each eigenvalue $\lambda \in \sigma(A)$ the ascending chain

$$\mathscr{N}(\lambda I - A) \subset \mathscr{N}((\lambda I - A)^2) \subset \cdots \subset \mathscr{N}((\lambda I - A)^l) \subset \cdots$$

for the noninvertible linear operator $\lambda I - A$ stabilizes when

$$l = \mathrm{ind}(\lambda I - A).$$

We will show that the subspaces

$$\mathscr{N}((\lambda I - A)^{\mathrm{ind}(\lambda I - A)}),\ \lambda \in \sigma(A),$$

do give a direct sum decomposition of $\mathbb{F}^n$ and the linear operator in the corresponding basis is a block diagonal matrix.

Definition 12.2.8. For $A \in M_n(\mathbb{C})$ and $\lambda \in \sigma(A)$, the subspace

$$\mathscr{E}_\lambda = \mathscr{N}((\lambda I - A)^{\mathrm{ind}(\lambda I - A)})$$

is called the generalized eigenspace of A corresponding to λ.

Every nonzero vector in $\mathscr{E}_\lambda$ is called a generalized eigenvector of A corresponding to λ.

Through the next four lemmas we develop the theory needed to prove that the generalized eigenspaces of a linear operator on a finite dimensional vector space do indeed give a direct sum decomposition of the the vector space.

Lemma 12.2.9. For $A \in M_n(\mathbb{C})$ and $\lambda \in \sigma(A)$, the generalized eigenspace $\mathscr{E}_\lambda$ is A-invariant.

What questions do you have?

Example (in lieu of 12.2.10). The matrix

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

has two distinct eigenvalues

- $\lambda_1 = 2$ of algebraic multiplicity 3 and
- $\lambda_2 = 5$ of algebraic multiplicity 1.

Since

$$\lambda_1 I - A = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -3 \end{bmatrix},$$

there we have $\mathscr{N}(\lambda_1 I - A) = \mathrm{span}(\mathrm{e}_1)$.

Since

$$(\lambda_1 I - A)^2 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 9 \end{bmatrix},$$

we have

$$\mathscr{N}((\lambda_1 I - A)^2) = \mathrm{span}(\mathrm{e}_1, \mathrm{e}_2).$$

Since

$$(\lambda_1 I - A)^3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & -27 \\ 0 & 0 & 0 & -27 \end{bmatrix},$$

we have

$$\mathscr{N}((\lambda_1 I - A)^3) = \mathrm{span}(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3).$$

Since

$$(\lambda_1 I - A)^4 = \begin{bmatrix} 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & -27 \\ 0 & 0 & 0 & -27 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & -27 \\ 0 & 0 & 0 & -81 \\ 0 & 0 & 0 & -81 \end{bmatrix}.$$

we have

$$\mathscr{N}((\lambda_1 I - A)^4) = \mathrm{span}(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3).$$

Thus

$$\mathrm{ind}(\lambda_1 I - A) = 3,$$

and the generalized eigenspace $\mathscr{E}_{\lambda_1}$ of A corresponding to λ_1 is $\mathscr{N}((\lambda_1 I - A)^3$ and has a basis of $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$.

Is it a coincidence that $\mathrm{ind}(\lambda_1 I - A)$ is equal to the algebraic multiplicity of λ_1?

Is it a coincidence that $\dim(\mathscr{E}_{\lambda_1})$ is equal to the algebraic multiplicity of λ_1?

Since

$$\lambda_2 I - A = \begin{bmatrix} 3 & -1 & 0 & 0 \\ 0 & 3 & -1 & 0 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

the (generalized) eigenspace $\mathscr{E}_{\lambda_2}$ has a basis of

$$\begin{bmatrix} 1 \\ 3 \\ 9 \\ 9 \end{bmatrix}.$$

Recall that $\mathscr{E}_{\lambda_1} = \mathrm{span}(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3)$.

Is it a coincidence that $\mathscr{E}_{\lambda_1} \cap \mathscr{E}_{\lambda_2} = \{0\}$?

Example. It is a coincidence when the index equals the algebraic multiplicity as shown here.

For the eigenvalue $\lambda_1 = 2$ with algebraic multiplicity 3 of the matrix

$$A = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

we have

$$(\lambda_1 I - A)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}, (\lambda_1 I - A)^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -27 \end{bmatrix},$$

so that $\mathscr{N}((\lambda_1 I - A)^2) = \mathscr{N}((\lambda_1 I - A)^3)$, implying

$$\mathrm{ind}(\lambda_1 I - A) = 2 \neq 3.$$

What questions do you have?

Lemma 12.2.11. If λ and μ are distinct eigenvalues of $A \in M_n(\mathbb{C})$, then

$$\mathscr{E}_\lambda \cap \mathscr{E}_\mu = \{0\}.$$

[Not a coincidence that generalized eigenspaces for distinct eigenvalues intersect trivially.]

Lemma 12.2.12. For $A \in M_n(\mathbb{C})$, suppose W_1 and W_2 are A-invariant subspaces of $\mathbb{C}^n$ with

$$W_1 \cap W_2 = \{0\}.$$

If, for $\lambda \in \sigma(A)$, the generalized eigenspace $\mathscr{E}_\lambda$ satisfies $\mathscr{E}_\lambda \cap W_i = \{0\}$ for all $i = 1, 2$, then

$$\mathscr{E}_\lambda \cap (W_1 \oplus W_2) = \{0\}.$$

Lemma 12.2.13. For $A \in M_n(\mathbb{C})$ and $\lambda \in \sigma(A)$, the dimension of the generalized eigenspace $\mathscr{E}_\lambda$ equals the algebraic multiplicity m_λ of λ.

[Not a coincidence that $\dim(\mathscr{E}_\lambda) = m_\lambda$.]

Theorem 12.2.14. For each $A \in M_n(\mathbb{C})$ there is decomposition of $\mathbb{C}^n$ into a direct sum of A-invariant subspaces

$$\mathbb{C}^n = \bigoplus_{\lambda \in \sigma(A)} \mathscr{E}_\lambda.$$

Remark 12.2.15. Theorem 12.2.14 implies that every $A \in M_n(\mathbb{C})$ is *similar* to a block diagonal matrix where each block is the representation of A on the A-invariant $\mathscr{E}_\lambda$.

There exists a basis for each block in which the block matrix is upper triangular with the eigenvalue in each diagonal entry and either zeros or ones on the super diagonal.

With each block put into this form, we obtain what is known as the Jordan Canonical Form of A.

Although useful in theory, the Jordan Canonical Form is poorly conditioned, meaning small errors in the floating-point arithmetic can compound into large errors in the final result.

Example (Continued). Recall that the generalized eigenspaces for the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

with spectrum $\sigma(A) = \{\lambda_1 = 2, \lambda_2 = 5\}$ are

$$\mathscr{E}_{\lambda_1} = \mathscr{N}(\lambda_1 I - A) = \mathrm{span}(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3)$$

and

$$\mathscr{E}_{\lambda_2} = \mathscr{N}(\lambda_2 I - A) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 3 \\ 9 \\ 9 \end{bmatrix}\right).$$

By Theorem 12.2.14 we have

$$\mathbb{C}^4 = \mathscr{E}_{\lambda_1} \oplus \mathscr{E}_{\lambda_2}.$$

Using the bases $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ for $\mathscr{E}_{\lambda_1}$ and $\{\mathrm{v}_4\} = (1, 3, 9, 9)^{\mathrm{T}}$ for $\mathscr{E}_{\lambda_2}$ the linear operator A has a block diagonal form which is obtained by using the transition matrix C_{ST} for $S = \{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3, \mathrm{e}_4\}$, and $T = \{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3, \mathrm{v}_4\}$, i.e.,

$$C_{ST} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 9 \end{bmatrix}, \; [x]_S = C_{ST}[x]_T.$$

The block diagonal matrix similar to A is

$$C_{ST}^{-1} A C_{ST} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

[This is the Jordan Canonical Form for A; you are NOT responsible for obtaining this.]

What questions do you have?