VI.D. Generalized eigenspaces

Let $T\colon \mathbb{C}^n \to \mathbb{C}^n$ be a *fixed* linear transformation. For this section and the next, all vector spaces are assumed to be over $\mathbb{C}$; in particular, we will often write V for $\mathbb{C}^n$.[17] In what follows, I will write "S" for an "arbitrary" transformation, which could be T, or $\sigma\mathbb{I} - T$, or its restriction to a subspace, etc.

We are looking for forms $A = [T]_{\hat{e}}$ can be put into (via $P_{\mathcal{B}}^{-1} A P_{\mathcal{B}}$) even if it is not diagonalizable. The structure underlying the rational canonical form was a direct-sum decomposition of $V = \mathbb{C}^n$ into T-cyclic subspaces in 1-to-1 correspondence with the nontrivial invariant factors $\Delta_r(\lambda), \ldots, \Delta_n(\lambda)$ of A. In the present section we describe the structure beneath the Jordan canonical form — which, unlike the rational form, actually reduces to D when A is diagonalizable ($= P_{\mathcal{B}} D P_{\mathcal{B}}^{-1}$). We can forget about most of the $\mathsf{F}[\lambda]$ stuff here; the theory is fortunately easier than that in the last two sections.

Recall that if A is diagonalizable with eigenvalues $\{\sigma_1, \ldots, \sigma_s\}$,[18] then V is the sum of the corresponding eigenspaces and in fact the geometric multplicities add to n:

$$\sum_i \dim E_{\sigma_i}(A) = n.$$

In the language of direct sums,

$$V = E_{\sigma_1}(A) \oplus \cdots \oplus E_{\sigma_s}(A).$$

What we claim is that there are "generalized" eigenspaces $\widetilde{E}_{\sigma_i}$ such that

$$V = \widetilde{E}_{\sigma_1}(A) \oplus \cdots \oplus \widetilde{E}_{\sigma_s}(A)$$

even if A is *not* diagonalizable. They contain the E_{σ_i}, so if we write $d_i = \dim(E_{\sigma_i})$ and $\tilde{d}_i = \dim(\widetilde{E}_{\sigma_i})$, then $d_i \leq \tilde{d}_i$ and $\sum_i \tilde{d}_i = n$. Indeed, the $\tilde{d}_i$ will just turn out to be the algebraic multiplicities k_i.

[17]The reason to take $\mathsf{F} = \mathbb{C}$ is so that the algebraic multiplicities of the eigenvalues of $A \in M_n(\mathsf{F})$ always sum to n, i.e. $f_A(\lambda)$ breaks into linear factors over F. The results below hold more generally (e.g. with $\mathsf{F} = \mathbb{R}$) whenever this is the case.

[18]Here we mean the list of *distinct* eigenvalues, i.e. *not* repeated according to multiplicity.

The proof will require a few facts about stable image/kernel, and nilpotent transformations ($S\colon U \to U$ such that S^k is the zero transformation for some k). Throughout it is important to remember that if $W \subseteq V$ is closed under the action of T then the restriction of T to W makes sense as a linear transformation and is written $T|_W$ (and read "T on W").

Stable Image and Kernel. Given a transformation $S\colon V \to V$, the series of subspaces of V

$$\{0\} = \ker(\mathbb{I}) \subseteq \ker(S) \subseteq \ker(S^2) \subseteq \ldots$$

and

$$V = \mathrm{im}(\mathbb{I}) \supseteq \mathrm{im}(S) \supseteq \mathrm{im}(S^2) \supseteq \ldots$$

both level off at some point (since V is finite dimensional). Let K be sufficiently large that

$$\mathrm{im}(S^K) = \mathrm{im}(S^{K+1}) = \ldots$$
$$\ker(S^K) = \ker(S^{K+1}) = \ldots\,;$$

these are called the *stable image* and *stable kernel* of S. An equivalent definition of these objects (subspaces of V) is:

$$\text{(VI.D.1)}\qquad \begin{aligned} \widetilde{\ker}(S) &= \left\{\vec{w} \in V \,\middle|\, S^k\vec{w} = 0 \text{ for some } k\right\} \\ \widetilde{\mathrm{im}}(S) &= \left\{\vec{w} \in V \,\middle|\, \text{for every } k,\ \exists\ \vec{v} \in V \text{ s.t. } \vec{w} = S^k\vec{v}\right\}. \end{aligned}$$

VI.D.2. REMARK. The $\vec{v}$ such that $S^k\vec{v} = \vec{w}$ in the second definition are in general different for each k (even for $k \geq K$).

We claim that

$$\text{(VI.D.3)}\qquad \text{(i) } \widetilde{\mathrm{im}}(S) \cap \widetilde{\ker}(S) = \{0\}\,, \quad \text{(ii) } \widetilde{\mathrm{im}}(S) + \widetilde{\ker}(S) = V.$$

To see (i), let $\vec{w} \in \widetilde{\mathrm{im}}(S) \cap \widetilde{\ker}(S)$; that is, $\vec{w} = S^K\vec{v}$ and $S^K\vec{w} = 0$, so that $0 = S^K(S^K\vec{v}) = S^{2K}\vec{v}$. But then $\vec{v} \in \ker(S^{2K}) = \widetilde{\ker}(S) = \ker(S^K)$, so that $(\vec{w} =)\, S^k\vec{v} = 0$.

To see (ii), apply rank-nullity to S^K to get

(VI.D.4)
$$\dim V = \dim(\operatorname{im} S^K) + \dim(\ker S^K) = \dim(\widetilde{\operatorname{im}}(S)) + \dim(\widetilde{\ker}(S)),$$

and the "modular law" $\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2$ (cf. Exercise II.C.3) for subspaces $W_1, W_2 \subseteq V$ to get

$$\begin{aligned}\dim(\widetilde{\operatorname{im}}(S)) + \dim(\widetilde{\ker}(S)) \\ &= \dim(\widetilde{\operatorname{im}}(S) \cap \widetilde{\ker}(S)) + \dim(\widetilde{\operatorname{im}}(S) + \widetilde{\ker}(S)) \\ &\overset{(i)}{=} \dim(\widetilde{\operatorname{im}}(S) + \widetilde{\ker}(S)).\end{aligned}$$

Combining this with (VI.D.4), $\dim(\widetilde{\operatorname{im}}(S) + \widetilde{\ker}(S)) = \dim V$ and (ii) follows.

We rewrite (VI.D.3)(i-ii) as

(VI.D.5)
$$V = \widetilde{\operatorname{im}}(S) \oplus \widetilde{\ker}(S).$$

This is *always true*, for *any* $S\colon V \to V$. Moreover, since S *respects* this decomposition (as you can check), one may speak of the restrictions $S\big|_{\widetilde{\ker} S}$ and $S\big|_{\widetilde{\operatorname{im}} S}$. By definition some power k of S annihilates $\widetilde{\ker}\, S$, and so $S\big|_{\widetilde{\ker} S}$ is nilpotent. On the other hand,

$$\ker\left(S\big|_{\widetilde{\operatorname{im}} S}\right) = \ker S \cap \widetilde{\operatorname{im}}\, S \subseteq \widetilde{\ker}\, S \cap \widetilde{\operatorname{im}}\, S = \{0\}$$

by (VI.D.3)(i), and thus $S\big|_{\widetilde{\operatorname{im}} S}$ is invertible. We have proved

VI.D.6. PROPOSITION. *Given any $S : V \to V$, there is a direct-sum decomposition*

$$V = U_0 \oplus W_0$$

respected by S, such that $S\big|_{W_0}$ is nilpotent and $S\big|_{U_0}$ is invertible.

Now let's look more generally at the situation where S respects a (possibly different) direct sum decomposition $V = U \oplus W$. We claim that

(a) $\ker S = (U \cap \ker S) + (W \cap \ker S)$, and
(b) $(U \cap \ker S) \cap (W \cap \ker S) = \{0\}$.

Now (b) is immediate since $U \cap W = \{0\}$. To see (a): take any $\vec{v} \in \ker S$ and write $\vec{v} = \vec{u} + \vec{w}$ (possible because $V = U \oplus W$); clearly $0 = S\vec{v} = S\vec{u} + S\vec{w}$. Since S respects $U \oplus W$, $S\vec{u} \in U$ and $S\vec{w} \in W$, but then $S\vec{u} = -S\vec{w}$ is a "problem" since $U \cap W = \{0\}$. So we must have $S\vec{u} = S\vec{w} = 0$! That means $\vec{u} \in U \cap \ker S$, $\vec{w} \in W \cap \ker S$, and since $\vec{v}$ is their sum we have proved (a).

Of course (a) + (b) $\implies$ $\ker S = (U \cap \ker S) \oplus (W \cap \ker S)$, so applying this to S^K we get

VI.D.7. PROPOSITION. *Given $S : V \to V$ respecting some direct-sum decomposition*

$$V = U \oplus W,$$

one has

$$\widetilde{\ker}\, S = \left(U \cap \widetilde{\ker}\, S\right) \oplus \left(W \cap \widetilde{\ker}\, S\right).$$

Nilpotent Transformations. Every $S : V \to V$ has an eigenvalue (unless $V = \{0\}$), since the characteristic polynomial $f_S(\lambda)$ has a root in $\mathbb{C}$. (This is where we really need $V = \mathbb{C}^n$.) This eigenvalue has at least one nonzero eigenvector. What if *zero* is the only one?

VI.D.8. PROPOSITION. *S is nilpotent* $\iff$ 0 *is its only eigenvalue.*

PROOF. ($\Leftarrow$) Suppose $0 =$ only eigenvalue of $S =$ only root of $f_S(\lambda)$. That is, $f_S(\lambda) = \lambda^n$. By Cayley-Hamilton, S satisfies its own characteristic polynomial, so $S^n = 0$.

($\Rightarrow$) Suppose $S^k = 0$, and also suppose λ is an eigenvalue of S. There is a *nonzero* $\vec{v}$ such that $S\vec{v} = \lambda\vec{v}$, and thus

$$0 = S^k\vec{v} = \lambda^k\vec{v} \implies \lambda^k = 0 \implies \lambda = 0. \qquad \square$$

Stable Eigenspace. Given λ an eigenvalue of $S \colon V \to V$ ($\Leftrightarrow$ λ any root of f_S in $\mathbb{C}$), recall the definition

$$E_\lambda(S) := \ker(\lambda\mathbb{I} - S) = \{\vec{v} \in V \mid (\lambda\mathbb{I} - S)\vec{v} = 0\}$$

of the eigenspace of λ. Define the *generalized* or *stable eigenspace*

$$\widetilde{E}_\lambda(S) := \widetilde{\ker}(\lambda\mathbb{I} - S) = \left\{\vec{v} \in V \;\middle|\; (\lambda\mathbb{I} - S)^k\vec{v} = 0 \text{ for some } k\right\}.$$

Clearly $\widetilde{E}_\lambda(S) \supseteq E_\lambda(S)$.

Now we return to our original $T: V \to V$ with distinct eigenvalues $\{\sigma_1, \ldots, \sigma_s\}$, and set

$$W_j = \widetilde{E}_{\sigma_j}(T).$$

(These are not the W_j's of §VI.C!) Clearly some power of $(\sigma_j \mathbb{I} - T)$ annihilates W_j, so that $(\sigma_j \mathbb{I} - T)\Big|_{W_j}$ is nilpotent and has only eigenvalue 0. That is, if $\vec{v} \in W_j$ satisfies

$$(\sigma_j \mathbb{I} - T)\vec{v} = \lambda \vec{v},$$

then $\lambda = 0$. Therefore, if $\vec{v} \in W_j$ satisfies

$$T\vec{v} = \sigma \vec{v},$$

then

$$(\sigma_j \mathbb{I} - T)\vec{v} = (\sigma_j - \sigma)\vec{v}$$

and $\sigma_j - \sigma$ must be 0, i.e. $\sigma = \sigma_j$.

Conclusion: the only eigenvalue of $T\Big|_{W_j}$ is σ_j.

Now consider for $i \neq j$ the intersection of two stable eigenspaces

$$W_i \cap W_j.$$

The only eigenvalue of $T\big|_{W_i}$ is σ_i, while the only eigenvalue of $T\Big|_{W_j}$ is σ_j. Since $\sigma_i \neq \sigma_j$, $T\Big|_{W_i \cap W_j}$ can have no eigenvalue. This is absurd unless $W_i \cap W_j = \{0\}$, proving the

VI.D.9. PROPOSITION. *$\widetilde{E}_{\sigma_i}(T) \cap \widetilde{E}_{\sigma_j}(T) = \{0\}$ for all $i \neq j$.*

We make one further observation concerning stable eigenspaces: how to find bases for them. You know how to find bases for kernels. Working in the standard basis of $\mathbb{C}^n$ (in terms of which $[T]_{\hat{e}} = A$ by definition), find bases for

$$\ker(\sigma_i \mathbb{I} - A) \subseteq \ker\left\{(\sigma_i \mathbb{I} - A)^2\right\} \subseteq \ker\left\{(\sigma_i \mathbb{I} - A)^3\right\} \subseteq \ldots.$$

You stop when two successive bases have the same number of elements (once $\ker(S^k) = \ker(S^{k+1})$, all the remaining ones are the same as well: see Exercise (4)).

The Jordan Structure Theorem. Here is what holds even when T is not semisimple ($\Leftrightarrow A$ is not diagonalizable). We emphasize that the $\{W_j\}$ have nothing to do with those in the preceding section.

VI.D.10. THEOREM. *Let $T\colon V \to V$ ($V = \mathbb{C}^n$) be a linear transformation, with distinct eigenvalues $\{\sigma_1, \ldots, \sigma_s\}$ and corresponding stable eigenspaces $W_j = \widetilde{E}_{\sigma_j}(T) = \widetilde{\ker}(\sigma_j \mathbb{I} - T)$. Then*

$$V = W_1 \oplus \cdots \oplus W_s$$

and $\dim W_j$ *= algebraic multiplicity of σ_j. Furthermore, T respects this decomposition.*

PROOF. We first prove the decomposition, by induction on s. Set $\tilde{d}_j = \dim W_j$ and $A = [T]_{\hat{e}}$; and let k_j denote the algebraic multiplicity of σ_j (as a root of the characteristic polynomial f_A).

• <u>Case $s = 1$</u>: σ_1 = the only eigenvalue of T on $V \implies 0$ = only eigenvalue of $(\sigma_1 \mathbb{I} - T)$ on $V \implies (\sigma_1 \mathbb{I} - T)$ nilpotent $\implies (\sigma_1 \mathbb{I} - T)^k = 0 \implies V = \widetilde{\ker}(\sigma_1 \mathbb{I} - T) = W_1$.

• <u>Inductive step</u>: Assume the Theorem holds for transformations with $s-1$ distinct eigenvalues, and let T be as above. Apply (VI.D.5) (and Exercise (3)) to $S = \sigma_s \mathbb{I} - T$ to get

$$V = \widetilde{\ker}(\sigma_s \mathbb{I} - T) \oplus \widetilde{\mathrm{im}}(\sigma_s \mathbb{I} - T) =: W_s \oplus U_s,$$

where $\sigma_s \mathbb{I} - T$ respects the decomposition. Moreover, since $\mathbb{I}$ also respects the direct sum (or, for that matter, any direct sum!), so do T and $\sigma_j \mathbb{I} - T$, $j \neq s$. So we may speak of $T|_{U_s} : U_s \to U_s$. Since $(\sigma_s \mathbb{I} - T)$ is invertible on U_s, σ_s cannot be an eigenvalue of T there.[19]

Thus $T|_{U_s}$ has eigenvalues $\subseteq \{\sigma_1, \ldots, \sigma_{s-1}\}$, and by induction

$$U_s = {}'W_1 \oplus \cdots \oplus {}'W_{s-1},$$

[19] $(\sigma_s \mathbb{I} - T|_{U_s})$ invertible $\implies$ $\det(\sigma_s \mathbb{I} - T|_{U_s}) \neq 0$ $\implies$ σ_s not a root of $\det(\lambda \mathbb{I} - T|_{U_s})$.

where

$$'W_j = \widetilde{\ker}\left(\sigma_j\mathbb{I} - T\,|_{U_s}\right) = \widetilde{\ker}(\sigma_j\mathbb{I} - T) \cap U_s = W_j \cap U_s.$$

We must show that $'W_j = W_j$.

Since $(j \neq s)$ $\sigma_j\mathbb{I} - T$ also respects the decomposition $V = W_s \oplus U_s$, we have (Prop. VI.D.7)

$$W_j = \widetilde{\ker}(\sigma_j\mathbb{I} - T) = \{W_s \cap \widetilde{\ker}(\sigma_j\mathbb{I} - T)\} \oplus \{U_s \cap \widetilde{\ker}(\sigma_j\mathbb{I} - T)\}$$
$$= W_s \cap W_j \,\oplus\, U_s \cap W_j.$$

By Prop. VI.D.9, $W_s \cap W_j = \{0\}$ and so

$$W_j = U_s \cap W_j = {'W_j}\,,$$

as desired.

• <u>T respects the direct sum</u>: We need to show $T(W_j) \subseteq W_j$. Take $\vec{w} \in \widetilde{\ker}(\sigma_j\mathbb{I} - T)$, so that for κ sufficiently large $(\sigma_j\mathbb{I} - T)^\kappa \vec{w} = 0$. But then $(\sigma_j\mathbb{I} - T)^\kappa T\vec{w} = T(\sigma_j\mathbb{I} - T)^\kappa \vec{w} = 0 \quad \Longrightarrow \quad T\vec{w} \in \widetilde{\ker}(\sigma_j\mathbb{I} - T)$.

• <u>$\tilde{d}_j = k_j$</u>: Let $\mathcal{B}_1, \ldots, \mathcal{B}_s$ be bases for $W_1, \ldots, W_s$; the collection $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_s\}$ then yields a basis for V "subordinate to the direct sum". Since T respects the direct sum, its matrix with respect to $\mathcal{B}$ splits into blocks down the diagonal (of dimensions $\tilde{d}_1 \times \tilde{d}_1, \ldots, \tilde{d}_s \times \tilde{d}_s$):

$$[T]_{\mathcal{B}} =: B = P_{\mathcal{B}}^{-1} A P_{\mathcal{B}} = \text{diag}\left\{[T\,|_{W_1}]_{\mathcal{B}_1}\,,\ \ldots\ ,\ [T\,|_{W_s}]_{\mathcal{B}_s}\right\}$$
$$\text{diag}\left\{B_1, \ldots, B_s\right\}.$$

Moreover, since $A \sim B$, $\lambda\mathbb{I} - A \sim \lambda\mathbb{I} - B$ and $f_A(\lambda) = f_B(\lambda)$. From

$$\lambda\mathbb{I} - B = \text{diag}\left\{\lambda\mathbb{I}_{\tilde{d}_1} - B_1, \ldots, \lambda\mathbb{I}_{\tilde{d}_s} - B_s\right\}$$

we have

$$f_B(\lambda) = \det(\lambda\mathbb{I} - B) \;=\; \prod_j \det(\lambda\mathbb{I}_{\tilde{d}_j} - B_j) \;=\; f_{B_1}(\lambda) \cdot \cdots \cdot f_{B_s}(\lambda).$$

Since the only eigenvalue of $T\,|_{W_j}$ is σ_j (and $B_j = [T\,|_{W_j}]_{\mathcal{B}_j}$) the only root of $f_{B_j}(\lambda)$ is σ_j. Since B_j is $\tilde{d}_j \times \tilde{d}_j$, it follows that $\deg(f_{B_j}) = \tilde{d}_j$

and so $f_{B_j}(\lambda) = (\lambda - \sigma_j)^{\tilde{d}_j}$. But then $(f_A(\lambda) =)$

$$f_B(\lambda) = \prod (\lambda - \sigma_j)^{\tilde{d}_j}$$

and we are done. □

Exercises

(1) Find the stable eigenspaces of

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

(2) Suppose A is an 8×8 matrix with $m_A(\lambda) = \lambda(\lambda-1)^2(\lambda-2)^3$ and $f_A(\lambda) = \lambda^2(\lambda-1)^2(\lambda-2)^4$. What are the dimensions of the eigenspaces and stable eigenspaces of A?

(3) Check that S respects the decomposition (VI.D.5) into stable image and kernel.

(4) For any endomorphism $S \colon \mathbb{C}^n \to \mathbb{C}^n$, show $\ker(S^k) = \ker(S^{k+1})$ implies
(a) $\ker(S^k) = \ker(S^\ell)$ for all $\ell \geq k$, and
(b) $\mathrm{im}(S^k) = \mathrm{im}(S^\ell)$ for all $\ell \geq k$. [Hint for (b): use Rank + Nullity and (a).]

(5) Show that a matrix $A \in M_n(\mathbb{C})$ is nilpotent if and only if it is similar to an upper-triangular matrix with diagonal entries zero. [Hint: given a nilpotent matrix, what does its rational canonical form look like?]