Chapter 8

Operators on Complex Vector Spaces

8.A Generalized Eigenvectors and Nilpotent Operators

10/22:

- In this chapter, we will assume that V is a finite-dimensional *nonzero* vector space over $\mathbb{F}$ (just to avoid dealing with some trivialities).
- Null spaces and powers of an operator.

Theorem 8.1. *Suppose* $T \in \mathcal{L}(V)$. *Then*

$$\{0\} = \operatorname{null} T^0 \subset \operatorname{null} T^1 \subset \cdots$$

Proof. We induct on the exponent k of T. For the base case $k = 0$, suppose $v \in \operatorname{null} T^0$. Then $v \in \operatorname{null} I$ since $T^0 = I$ by definition. It follows that

$$0 = Iv = v$$

so $\{0\} = \operatorname{null} T^0$, as desired. Now suppose inductively that we have proven the claim for k; we now wish to show that $\operatorname{null} T^k \subset \operatorname{null} T^{k+1}$. Suppose $v \in \operatorname{null} T^k$. Then $T^k v = 0$. It follows that

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

so $v \in T^{k+1}$, as desired. ∎

Theorem 8.2. *Let* $T \in \mathcal{L}(V)$, *and suppose* m *is a nonnegative integer such that* $\operatorname{null} T^m = \operatorname{null} T^{m+1}$. *Then*

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} = \operatorname{null} T^{m+2} = \cdots$$

Proof. We induct on k, defined as follows. For the base case $k = 0$, we have that

$$\operatorname{null} T^{m+0} = \operatorname{null} T^m = \operatorname{null} T^{m+1} = \operatorname{null} T^{m+0+1}$$

by hypothesis, as desired. Now suppose inductively that we have proven that $\operatorname{null} T^{m+k-1} = \operatorname{null} T^{m+k}$; we now wish to show that $\operatorname{null} T^{m+k} = \operatorname{null} T^{m+k+1}$. By Theorem 8.1, we have that $\operatorname{null} T^{m+k} \subset \operatorname{null} T^{m+k+1}$. On the other hand, suppose that $v \in \operatorname{null} T^{m+k+1}$. Then

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

But this implies that $T^k v \in \operatorname{null} T^{m+1} = \operatorname{null} T^m$ by hypothesis. Therefore,

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

so $v \in \operatorname{null} T^{m+k}$, as desired. ∎

- Theorem 8.2 raises the question how to characterize/define/find nonnegative integers m such that the null space stops growing. We tackle begin to tackle this question with the following.

  Theorem 8.3. *Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then*

  $$\operatorname{null} T^n = \operatorname{null} T^{n+1} = \cdots$$

  Proof. To prove the claim, Theorem 8.2 tells us that we need only verify that $\operatorname{null} T^n = \operatorname{null} T^{n+1}$. Suppose for the sake of contradiction that $\operatorname{null} T^n \neq \operatorname{null} T^{n+1}$. Then by Theorem 8.2, we cannot have $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for any $0 \leq k \leq n$. However, by Theorem 8.1, we must still have that $\operatorname{null} T^k \subset \operatorname{null} T^{k+1}$ for each $k = 1, \dots, n$. Combining the last two results, we must have the following.

  $$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^n \subsetneq \operatorname{null} T^{n+1}$$

  At each of these strict inclusions, the dimension from the previous to the next null space must increase by at least one. Thus, $\dim \operatorname{null} T^{n+1} \geq n+1$. But since $\operatorname{null} T^{n+1} \subset V$, Theorem 2.11 asserts that $\dim \operatorname{null} T^{n+1} \leq n$, so we have that

  $$n + 1 \leq \dim \operatorname{null} T^{n+1} \leq n$$

  a contradiction. ■

- While it is not true that $V = \operatorname{null} T \oplus \operatorname{range} T$ for each $T \in \mathcal{L}(V)$, we can prove the following related theorem.

  Theorem 8.4. *Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then*

  $$V = \operatorname{null} T^n \oplus \operatorname{range} T^n$$

  Proof. To prove that $V = \operatorname{null} T^n \oplus \operatorname{range} T^n$, it will suffice to show that $(\operatorname{null} T^n) \cap (\operatorname{range} T^n) = \{0\}$ and that $\dim(\operatorname{null} T^n \oplus \operatorname{range} T^n) = \dim V$ (see Exercise 2.C.1). Let's begin.

  Suppose $v \in (\operatorname{null} T^n) \cap (\operatorname{range} T^n)$. Then $T^n v = 0$ and there exists $u \in V$ such that $v = T^n u$. Combining these results reveals that

  $$T^{2n} u = T^n v = 0$$

  so $u \in \operatorname{null} T^{2n} = \operatorname{null} T^n$ by Theorem 8.3. Therefore, $v = T^n u = 0$, as desired.

  As to the other equality, we have that

  $$\begin{aligned} \dim(\operatorname{null} T^n \oplus \operatorname{range} T^n) &= \dim \operatorname{null} T^n + \dim \operatorname{range} T^n && \text{Theorem 3.21} \\ &= \dim V && \text{Fundamental Theorem of Linear Maps} \end{aligned}$$

  as desired. ■

- Although many operators can be described by their eigenvectors, not all can. Thus, we introduce the following more general descriptor.

- **Generalized eigenvector** (of $T \in \mathcal{L}(V)$): A nonzero vector $v \in V$ such that

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

  for some positive integer j, where λ is an eigenvalue of T.

  - Although this definition lets j be arbitrary, we will soon prove that if $j = \dim V$, every generalized eigenvector satisfies the above equation.
  - Note that we do not define generalized eigenvalues because generalized eigenvectors still pertain to the original set of eigenvalues.

- Every eigenvector of T is a generalized eigenvector of T (take $j = 1$ in the definition).
- **Generalized eigenspace** (of $T \in \mathcal{L}(V)$ and λ): The set of all generalized eigenvectors of T corresponding to λ, and the 0 vector. *Denoted by* $\boldsymbol{G(\lambda, T)}$.
- Since every eigenvector of T is a generalized eigenvector of T, we have that $E(\lambda, T) \subset G(\lambda, T)$.
- We now characterize generalized eigenspaces.

  Theorem 8.5. *Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then $G(\lambda, T) = \text{null}(T - \lambda I)^{\dim V}$.*

  Proof. Suppose first that $v \in (T - \lambda I)^{\dim V}$. Then by the definition of $G(\lambda, T)$, $v \in G(\lambda, T)$, as desired. Now suppose that $v \in G(\lambda, T)$. Then $(T - \lambda I)^j v = 0$ for some positive integer j. Thus, $v \in \text{null}(T - \lambda I)^j$. We divide into two cases ($j < \dim V$ and $j \geq \dim V$). If $j < \dim V$, then by Theorem 8.1, $v \in \text{null}(T - \lambda I)^j \subset \text{null}(T - \lambda I)^{\dim V}$, as desired. On the other hand, if $j \geq \dim V$, then by Theorem 8.3 $v \in \text{null}(T - \lambda I)^j = \text{null}(T - \lambda I)^{\dim V}$, as desired. ∎

- We now prove an analogous result to Theorem 5.2 for generalized eigenvectors.

  Theorem 8.6. *Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and $v_1, \dots, v_m$ are corresponding generalized eigenvectors. Then $v_1, \dots, v_m$ is linearly independent.*

  Proof. Suppose $a_1, \dots, a_m \in \mathbb{F}$ are numbers such that

$$0 = a_1 v_1 + \cdots + a_m v_m$$

  We will prove that each $a_j = 0$ one at a time. Let's begin.

  Let $j \in \{1, \dots, n\}$ be arbitrary, and let k be the largest nonnegative integer such that $(T - \lambda_j I)^k v_j \neq 0$. Let

$$w = (T - \lambda_j I)^k v_j$$

  Then by the definition of k,

$$\begin{aligned} (T - \lambda_j I)w = (T - \lambda_j I)^{k+1} v_1 = 0 \\ Tw = \lambda_j w \end{aligned}$$

  It follows that for any $\lambda \in \mathbb{F}$, $(T - \lambda I)w = (\lambda_j - \lambda)w$, which in turn implies that

$$(T - \lambda I)^n w = (\lambda_j - \lambda)^n w$$

  for any $\lambda \in \mathbb{F}$ where $n = \dim V$. Thus, we have that

$$\begin{aligned} (T - \lambda_j I)^k \prod_{\substack{i=1 \\ i \neq j}}^m (T - \lambda_i I)^n 0 &= (T - \lambda_j I)^k \prod_{\substack{i=1 \\ i \neq j}}^m (T - \lambda_i I)^n (a_1 v_1 + \cdots + a_m v_m) \\ 0 &= a_j (T - \lambda_j I)^k \prod_{\substack{i=1 \\ i \neq j}}^m (T - \lambda_i I)^n v_j \\ &= a_j (T - \lambda_j I)^k \prod_{\substack{i=1 \\ i \neq j}}^m (\lambda_j - \lambda_i)^n v_j \\ &= a_j \prod_{\substack{i=1 \\ i \neq j}}^m (\lambda_j - \lambda_i)^n (T - \lambda_j I)^k v_j \\ &= a_j \prod_{\substack{i=1 \\ i \neq j}}^m (\lambda_j - \lambda_i)^n w \end{aligned}$$

  so $a_j = 0$, as desired. ∎

- **Nilpotent** (operator): An operator T such that $T^j = 0$ for some positive integer j.
- We now show that we never need to raise a nilpotent operator to a $j > \dim V$ to make it equal to zero.

  Theorem 8.7. *Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $N^{\dim V} = 0$.*

  Proof. Since N is nilpotent, we know that there exists a nonnegative integer j such that

  $$(N - 0I)^j v = N^j v = 0 = 0v$$

  for any $v \in V$. Thus, $G(0, N) = V$. It follows by Theorem 8.5 that $V = G(0, N) = \operatorname{null}(N - 0I)^{\dim V} = \operatorname{null} N^{\dim V}$. Consequently, for any $v \in V$, $N^{\dim V} v = 0$, so $N^{\dim V} = 0$, as desired. ■

- We now show that if N is nilpotent, there exists a basis of V such that $\mathcal{M}(N)$ is more than half zeroes.

  Theorem 8.8. *Suppose N is a nilpotent operator on V. Then there is a basis of V with respect to which the matrix of N has the form*

  $$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

  , i.e., where all entries on and below the diagonal are zeroes.

  Proof. First choose a basis of $\operatorname{null} N$. Then extend this to a basis of $\operatorname{null} N^2$, then to a basis of $\operatorname{null} N^3$, on and on up until we have extended it to a basis $v_1, \dots, v_n$ of $\operatorname{null} N^{\dim V}$ (which, incidentally, will be a basis of V since $\operatorname{null} N^{\dim V} = V$ by Theorem 8.7). We will prove that $\mathcal{M}(N, (v_1, \dots, v_n))$ has the desired form.

  Let k be the smallest positive integer such that $v_1 \in \operatorname{null} N^k$. Then $0 = N^k v_1 = N^{k-1} N v_1$, so $N v_1 \in \operatorname{null} N^{k-1} = \{0\}$ by the condition on k. It follows that $N v_1 = 0$, so since $v_1, \dots, v_n$ is linearly independent (as a basis), $\mathcal{M}(N, (v_1, \dots, v_n))_{\cdot,1} = \mathcal{M}(N v_1)$ has only zero entries. Apply the same argument to any other vector in $\operatorname{null} N^k$, getting all zero columns for some number of columns. Having done this, move onto the first vector in the basis that is not in $\operatorname{null} N^k$. Let this vector be v_i. Then in a similar fashion to before, $N v_i \in \operatorname{null} N^k$, so $N v_i$ is a linear combination of all vectors before v_i. Thus, all nonzero entries in $\mathcal{M}(()N, (v_1, \dots, v_n))_{\cdots,i} = \mathcal{M}(()N v_i)$ are above the diagonal. We continue in this fashion for the whole basis. ■