

10.1 Generalized Eigenspaces

Chapter 9 left one gap wide open. When an operator has too few eigenvectors, its eigenspaces are too small to fill the space, and diagonalization fails. The remedy of this chapter is to enlarge each eigenspace until the enlarged pieces do fill the space. This section builds the enlargement, shows that the enlarging process stops after finitely many steps, and proves that the pieces split V whenever the characteristic polynomial splits.

10.1.1 Enlarging an eigenspace

Take the standing counterexample of Chapter 9,

$$\mathbf{J} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \in M_2(F),$$

and let $\mathbf{N} = \mathbf{J} - 2\mathbf{I}$. The only eigenvalue is 2, and $E_2(\mathbf{J}) = \ker \mathbf{N} = \text{span}(\mathbf{e}_1)$ is a line: one eigenvector direction where we need two, so $\mathbf{J}$ is not diagonalizable over any field (Theorem 9.4.1 (e)).

Now look at what $\mathbf{N}$ does to the missing direction $\mathbf{e}_2$. One application does not kill $\mathbf{e}_2$, but two do:

$$\mathbf{N}\mathbf{e}_2 = \mathbf{e}_1, \quad \mathbf{N}^2\mathbf{e}_2 = \mathbf{N}\mathbf{e}_1 = \mathbf{0}.$$

So $\mathbf{e}_2$ is killed by $\mathbf{N}^2$, and $\ker \mathbf{N}^2 = F^2$ is everything. The eigenspace was too small, but the kernel of the **square** is as big as we could wish.

This is the whole idea. Instead of asking for vectors killed by $T - \lambda \text{id}_V$, ask for vectors killed by **some power** of $T - \lambda \text{id}_V$. The eigenvectors are still there (they are killed at the first step), and there may be more.

A generalized eigenvector is a vector that some power of $T - \lambda \text{id}_V$ kills, not necessarily the first power.

10.1.2 Generalized eigenvectors and generalized eigenspaces

Definition 10.1.1. Generalized Eigenvector

Let V be a vector space over F , let $T \in \mathcal{L}(V)$ and $\lambda \in F$. A **generalized eigenvector** of T for λ is a vector $v \in V$ with $v \neq 0$ such that

$$(T - \lambda \text{id}_V)^k v = 0 \quad \text{for some integer } k \geq 1.$$

In words: the vector must be **nonzero**, exactly as for an eigenvector; the scalar λ must lie **in** F ; and the power k is not fixed in advance — it may depend on v . Taking $k = 1$ recovers the eigenvector condition, so **every eigenvector for λ is a generalized eigenvector for λ** .

Collecting all of them, together with 0, gives the space we are after.

Definition 10.1.2. Generalized Eigenspace

Let V be a vector space over F , let $T \in \mathcal{L}(V)$ and $\lambda \in F$. The **generalized eigenspace** of

T for λ is

$$\begin{aligned} G_\lambda(T) &:= \{v \in V : (T - \lambda \operatorname{id}_V)^k v = 0 \text{ for some integer } k \geq 1\} \\ &= \bigcup_{k \geq 1} \ker(T - \lambda \operatorname{id}_V)^k. \end{aligned}$$

For $\mathbf{A} \in M_n(F)$ we write $G_\lambda(\mathbf{A}) = G_\lambda(T_{\mathbf{A}})$.

So $G_\lambda(T)$ consists of 0 together with all generalized eigenvectors for λ . Writing $N = T - \lambda \operatorname{id}_V$ for the rest of this subsection, the definition says $G_\lambda(T) = \ker N \cup \ker N^2 \cup \ker N^3 \cup \dots$

Well-definedness: it really is a subspace. A union of subspaces is almost never a subspace (Proposition 1.2.3), so this needs a check. What saves us is that the union is **increasing**: if $N^k v = 0$, then $N^{k+1} v = N(N^k v) = N0 = 0$, so

$$\ker N \subseteq \ker N^2 \subseteq \ker N^3 \subseteq \dots$$

Now use the subspace test (Theorem 1.2.1). We have $0 \in \ker N \subseteq G_\lambda(T)$, so the set is nonempty. If $u, v \in G_\lambda(T)$, say $N^j u = 0$ and $N^k v = 0$, put $m = \max(j, k)$; then $u, v \in \ker N^m$ by the inclusions above, and $\ker N^m$ is a subspace, so $u + v$ and cu lie in $\ker N^m \subseteq G_\lambda(T)$ for every $c \in F$. Hence $G_\lambda(T)$ is a subspace of V .

Examples.

- **The standing counterexample.** For $\mathbf{J} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, we computed $\ker(\mathbf{J} - 2\mathbf{I}) = \operatorname{span}(\mathbf{e}_1)$ and $\ker(\mathbf{J} - 2\mathbf{I})^2 = F^2$. So $G_2(\mathbf{J}) = F^2$, while $E_2(\mathbf{J})$ is only a line. The generalized eigenspace is strictly bigger than the eigenspace, and it is all of F^2 .
- **A diagonalizable operator gains nothing.** Let $\mathbf{A} = \operatorname{diag}(2, 5)$ over $\mathbb{Q}$. Then $(\mathbf{A} - 2\mathbf{I})^k = \operatorname{diag}(0, 3^k)$, whose kernel is $\operatorname{span}(\mathbf{e}_1)$ for every $k \geq 1$. So $G_2(\mathbf{A}) = E_2(\mathbf{A}) = \operatorname{span}(\mathbf{e}_1)$, and likewise $G_5(\mathbf{A}) = E_5(\mathbf{A}) = \operatorname{span}(\mathbf{e}_2)$. Enlarging an eigenspace of a diagonalizable operator does nothing, a fact we prove in general below.
- **Differentiation.** Let $D \in \mathcal{L}(F[x]_{\leq n})$ be differentiation. Then $D^{n+1} = 0$ (Example (Differentiation Is Nilpotent)), so **every** polynomial is killed by a power of D , that is, of $D - 0 \operatorname{id}$, and $G_0(D) = F[x]_{\leq n}$. By contrast $E_0(D) = \ker D$ is the line of constants when $F = \mathbb{R}$. Here the generalized eigenspace for the single eigenvalue 0 is the whole space, of dimension $n + 1$.
- **A scalar not among the eigenvalues.** For the same $\mathbf{A} = \operatorname{diag}(2, 5)$ and $\lambda = 7$, the matrix $(\mathbf{A} - 7\mathbf{I})^k = \operatorname{diag}((-5)^k, (-2)^k)$ is invertible, so its kernel is $\{\mathbf{0}\}$ for every k . Hence $G_7(\mathbf{A}) = \{\mathbf{0}\}$. The definition makes sense for every $\lambda \in F$, and gives the zero space unless λ is an eigenvalue.
- **A degenerate case.** On a space of dimension 1, every operator is $\lambda \operatorname{id}$ for one scalar λ , and $G_\lambda = V$, $G_\mu = \{\mathbf{0}\}$ for $\mu \neq \lambda$. Small as it is, this is the model for a 1×1 Jordan block.
- **No eigenvalues at all.** The rotation R of $\mathbb{R}^2$ by a right angle has no real eigenvalue (Example (A Rotation Has No Invariant Lines)), and indeed $G_\lambda(R) = \{\mathbf{0}\}$ for every $\lambda \in \mathbb{R}$: if $(R - \lambda \mathbf{I})^k \mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$, then $R - \lambda \mathbf{I}$ would not be injective, so λ would be an eigenvalue. Over $\mathbb{C}$, where $p_R = x^2 + 1 = (x - i)(x + i)$ splits, $G_i(R)$ and $G_{-i}(R)$ are the two eigenlines.

Nonexample by minimal change. Keep $\mathbf{A} = \operatorname{diag}(2, 5)$ over $\mathbb{Q}$ and take $\mathbf{v} = (1, 1)$. This vector is killed by a polynomial in $\mathbf{A}$ whose roots are eigenvalues:

$$(\mathbf{A} - 2\mathbf{I})(\mathbf{A} - 5\mathbf{I})\mathbf{v} = \mathbf{0},$$

as one checks entry by entry. But it is a generalized eigenvector for **neither** eigenvalue: $(\mathbf{A} - 2\mathbf{I})^k \mathbf{v} = (0, 3^k) \neq \mathbf{0}$ and $(\mathbf{A} - 5\mathbf{I})^k \mathbf{v} = ((-3)^k, 0) \neq \mathbf{0}$ for every $k \geq 1$. The failing clause is that

the definition demands a power of a **single** factor $T - \lambda \text{id}_V$, not a product of different factors. (Of course $\mathbf{v} = \mathbf{e}_1 + \mathbf{e}_2$ is a *sum* of two generalized eigenvectors; that is the decomposition theorem below, not the definition.)

Why this definition. Three clauses are doing work. The vector must be nonzero, or every G_λ would contain a “generalized eigenvector” and the word would carry no information. The exponent must be allowed to grow, because nothing tells us in advance how many steps a vector needs: for D on $\mathbb{R}[x]_{\leq n}$ the polynomial x^n needs $n + 1$ steps. And the power is a power of one factor, which is what makes G_λ one piece attached to one eigenvalue. The name is the honest one: these are the vectors that behave like eigenvectors after enough applications of $T - \lambda \text{id}_V$.

Warning. $G_\lambda(T)$ is usually bigger than $E_\lambda(T)$, and “generalized eigenvector” includes the ordinary ones. The inclusion $E_\lambda(T) \subseteq G_\lambda(T)$ always holds, and it is strict exactly when T has too few eigenvectors for λ : for $\mathbf{J} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $\dim E_2 = 1$ and $\dim G_2 = 2$. So a generalized eigenvector need not satisfy $Tv = \lambda v$: the vector $\mathbf{e}_2$ satisfies $\mathbf{J}\mathbf{e}_2 = \mathbf{e}_1 + 2\mathbf{e}_2$. Conversely, do not read “generalized” as “not an eigenvector”: every eigenvector is a generalized eigenvector.

10.1.3 The kernel chain stabilizes

The definition of $G_\lambda(T)$ is an infinite union, which is awkward: to decide membership we would have to try every k . In finite dimension this is an illusion. The chain $\ker N \subseteq \ker N^2 \subseteq \dots$ is a chain of subspaces of V , and their dimensions cannot keep growing. Once the chain stops growing, it stops forever.

Theorem 10.1.1. The Kernel Chain Stabilizes

Let V be a finite-dimensional vector space over F with $\dim V = n$, let $N \in \mathcal{L}(V)$, and write $K_k = \ker N^k$ for $k \geq 0$, so that $K_0 = \{0\}$. Then:

- (a) $K_0 \subseteq K_1 \subseteq K_2 \subseteq \dots$, and every K_k is N -invariant;
- (b) if $K_m = K_{m+1}$ for some $m \geq 0$, then $K_{m+j} = K_m$ for **every** $j \geq 0$;
- (c) there is a **least** index $s \geq 0$ with $K_s = K_{s+1}$; it satisfies $s \leq n$, the inclusions

$$K_0 \subsetneq K_1 \subsetneq \dots \subsetneq K_s$$

are **strict**, and $K_k = K_s$ for every $k \geq s$;

- (d) in particular $\ker N^n = \ker N^{n+j}$ for every $j \geq 0$.

Idea. The only real point is (b): equality at one step propagates. Read it backwards. To put $v \in K_{m+j+1}$ into K_{m+j} , look at $w = N^j v$. It lies in K_{m+1} , which by hypothesis is K_m , and that is exactly the statement $N^{m+j} v = 0$. Then (c) and (d) are bookkeeping: while the inclusions are strict the dimensions go up by at least 1, so strictness cannot survive past step n .

Proof.

- (a) If $N^k v = 0$, then $N^{k+1} v = N(N^k v) = 0$, so $K_k \subseteq K_{k+1}$. Each $K_k = \ker N^k$ is the kernel of a polynomial in N , hence N -invariant by Theorem 6.7.4 (b).
- (b) Suppose $K_m = K_{m+1}$. We prove $K_{m+j} = K_m$ by induction on $j \geq 0$. For $j = 0$ there is nothing to prove. Assume $K_{m+j} = K_m$, and let $v \in K_{m+j+1}$. Then

$$N^{m+1}(N^j v) = N^{m+j+1} v = 0,$$

so $N^j v \in K_{m+1} = K_m$, which says $N^{m+j} v = N^m(N^j v) = 0$. Hence $v \in K_{m+j} = K_m$. Together with $K_m \subseteq K_{m+j+1}$ from (a), this gives $K_{m+j+1} = K_m$.

- (c) First, some m with $0 \leq m \leq n$ satisfies $K_m = K_{m+1}$. Otherwise $K_m \subsetneq K_{m+1}$ for every $m = 0, 1, \dots, n$, so $\dim K_{m+1} \geq \dim K_m + 1$ for those m , and hence $\dim K_{n+1} \geq n + 1$, contradicting $\dim K_{n+1} \leq \dim V = n$ (Theorem 1.6.7). By the well-ordering principle (Theorem 0.2.11) there is a least such index; call it s , and $s \leq n$. For $k < s$ minimality gives $K_k \neq K_{k+1}$, so the inclusion $K_k \subseteq K_{k+1}$ of (a) is strict. For $k \geq s$, (b) gives $K_k = K_s$.
- (d) Since $s \leq n$, (c) gives $K_n = K_s = K_{n+j}$ for every $j \geq 0$.

■

Quick check. Can $\ker N \subsetneq \ker N^2 = \ker N^3$ happen? Can $\ker N = \ker N^2 \subsetneq \ker N^3$ happen?

Solution. The first can: for $\mathbf{N} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on F^2 we have $\ker \mathbf{N} = \text{span}(\mathbf{e}_1)$ and $\ker \mathbf{N}^2 = F^2 = \ker \mathbf{N}^3$, since $\mathbf{N}^2 = \mathbf{O}$. Here $s = 2$.

The second cannot, by Theorem 10.1.1 (b) with $m = 1$: equality at one step forces equality at every later step. This is why the theorem says the strict inclusions come **first** and the equalities **after**; the chain never resumes growing.

The index s will come up often enough to name. We call the least $s \geq 0$ with $\ker N^s = \ker N^{s+1}$ the **stabilization index** of N . Part (d) is the useful form: the chain has certainly stopped by step $n = \dim V$, whatever N is. Applying this to $N = T - \lambda \text{id}_V$ turns the infinite union in Definition 10.1.2 into a single kernel.

Corollary 10.1.2. The Generalized Eigenspace Is One Kernel

Let V be a finite-dimensional vector space over F with $\dim V = n$, let $T \in \mathcal{L}(V)$ and $\lambda \in F$, and put $N = T - \lambda \text{id}_V$ with stabilization index s . Then:

- (a) $G_\lambda(T) = \ker N^s = \ker N^n$;
- (b) $G_\lambda(T)$ is T -invariant, and $E_\lambda(T) \subseteq G_\lambda(T)$;
- (c) $G_\lambda(T) \neq \{0\}$ if and only if λ is an eigenvalue of T ;
- (d) if $v \in G_\lambda(T)$ is nonzero and $k \geq 1$ is **least** with $N^k v = 0$, then $N^{k-1} v$ is an eigenvector of T for λ .

Proof.

- (a) Every $\ker N^k$ with $k \geq 1$ is contained in $\ker N^{\max(k,n)} = \ker N^n$ by Theorem 10.1.1 (a) and (d), so the union $G_\lambda(T)$ equals $\ker N^n$; and $\ker N^n = \ker N^s$ by (c) of the same theorem, since $s \leq n$.
- (b) By (a), $G_\lambda(T) = \ker N^n$ is the kernel of a polynomial in T , so T maps it into itself (Theorem 6.7.4 (b)). And $E_\lambda(T) = \ker N \subseteq \ker N^n$.
- (c) ($\Leftarrow$) If λ is an eigenvalue, then $E_\lambda(T) \neq \{0\}$ (Theorem 9.2.1), and $E_\lambda(T) \subseteq G_\lambda(T)$ by (b). ($\Rightarrow$) If $v \in G_\lambda(T)$ is nonzero, take $k \geq 1$ least with $N^k v = 0$, which exists by the definition of $G_\lambda(T)$ and Theorem 0.2.11. Then $w = N^{k-1}v \neq 0$ by minimality of k (for $k = 1$ this reads $w = v \neq 0$), while $Nw = N^k v = 0$. So $w \in \ker N$ is nonzero and λ is an eigenvalue.
- (d) This is the computation just made: $w = N^{k-1}v$ is nonzero and lies in $\ker N = E_\lambda(T)$.

■

Part (d) is a move worth remembering: **to get a genuine eigenvector out of a generalized one, apply $T - \lambda \text{id}_V$ as many times as possible without reaching 0.** It is the reason the enlargement never produces eigenvalues out of thin air, and it will start every Jordan chain in Section 2.

10.1.4 The generalized eigenspace decomposition

We can now state the payoff. Chapter 9's primary decomposition (Theorem 9.8.1) already splits V into invariant pieces, one for each irreducible factor of m_T . When p_T splits, all those factors are linear, the pieces are kernels of powers of $T - \lambda_i \text{id}_V$ — and those are exactly the generalized eigenspaces. The extra information here is that the piece attached to λ_i has dimension $a_T(\lambda_i)$, so the sizes are read off the characteristic polynomial.

When the characteristic polynomial splits, the generalized eigenspaces fill the space, and the one for λ has dimension the algebraic multiplicity of λ .

Theorem 10.1.3. Generalized Eigenspace Decomposition

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, let $T \in \mathcal{L}(V)$, and suppose p_T **splits** over F , with distinct eigenvalues $\lambda_1, \dots, \lambda_k$ and

$$p_T = (x - \lambda_1)^{a_1} \cdots (x - \lambda_k)^{a_k}, \quad m_T = (x - \lambda_1)^{s_1} \cdots (x - \lambda_k)^{s_k},$$

where $a_i = a_T(\lambda_i)$ and $1 \leq s_i \leq a_i$. Write $G_i = G_{\lambda_i}(T)$ and $N_i = (T - \lambda_i \text{id}_V)|_{G_i}$. Then:

- (a) $G_i = \ker(T - \lambda_i \text{id}_V)^{s_i}$, and

$$V = G_1 \oplus G_2 \oplus \cdots \oplus G_k;$$

- (b) each G_i is T -invariant and nonzero, with $E_{\lambda_i}(T) \subseteq G_i$, and $E_{\lambda_i}(T) = G_i$ if and only if $s_i = 1$;
- (c) $\dim G_i = a_i$ and $p_{T|_{G_i}} = (x - \lambda_i)^{a_i}$;
- (d) $N_i^{a_i} = 0$; that is, $T|_{G_i} = \lambda_i \text{id}_{G_i} + N_i$ with N_i nilpotent, of stabilization index s_i ;
- (e) if $\mathcal{B}$ is a basis of V obtained by writing bases $\mathcal{B}_1, \dots, \mathcal{B}_k$ of $G_1, \dots, G_k$ one after another,

then

$$[T]_{\mathcal{B}}^{\mathcal{B}} = [T|_{G_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} \oplus \cdots \oplus [T|_{G_k}]_{\mathcal{B}_k}^{\mathcal{B}_k}$$

is block diagonal, with the i -th block of size a_i .

Idea. Almost everything is already proved; the one new step is (a), the identification $G_i = \ker(T - \lambda_i \text{id}_V)^{s_i}$. One inclusion is free, since $\ker(T - \lambda_i)^{s_i}$ sits inside the union defining G_i . For the other, count. Chapter 9 says the primary piece $\ker(T - \lambda_i)^{s_i}$ has dimension a_i , so it is enough to show $\dim G_i \leq a_i$. Now G_i is invariant, and the only eigenvalue of T on G_i is λ_i : an eigenvector in G_i for μ is killed by $(\mu - \lambda_i)^n$, which forces $\mu = \lambda_i$. So the characteristic polynomial of the restriction is a power of $x - \lambda_i$ dividing p_T , and its degree $\dim G_i$ is at most a_i .

Proof. Since p_T splits and $m_T \mid p_T$, Corollary 9.6.2 (b) gives the displayed factorization of m_T with $1 \leq s_i \leq a_i$. Put $V_i = \ker(T - \lambda_i \text{id}_V)^{s_i}$. The polynomials $x - \lambda_i$ are distinct monic irreducibles, so Theorem 9.8.1 applies to $m_T = \prod_i (x - \lambda_i)^{s_i}$ and gives

$$V = V_1 \oplus \cdots \oplus V_k,$$

with each V_i invariant and nonzero. By Corollary 9.8.4, $\dim V_i = a_i$, $p_{T|_{V_i}} = (x - \lambda_i)^{a_i}$, and $E_{\lambda_i}(T) \subseteq V_i$ with equality if and only if $s_i = 1$.

Claim: $G_i = V_i$. Fix i and write $G = G_i$, $\lambda = \lambda_i$, $a = a_i$, $N = T - \lambda \text{id}_V$.

($\supseteq$) $V_i = \ker N^{s_i} \subseteq \bigcup_{j \geq 1} \ker N^j = G$.

($\subseteq$) It suffices to prove $\dim G \leq a$: then $V_i \subseteq G$ and $\dim V_i = a \geq \dim G$ force $\dim V_i = \dim G$, and Theorem 1.6.8, applied to the subspace V_i of G , gives $V_i = G$.

First, *the only eigenvalue of $T|_G$ is λ* . Indeed, G is T -invariant (Corollary 10.1.2 (b)), so the restriction makes sense (Definition 9.1.2). Let $\mu \in F$ be an eigenvalue of $T|_G$, with eigenvector $v \in G$, $v \neq 0$. From $Tv = \mu v$ we get $Nv = (\mu - \lambda)v$, hence $N^n v = (\mu - \lambda)^n v$ by induction. But $G = \ker N^n$ by Corollary 10.1.2 (a), so $(\mu - \lambda)^n v = 0$ with $v \neq 0$, which forces the scalar $(\mu - \lambda)^n$ to be 0 (Theorem 1.1.9), and so $\mu = \lambda$, since a field has no zero divisors (Theorem 0.7.1).

Now compare characteristic polynomials. $G \neq \{0\}$, since λ is an eigenvalue of T (Corollary 10.1.2 (c)). If $G = V$, then every root of p_T is an eigenvalue of $T = T|_G$ (Theorem 9.2.1), so λ is the only root; as p_T splits and is monic of degree n (Theorem 7.10.2), $p_T = (x - \lambda)^n$ and $a = n = \dim G$. If $G \neq V$, then Theorem 9.1.3 (b) applies to the invariant subspace G and gives

$$p_T = p_{T|_G} p_{\bar{T}},$$

so $p_{T|_G}$ is a monic divisor of $p_T = \prod_j (x - \lambda_j)^{a_j}$. By Lemma 9.5.2, $p_{T|_G} = \prod_j (x - \lambda_j)^{c_j}$ with $0 \leq c_j \leq a_j$. Its roots are the eigenvalues of $T|_G$ (Theorem 9.2.1, applied to $T|_G$), which we just showed to be λ alone, so $c_j = 0$ for $j \neq i$. Hence $p_{T|_G} = (x - \lambda)^{c_i}$, and comparing degrees (Theorem 7.10.2) gives $\dim G = c_i \leq a$. This proves the claim, and with it (a), (b), and (c), since $G_i = V_i$.

(d) By Theorem 9.8.1 (c) and the claim, $m_{T|_{G_i}} = (x - \lambda_i)^{s_i}$. For an integer $j \geq 0$, the operator N_i^j is $(x - \lambda_i)^j$ evaluated at $T|_{G_i}$, so $N_i^j = 0$ if and only if $(x - \lambda_i)^{s_i}$ divides $(x - \lambda_i)^j$ (Theorem 9.5.1), that is, if and only if $j \geq s_i$. Since $s_i \leq a_i$, this gives

$N_i^{a_i} = 0$, and it identifies s_i as the least exponent killing N_i . Hence $\ker N_i^{s_i} = G_i$ while $\ker N_i^{s_i-1} \subsetneq G_i$, so the stabilization index of N_i is s_i .

- (e) Each G_i is invariant and nonzero, and $V = \bigoplus_i G_i$, so Theorem 9.1.5 gives the block diagonal shape, the i -th block being $[T|_{G_i}]_{\mathcal{B}_i}^{s_i}$, of size $\dim G_i = a_i$.

■

Read (d) again, since it is the plan for the rest of the chapter. **On its generalized eigenspace, T is a scalar plus a nilpotent operator.** The scalar part is as simple as an operator can be. So the only question left is what a nilpotent operator looks like in a well-chosen basis, and that is the subject of the next section. Once it is answered, (e) assembles the answers for the different eigenvalues into one matrix for T .

Two consistency checks. Summing dimensions in (a) and using (c) gives $\sum_i a_i = n$, which is the degree count for a split p_T . And if T is diagonalizable, then every $s_i = 1$ (Theorem 9.8.2), so (b) says $G_i = E_{\lambda_i}(T)$ and (a) becomes the eigenspace decomposition $V = \bigoplus_i E_{\lambda_i}(T)$ of Theorem 9.4.1 (c). Nothing is lost, and nothing new is gained, in the diagonalizable case.

Example. Generalized Eigenspaces of a 4×4 Matrix

Let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & -1 & -2 \\ -3 & 3 & -1 & 1 \\ 3 & -1 & 3 & -1 \\ -3 & 1 & -1 & 3 \end{pmatrix} \in M_4(\mathbb{Q}), \quad p_{\mathbf{A}} = (x-2)^3(x-5).$$

Find $E_2(\mathbf{A})$, $G_2(\mathbf{A})$, $G_5(\mathbf{A})$, the stabilization index of $\mathbf{A} - 2\mathbf{I}$, and the minimal polynomial. Verify Theorem 10.1.3 on this matrix.

Solution. The eigenspace. Write $\mathbf{N} = \mathbf{A} - 2\mathbf{I}$, so

$$\mathbf{N} = \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & 1 & -1 & 1 \\ 3 & -1 & 1 & -1 \\ -3 & 1 & -1 & 1 \end{pmatrix}.$$

Rows 2 and 4 are equal and row 3 is their negative, so the row space is spanned by rows 1 and 2, which are not proportional; hence $\text{rank } \mathbf{N} = 2$ and $\text{nullity } \mathbf{N} = 2$ (Theorem 3.5.7). The kernel is cut out by

$$x_2 - x_3 - 2x_4 = 0, \quad -3x_1 + x_2 - x_3 + x_4 = 0.$$

Subtracting the first equation from the second gives $-3x_1 + 3x_4 = 0$, so $x_1 = x_4$, and then $x_2 = x_3 + 2x_4$. With $x_3 = s$ and $x_4 = t$ free,

$$E_2(\mathbf{A}) = \{(t, s + 2t, s, t)\} = \text{span}((0, 1, 1, 0), (1, 2, 0, 1)), \quad g_{\mathbf{A}}(2) = 2.$$

Since $a_{\mathbf{A}}(2) = 3 > 2$, the matrix is not diagonalizable (Theorem 9.4.1 (e)).

The generalized eigenspace. Squaring,

$$\mathbf{N}^2 = 9 \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \end{pmatrix},$$

which has rank 1, so $\ker \mathbf{N}^2$ has dimension 3 and is the hyperplane

$$\ker \mathbf{N}^2 = \{ \mathbf{x} : x_1 = x_4 \} = \text{span}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_4).$$

The chain of kernels has dimensions 0, 2, 3 so far. Since $\ker \mathbf{N}^2 \subseteq G_2(\mathbf{A})$ and $\dim G_2(\mathbf{A}) = a_{\mathbf{A}}(2) = 3 = \dim \ker \mathbf{N}^2$ by part (c) of the theorem, we get $G_2(\mathbf{A}) = \ker \mathbf{N}^2$, and the stabilization index is $s = 2$. One can see the stabilization by hand as well: multiplying out gives $\mathbf{N}^3 = 3\mathbf{N}^2$, so $\ker \mathbf{N}^3 = \ker \mathbf{N}^2$ directly.

The second generalized eigenspace. Here $a_{\mathbf{A}}(5) = 1$, so $\dim G_5(\mathbf{A}) = 1$ and $G_5(\mathbf{A}) = E_5(\mathbf{A})$ by part (b) — its s must be 1. Solving $(\mathbf{A} - 5\mathbf{I})\mathbf{x} = \mathbf{0}$ gives $G_5(\mathbf{A}) = E_5(\mathbf{A}) = \text{span}((0, 1, -1, 1))$. As a check, $\mathbf{A}(0, 1, -1, 1) = (1+1-2, 3+1+1, -1-3-1, 1+1+3) = (0, 5, -5, 5)$.

The decomposition. $\mathbb{Q}^4 = G_2(\mathbf{A}) \oplus G_5(\mathbf{A})$, of dimensions $3 + 1 = 4$, and the sum is direct because $(0, 1, -1, 1) \notin G_2(\mathbf{A})$: its first and last coordinates are $0 \neq 1$.

The minimal polynomial. The stabilization indices are $s_1 = 2$ at $\lambda = 2$ and $s_2 = 1$ at $\lambda = 5$, and part (d) of the theorem identifies these indices as the exponents in $m_{\mathbf{A}}$, so $m_{\mathbf{A}} = (x - 2)^2(x - 5)$. Consistently, $(\mathbf{A} - 2\mathbf{I})(\mathbf{A} - 5\mathbf{I}) \neq \mathbf{0}$: the two factors are coprime, so the kernel of the product is $\ker(\mathbf{A} - 2\mathbf{I}) \oplus \ker(\mathbf{A} - 5\mathbf{I})$ (Theorem 6.7.5 (a)), of dimension $2 + 1 = 3 < 4$.

A basis adapted to the splitting. Take $\mathcal{B} = (\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_4, (0, 1, -1, 1))$, a basis of G_2 followed by one of G_5 . By part (e), $[T_{\mathbf{A}}]_{\mathcal{B}}^{\mathcal{B}}$ is block diagonal with a 3×3 block and the 1×1 block (5).

Warning. The dimension of G_{λ} is the algebraic multiplicity, not the number of steps the chain takes. In the example, $\dim G_2 = 3$ while the chain $\{\mathbf{0}\} \subsetneq \ker \mathbf{N} \subsetneq \ker \mathbf{N}^2 = G_2$ stabilizes after $s = 2$ steps. The two numbers measure different things: $a(\lambda) = \dim G_{\lambda}$ is the total size of the piece, and s is the exponent of $x - \lambda$ in the minimal polynomial. All that holds in general is $1 \leq s \leq a(\lambda)$ and $g(\lambda) \leq a(\lambda)$ (Theorem 9.3.1).

Finally, a word on what we may not yet assume. The decomposition needs p_T to split, which is automatic over $\mathbb{C}$ (Corollary 6.5.5) but fails, for instance, for a rotation of $\mathbb{R}^2$, whose generalized eigenspaces over $\mathbb{R}$ are all $\{\mathbf{0}\}$ and certainly do not fill $\mathbb{R}^2$. Operators over an arbitrary field need the coarser primary decomposition of Chapter 9, whose pieces we return to in Section 6.

10.1.5 Exercises

10.1.5.1 A. Check your understanding

Exercise. A1

- (a) Define a generalized eigenvector and the generalized eigenspace $G_\lambda(T)$.
- (b) State what [Theorem 10.1.1](#) says about the chain $\ker N \subseteq \ker N^2 \subseteq \dots$, including the bound on where it stabilizes.
- (c) True or false: every generalized eigenvector of T for λ satisfies $Tv = \lambda v$. Justify your answer.
- (d) True or false: if λ is not an eigenvalue of T , then $G_\lambda(T) = \{0\}$. Justify your answer.
- (e) Let $\dim V = 6$ and let $p_T = (x-1)^4(x-3)^2$ split over F . What are $\dim G_1(T)$ and $\dim G_3(T)$?
- (f) In the situation of (e), what can you say about $\dim E_1(T)$?

Solution.

- (a) A generalized eigenvector of T for $\lambda \in F$ is a **nonzero** $v \in V$ with $(T - \lambda \operatorname{id}_V)^k v = 0$ for **some** $k \geq 1$ ([Definition 10.1.1](#)); $G_\lambda(T)$ is the set of all of them together with 0, that is, $\bigcup_{k \geq 1} \ker(T - \lambda \operatorname{id}_V)^k$ ([Definition 10.1.2](#)).
- (b) The kernels increase; once two consecutive ones are equal, all later ones equal them; the strict inclusions come first, and the chain has stabilized by step $\dim V$, so $G_\lambda(T) = \ker(T - \lambda \operatorname{id}_V)^{\dim V}$.
- (c) False. For $\mathbf{J} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, the vector $\mathbf{e}_2$ lies in $G_2(\mathbf{J})$ but $\mathbf{J}\mathbf{e}_2 = \mathbf{e}_1 + 2\mathbf{e}_2 \neq 2\mathbf{e}_2$. Only the vectors of $E_\lambda(T) \subseteq G_\lambda(T)$ satisfy the eigenvector equation.
- (d) True, by [Corollary 10.1.2](#) (c).
- (e) $\dim G_1(T) = a_T(1) = 4$ and $\dim G_3(T) = a_T(3) = 2$ ([Theorem 10.1.3](#) (c)).
- (f) Only $1 \leq \dim E_1(T) \leq 4$, since $E_1(T) \subseteq G_1(T)$ and $1 \leq g_T(1) \leq a_T(1)$ ([Theorem 9.3.1](#)). All four values occur, as the Jordan forms of Section 3 will show.

10.1.5.2 B. Practice**Exercise. B1: Computing generalized eigenspaces**

For each operator, find all generalized eigenspaces, and compare each with the corresponding eigenspace.

- (a) $\mathbf{A}_1 = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \in M_2(\mathbb{Q})$.
- (b) $\mathbf{A}_2 = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \in M_3(\mathbb{R})$, which has $m_{\mathbf{A}_2} = (x-1)^2(x-2)$ ([Exercise \(B2: Decomposing \$\mathbb{R}^3\$ \)](#)).
- (c) $D \in \mathcal{L}(\mathbb{R}[x]_{\leq 3})$, differentiation.
- (d) The rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, first over $\mathbb{R}$ and then over $\mathbb{C}$.

Solution.

- (a) $p_{\mathbf{A}_1} = (x - 3)^2$, so 3 is the only eigenvalue, and $G_\lambda = \{\mathbf{0}\}$ for $\lambda \neq 3$ (Corollary 10.1.2 (c)). Since $(\mathbf{A}_1 - 3\mathbf{I})^2 = \mathbf{0}$, $G_3(\mathbf{A}_1) = \mathbb{Q}^2$, while $E_3(\mathbf{A}_1) = \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{span}(\mathbf{e}_1)$ is a line.
- (b) From $m_{\mathbf{A}_2} = (x - 1)^2(x - 2)$ and Theorem 10.1.3 (a), $G_1 = \ker(\mathbf{A}_2 - \mathbf{I})^2$ and $G_2 = \ker(\mathbf{A}_2 - 2\mathbf{I})$. Chapter 9 computed these: $G_1 = \text{span}((-1, 1, 0), (0, 0, 1))$, the plane $x + y = 0$, and $G_2 = \text{span}((0, -1, 1))$. The eigenspace $E_1(\mathbf{A}_2) = \text{span}((-1, 1, 0))$ is a line strictly inside G_1 , while $E_2(\mathbf{A}_2) = G_2$ because $s_2 = 1$. All other G_λ are $\{\mathbf{0}\}$.
- (c) $D^4 = 0$ (Example (Differentiation Is Nilpotent)), so $G_0(D) = \mathbb{R}[x]_{\leq 3}$, of dimension 4, while $E_0(D) = \ker D = \mathbb{R}$ is the line of constants. Since $p_D = x^4$ has 0 as its only root, every other generalized eigenspace is $\{\mathbf{0}\}$.
- (d) Over $\mathbb{R}$, $p_{\mathbf{R}} = x^2 + 1$ has no root, so $\mathbf{R}$ has no eigenvalue and $G_\lambda(\mathbf{R}) = \{\mathbf{0}\}$ for every $\lambda \in \mathbb{R}$. The generalized eigenspaces do **not** fill $\mathbb{R}^2$; the hypothesis that $p_{\mathbf{R}}$ splits fails. Over $\mathbb{C}$, $p_{\mathbf{R}} = (x - i)(x + i)$ splits with $a(i) = a(-i) = 1$, so $G_i(\mathbf{R}) = E_i(\mathbf{R}) = \text{span}((1, -i))$ and $G_{-i}(\mathbf{R}) = E_{-i}(\mathbf{R}) = \text{span}((1, i))$, and $\mathbb{C}^2 = G_i \oplus G_{-i}$. (Check: $\mathbf{R}(1, -i) = (i, 1) = i(1, -i)$.)

Exercise. B2: The image chain stabilizes too

Let V be finite-dimensional with $\dim V = n$, and let $N \in \mathcal{L}(V)$.

- (a) Prove that $\text{im } N \supseteq \text{im } N^2 \supseteq \text{im } N^3 \supseteq \dots$, and that each $\text{im } N^k$ is N -invariant.
- (b) Prove that $\text{im } N^m = \text{im } N^{m+1}$ if and only if $\ker N^m = \ker N^{m+1}$.
- (c) Deduce that $\text{im } N^n = \text{im } N^{n+j}$ for every $j \geq 0$, and that the image chain stabilizes at exactly the same index s as the kernel chain.

Solution.

- (a) $\text{im } N^{k+1} = \{N^{k+1}v : v \in V\} = \{N^k(Nv) : v \in V\} \subseteq \text{im } N^k$. Each $\text{im } N^k$ is the image of a polynomial in N , hence N -invariant (Theorem 6.7.4 (b)).
- (b) By Theorem 2.3.1, $\dim \text{im } N^k + \dim \ker N^k = n$ for every k . The inclusions $\text{im } N^{m+1} \subseteq \text{im } N^m$ and $\ker N^m \subseteq \ker N^{m+1}$ are equalities exactly when the dimensions agree (Theorem 1.6.8), and $\dim \text{im } N^m = \dim \text{im } N^{m+1}$ if and only if $\dim \ker N^m = \dim \ker N^{m+1}$ by the displayed identity.
- (c) By Theorem 10.1.1 (d), $\ker N^n = \ker N^{n+j}$; iterating (b) at each step from n to $n + j$ gives $\text{im } N^n = \text{im } N^{n+j}$. And by (b), the least index at which two consecutive images agree is the least index at which two consecutive kernels agree, namely s .

Exercise. B3: Reading the dimensions off a triangular matrix

Let

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in M_4(\mathbb{Q}).$$

- (a) Write down $p_{\mathbf{A}}$ and the algebraic multiplicities.
- (b) Compute $E_1(\mathbf{A})$, $G_1(\mathbf{A})$, $E_2(\mathbf{A})$, and $G_2(\mathbf{A})$, and check $\dim G_{\lambda} = a_{\mathbf{A}}(\lambda)$.
- (c) Verify that $\mathbb{Q}^4 = G_1(\mathbf{A}) \oplus G_2(\mathbf{A})$, and find $m_{\mathbf{A}}$.

Solution.

- (a) $\mathbf{A}$ is upper triangular, so $p_{\mathbf{A}} = (x-1)^2(x-2)^2$ (Theorem 7.4.6), and $a_{\mathbf{A}}(1) = a_{\mathbf{A}}(2) = 2$.

- (b) $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ has rank 3 (rows 1, 2, 4 are independent and row 3 is the sum of rows 2 and 4), so $E_1(\mathbf{A}) = \text{span}(\mathbf{e}_1)$. Squaring,

$$(\mathbf{A} - \mathbf{I})^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

of rank 2, with kernel $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$; so $G_1(\mathbf{A}) = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, of dimension $2 = a_{\mathbf{A}}(1)$. (Membership can be seen directly: $(\mathbf{A} - \mathbf{I})\mathbf{e}_2 = \mathbf{e}_1$ and $(\mathbf{A} - \mathbf{I})\mathbf{e}_1 = \mathbf{0}$.)

Similarly $\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ has rank 3, so $E_2(\mathbf{A})$ is a line; solving

gives $E_2(\mathbf{A}) = \text{span}((1, 1, 1, 0))$. And $(\mathbf{A} - 2\mathbf{I})^2 = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ has rank 2;

its kernel is cut out by $x_1 - 2x_2 + x_3 = 0$ and $x_2 - x_3 + x_4 = 0$, so $G_2(\mathbf{A}) = \text{span}((1, 1, 1, 0), (-2, -1, 0, 1))$, of dimension $2 = a_{\mathbf{A}}(2)$. (Check: $(\mathbf{A} - 2\mathbf{I})(-2, -1, 0, 1) = (1, 1, 1, 0)$, the eigenvector.)

- (c) The four vectors $\mathbf{e}_1, \mathbf{e}_2, (1, 1, 1, 0), (-2, -1, 0, 1)$ form a basis of $\mathbb{Q}^4$: the matrix with these columns is upper triangular with diagonal 1, 1, 1, 1, hence invertible. So the sum is direct and equals $\mathbb{Q}^4$, as Theorem 10.1.3 (a) predicts. Both stabilization indices are 2 (the eigenspaces are lines, the generalized eigenspaces planes), so $m_{\mathbf{A}} = (x-1)^2(x-2)^2 = p_{\mathbf{A}}$.

10.1.5.3 C. Going deeper

Exercise. C1: The Fitting decomposition

Let V be finite-dimensional with $\dim V = n$, and let $N \in \mathcal{L}(V)$ be **any** operator. Put

$U = \ker N^n$ and $W = \operatorname{im} N^n$.

- (a) Prove that $U \cap W = \{0\}$, and deduce $V = U \oplus W$.
- (b) Prove that U and W are N -invariant, that $N|_U$ is nilpotent, and that $N|_W$ is invertible.
- (c) Identify U as a generalized eigenspace of N .
- (d) Compute U and W for $\mathbf{N} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(\mathbb{Q})$.

Hint: for (a), if $v = N^n w$ and $N^n v = 0$, look at which kernel w lies in.

Solution.

- (a) Let $v \in U \cap W$, say $v = N^n w$ with $N^n v = 0$. Then $N^{2n} w = N^n v = 0$, so $w \in \ker N^{2n} = \ker N^n$ by Theorem 10.1.1 (d), and therefore $v = N^n w = 0$. By Theorem 2.3.1, $\dim U + \dim W = n$, so $\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = n$ (Theorem 1.7.3), whence $U + W = V$ (Theorem 1.6.8) and the sum is direct (Theorem 1.7.4).
- (b) Both are invariant by Theorem 6.7.4 (b). On U we have $(N|_U)^n = 0$ by definition of U , so $N|_U$ is nilpotent. On W , the map $N|_W: W \rightarrow W$ is surjective: $N(W) = N(\operatorname{im} N^n) = \operatorname{im} N^{n+1} = \operatorname{im} N^n = W$ by Exercise (B2: The image chain stabilizes too).
- (c). A surjective operator on a finite-dimensional space is invertible (Corollary 2.3.2 (e)).
- (c) $U = \ker N^n = G_0(N)$, by Corollary 10.1.2 (a) with $\lambda = 0$. So the Fitting decomposition splits V into the generalized eigenspace for 0 and a part on which N is invertible.
- (d) Here $n = 2$ and $\mathbf{N}^2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \mathbf{N}$, so $U = \ker \mathbf{N} = \operatorname{span}(\mathbf{e}_1)$ and $W = \operatorname{im} \mathbf{N} = \operatorname{span}((1, 1))$. Indeed $\mathbb{Q}^2 = \operatorname{span}(\mathbf{e}_1) \oplus \operatorname{span}((1, 1))$, $\mathbf{N}$ is zero on U , and $\mathbf{N}(1, 1) = (1, 1)$, so $\mathbf{N}|_W$ is the identity.

Exercise. C2: Commuting operators preserve generalized eigenspaces

Let V be finite-dimensional, and let $S, T \in \mathcal{L}(V)$ with $ST = TS$.

- (a) Prove that $S(G_\lambda(T)) \subseteq G_\lambda(T)$ for every $\lambda \in F$.
- (b) Deduce that if p_T splits, then every generalized eigenspace of T is invariant under every polynomial in T and under S , and that V has a basis in which both S and T have block diagonal matrices with blocks of sizes $a_T(\lambda_1), \dots, a_T(\lambda_k)$.
- (c) Show by example that part (a) can be vacuous: find commuting S, T on $\mathbb{Q}^2$ and a λ with $E_\lambda(T)$ a line, $G_\lambda(T) = \mathbb{Q}^2$, and S not a scalar multiple of the identity.

Solution.

- (a) Since $ST = TS$, S commutes with every polynomial in T (Theorem 6.7.1 (d)), in particular with N^k , where $N = T - \lambda \operatorname{id}_V$. Let $v \in G_\lambda(T)$, say $N^k v = 0$. Then

$$N^k(Sv) = S(N^k v) = S0 = 0,$$

so $Sv \in \ker N^k \subseteq G_\lambda(T)$. (This is Theorem 6.7.4 (a) with $R = N^k$.)

- (b) Invariance under polynomials in T is Theorem 6.7.4 (b), and invariance under S is (a). If p_T splits, then $V = \bigoplus_i G_{\lambda_i}(T)$ with $\dim G_{\lambda_i}(T) = a_T(\lambda_i)$ (Theorem 10.1.3), and a basis assembled from bases of the summands makes both matrices block diagonal, by Theorem 9.1.5 applied to T and to S .
- (c) Take $T = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ and $S = T$. They commute, $E_2(T) = \text{span}(\mathbf{e}_1)$ is a line, $G_2(T) = \mathbb{Q}^2$, and S is not a scalar matrix. Consistently with Theorem 9.9.1, S maps the line $\text{span}(\mathbf{e}_1)$ into itself, and by (a) it maps $G_2(T) = \mathbb{Q}^2$ into itself, which says nothing. The lesson is that (a) is informative only when $G_\lambda(T) \neq V$.