The basic goal we have is to understand linear operators $T : V \to V$ of a finite-dimensional vector space V. We assume that our vector space is over the field $\mathbb{C}$. We begin by reviewing ideas that we've considered before in specific examples, and then proceed to the theory of generalized eigenspaces.

1 Review

We present some results that we have seen before.

Theorem 1. *Every linear operator of a finite-dimensional vector sapce $T : V \to V$ has an eigenvalue.*

Proof. Suppose that $\dim V = n < \infty$ and $T \in \mathcal{L}(V, V)$. Take any $0 \neq v \in V$. The set

$$\left\{v, Tv, T^2v, \ldots, T^{n-1}v, T^nv\right\}$$

is linearly dependent because its cardinality is $n + 1$ and $\dim V = n$. Hence, there exist coefficients α_i, not all of which are zero, such that

$$\alpha_n T^n v + \alpha_{n-1}T^{n-1}v + \cdots + \alpha_2 T^2 v, \alpha_1 Tv, +\alpha_0 v = 0$$

and therefore Tv vanishes on the polynomial

$$p(z) = \alpha_n z^n + \alpha_{n-1}z^{n-1} + \cdots + \alpha_2 z^2 + \alpha_1 z + \alpha_0$$

Since the field of complex numbers is algebraically closed, $p(z)$ splits into linear factors up to a nonzero constant

$$p(z) = \alpha(z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_m).$$

We now have that

$$c(T - \lambda_1 I)(T - \lambda_2 I)\cdots(T - \lambda_m I)v = 0.$$

This gives us that $T - \lambda_j I$ is not injective for some j. That is, T has an eigenvalue. $\square$

Proposition 1. *Nonzero eigenvectors corresponding to distinct eigenvalues of T are linearly independent.*

Proof. Suppose that $v_1, \ldots, v_m$ are eigenvectors of an operator T corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. If $v_1, \ldots, v_m$ are linearly dependent there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$, not all of which are zero, such that

$$w = \alpha_1 v_1 + \cdots + \alpha_m v_m$$

is the zero vector.

This gives us that

$$
\begin{aligned}
0 &= \alpha_1 T v_1 + \cdots + \alpha_m T v_m \\
&= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_m \lambda_m v_m
\end{aligned}
$$

The operator $S = \prod_{i \neq j} (T - \lambda_i I) w = \alpha_j v_j = 0$ which yields $\alpha_j = 0$ for all j. $\square$

The following example reminds us of the ideal situation where the eigenvectors of the matrix span the vector space.

Example 1. Consider the operator on $\mathbb{C}^3$ given by the matrix

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

We first wish to produce the eigenvalues for A. We note that if $(A - \lambda I)$ is not injective, then

$$
\begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -2 \\ 0 & 1 & 3-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} (1-\lambda)v_1 \\ -\lambda v_2 - 2v_3 \\ v_2 + (3-\lambda)v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

which gives us the system

$$
\begin{aligned}
(1-\lambda)v_1 &= 0 \\
-\lambda v_2 - 2v_3 &= 0 \\
v_2 + (3-\lambda)v_3 &= 0
\end{aligned}
$$

The first is satisfied by $\lambda = 1$ or $v_1 = 0$. The second and third is satisfied when $v_2 = v_3 = 0$. Let's see if we can get some mileage from these observations. We discuss them below:

$\lambda = \mathbf{1}$:

$$
\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -v_2 - v_3 \\ v_2 + 2v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

which gives us the system:

$$-v_2 - 2v_3 = 0$$
$$v_2 + 2v_3 = 0$$

This implies that $v_2 = v_3 = 0$. Hence, $E_{\lambda=1}$ is spanned by $(1, 0, 0)^T$ and $(0, -2, 1)^T$.

$\mathbf{v_1} = \mathbf{0}$:

$$\begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -2 \\ 0 & 1 & 3-\lambda \end{pmatrix} \begin{pmatrix} 0 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -\lambda v_2 - 2v_3 \\ v_2 + (3-\lambda)v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

yields the system

$$-\lambda v_2 = 2v_3$$
$$v_2 = (\lambda - 3)v_3$$

Solving for λ, one obtains either $\lambda = 1$ or $\lambda = 2$.

$\lambda = \mathbf{2}$:

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & -2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -v_1 \\ -2v_2 - 2v_3 \\ v_2 + v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which gives the system

$$-v_1 = 0$$
$$-2v_2 - 2v_3 = 0$$
$$v_2 + v_3 = 0$$

with solution $v_1 = 0, v_2 = -v_3$. Hence $E_{\lambda=2}$ is spanned by $(0, 1, -1)^T$.

We have the eigenvalues $\lambda = 1$ and $\lambda = 2$. We also see that the eigenvectors span $\mathbb{C}^3$.

The *eigenspaces* are

$$E_{\lambda=1} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\rangle$$

$$E_{\lambda=2} = \left\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\rangle.$$

It isn't difficult to show that these three vectors span $\mathbb{C}^3$. In fact,

$$\mathbb{C}^3 = E_{\lambda=1} \oplus E_{\lambda=2}.$$

2 Geralized Eigenvectors

Consider the operator on $\mathbb{C}^2$ given by the matrix

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

If M has an eigenvector v, then there exists λ such that $v \in \operatorname{Ker}(M - \lambda I)$. Writing $(v_1, v_2)^T$,

$$\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} (1-\lambda)v_1 + v_2 \\ (1-\lambda)v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which yields the system

$$\begin{aligned} (1-\lambda)v_1 + v_2 &= 0 \\ (1-\lambda)v_2 &= 0 \end{aligned}$$

This reduces to $\lambda = 1$ or $v_2 = 0$. Each case is analyzed below:

$\boldsymbol{\lambda = 1}$:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2 \\ 0 \end{pmatrix}$$

which implies that $v_2 = 0$.

$\mathbf{v_2 = 0}$:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ 0 \end{pmatrix} = \begin{pmatrix} (1-\lambda)v_1 \\ 0 \end{pmatrix}$$

which implies that $\lambda = 1$.

Hence, the eigenspace $E_{\lambda=1}$ is 1-dimensional and spanned by $(1,0)^T$. Our analysis shows that the eigenvectors do not span V. Such a matrix is said to be *defective*. We wish to find a basis for V related to $M - I$. We note the operator $(M - I)^2$ is the zero operator.

We have $(1,0)^T \in \operatorname{Ker}(M - I)$. Can we get a vector from $\operatorname{Ker}(M - I)^2$ to obtain a basis for V? Well, $\operatorname{Ker}(M - I)^2 = V$, so we can take $(0,1)^T$.

Definition 1. Let λ be an eigenvalue for the operator $T : V \to V$. A vector $v \in \operatorname{Ker}(T - \lambda I)^k$, for some $k \in \mathbb{N}$, is said to be a *generalized eigenvector*.

In our example above, $(1,0)^T$ and $(0,1)^T$ form the *generalized eigenspace* for $\lambda = 1$, and the generalized eigenspace is $\mathbb{C}^2$.

Example 2. Consider the operator on $\mathbb{C}^3$ given by the matrix

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

We first with to produce the eigenvalues for B, which occur when the $(B - \lambda I)$ is not injective:

$$\begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 2-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which yields the system:

$$\begin{aligned} (1-\lambda)v_1 + v_2 + v_3 &= 0 \\ (2-\lambda)v_2 + 2v_3 &= 0 \\ (2-\lambda)v_3 &= 0 \end{aligned}$$

Notice that this system is satisfied when $\lambda = 1$ and $v_2 = v_3 = 0$. Another solution is given by $\lambda = 2$ and $v_3 = 0$. Here's the discussion of these two situations:

$\lambda = \mathbf{1}$:

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_2 + v_3 \\ v_2 + 2v_3 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which gives us that $0 = v_2 = v_3$ and v_1 is independent. So, the eigenspace $E_{\lambda=1}$ is spanned by $(1,0,0)^T$.

$\lambda = \mathbf{2}$:

$$\begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -v_1 + v_2 + v_3 \\ 2v_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We see that $v_3 = 0$ and $v_1 = v_2$. So the eigenspace $E_{\lambda=2}$ is spanned by $(1,1,0)^T$. Here we have the situation where the direct sum of the eigenspaces is a proper subspace of $\mathbb{C}^3$.

Now, we wish to take a look at $\operatorname{Ker}(B-2I)^3$:

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

which has null space spanned by

$$(1,1,0)^T \text{ and } (0,1,-1)^T.$$

The *generalized eigenspaces* are

$$U_{\lambda=1} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle$$

and

$$U_{\lambda=2} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle.$$

One can check that

$$\mathbb{C}^3 = U_{\lambda=1} \oplus U_{\lambda=2}$$

This example demonstrates three ideas that come out in the general theory:

- The generalized eigenvectors of T associated to the eigenvalue $\lambda = 2$ equals $\operatorname{Ker}(T-2I)^3$.
- Generalized eigenvectors associated to distinct eigenvalues of T are linearly independent.
- The generalized eigenvectors of T span V.

3 The Theory of Generalized Eigenspaces

Lemma 1. *The set of generalized eigenvectors of T corresponding to an eigenvalue λ equals* $\operatorname{Ker}(T-\lambda I)^n$, *where n is the dimension of V.*

Proof. Every element of $\operatorname{Ker}(T-\lambda I)^n$ is a generalized eigenvector by definition.

Suppose that $v \in \operatorname{Ker}(T-\lambda I)^k$ for some $k < n$ the least natural number that gets the job done. Consider the set

$$\left\{v, (T-\lambda I)v, (T-\lambda I)^2 v, \ldots, (T-\lambda I)^{k-1} v\right\}.$$

If $\alpha_i \in \mathbb{C}$, $1 \leq i \leq k-1$, not all zero, such that $0 = \sum\limits_{m=0}^{k-1} \alpha_i (T-\lambda I)^m v$. Then

$$\begin{aligned} 0 &= (T-\lambda I)^{k-1}\left(\sum_{m=0}^{k-1}(T-\lambda I)^m\right) \\ &= \sum_{m=0}^{k-1} \alpha_i (T-\lambda I)^{k+m-1} \\ &= a_0 (T-\lambda I)^{k-1} v \end{aligned}$$

which implies that $a_0 = 0$.

Applying $(T-\lambda I)^{k-2}$ in this manner we get $\alpha_1 = 0$. Carrying on, we see that $\alpha_i = 0$ for $0 \leq i \leq k-1$. The statement follows. $\square$

Proposition 2. *Nonzero generalized eigenvectors corresponding to distinct eigenvalues are linearly independent.*

Proof. Suppose that $v_1, \ldots, v_m$ are nonzero generalized eigenvectors of T corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. If the v_i are linearly dependent there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$, not all of which are zero, such that $w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m$ is the zero vector.

Let $k \in \mathbb{N}$ be the smallest such that $(T-\lambda I)^k v_1 = 0$. Apply the linear operator

$$(T-\lambda_1 I)^{k-1}(T-\lambda_2 I)^{\dim V} \cdots (T-\lambda_m I)^{\dim V}$$

to get

$$0 = \alpha_1 (T-\lambda_1)^{k-1}(T-\lambda_2 I)^{\dim V} \cdots (T-\lambda_m I)^{\dim V} v_1.$$

Rewriting the operator

$$(T-\lambda_2 I)^{\dim V} \cdots (T-\lambda_m I)^{\dim V}$$

as

$$((T-\lambda_1 I)+(\lambda_1-\lambda_2)I)^{\dim V}\cdots((T-\lambda_1 I)+(\lambda_1-\lambda_m)I)^{\dim V}$$

we can expand each factor via the binomial theorem to observe that each term contains some power of $(T-\lambda_1 I)$ except for the term

$$(\lambda_1-\lambda_2)^{\dim V}\cdots(\lambda_1-\lambda_m)^{\dim V} I.$$

Applying $(T-\lambda_1 I)^{k-1}$ on the left and v_1 to the right gives us zero and the relation

$$\alpha_1(\lambda_1-\lambda_2)^{\dim V}\cdots(\lambda_1-\lambda_m)^{\dim V}(T-\lambda_1 I)^{k-1}v_1=0.$$

This gives us that $\alpha_1=0$ due to our choice of k. One can get that $\alpha_j=0$ for $1\leq j\leq m$ in a similar fashion. $\square$

Theorem 2. *The generalized eigenvectors of T span V.*

Proof. The proof will proceed by induction on the dimension of V. The result holds for $\dim V=1$ because every operator $V\to V$ has an eigenvalue. Hence, the associated eigenvector will span the space.

Suppose that $\dim V>1$ and that the result holds for vector spaces of smaller dimension. Let λ be an eigenvalue of T.

Claim: $V=\operatorname{Ker}(T-\lambda I)^{\dim V}\oplus\operatorname{Im}(T-\lambda I)^{\dim V}$.

Suppose V is in the intersetion of the two summands. Then $(T-\lambda I)^{\dim V}=0$ and there exists $w\in V$ such that $(T-\lambda I)^{\dim V}w=v$. This gives us that

$$(T-\lambda I)^{2\dim V}w=(T-\lambda I)v=0.$$

That is w is a generalized eigenvector of T. Since the geralized eigenvectors all lie in $\operatorname{Ker}(T-\lambda I)^{\dim V}$, we have that

$$(T-\lambda I)^{\dim V}w=0$$

which is impossible since we took v to be nonzero and the image of w. We conclude that the intersection is trivial. The dimensions of the direct summands add up properly as a consequence of the rank-nullity theorem. We've proven the claim.

Because $(T-\lambda I)^{\dim V}$ is not injective by assumption, $\operatorname{Ker}(T-\lambda I)^{\dim V}$ is nontrivial. Further, $\operatorname{Im}(T-\lambda I)^{\dim V}$ has dimension strictly less than $\dim V$. Applying the inductive hypothesis, we have that $\operatorname{Im}(T-\lambda I)^{\dim V}$ is spaned by the generalized eigenvectors of T restricted to $\operatorname{Im}(T-\lambda I)^{\dim V}$. $\square$

We end this section with a theorem describing a decomposition of V into generalized eigenspaces. This will be important in obtaining the Jordan form of a matrix.

Theorem 3. *Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of an operator T, with $U_{\lambda_1}, \ldots, U_{\lambda_m}$ the corresponding generalized eigenspaces. Then*

(a) $V = U_{\lambda_1} \oplus U_{\lambda_2} \oplus \cdots \oplus U_{\lambda_m}$

(b) Each U_{λ_j} is T-invariant.

(c) Each $(T - \lambda_j I)^{\dim V}$ restricted to U_j is nilpotent.

(d) For each j, T restricted to U_j has only one eigenvalue, namely λ_j.

Proof. (a) We now that the generalized eigenvectors of T span V. Further, the sum is direct because the generalized eigenvectors associated to distinct eigenvalues are independent.

(b) Suppose that $v \in U_{\lambda_j}$. Then there exists some $k \in \mathbb{N}$ such that $(T - \lambda_j I)^k v = 0$. We need to show that $Tv \in U_{\lambda_j}$. This follows from

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

and therefore $Tv \in \operatorname{Ker}(T - \lambda_j)^k$, which implies that $Tv \in U_{\lambda_j}$.

(c) We've seen that

$$\operatorname{Ker}(T - \lambda_j I)^{\dim V} = U_{\lambda_j}.$$

This gives us that $(T - \lambda_j I)^{\dim V}$ is the zero operator on U_{λ_j}. The statement follows.

(d) Suppose that $\lambda \neq \lambda_j$ is some eigenvalue of T restricted to U_{λ_j}. Then $v \in U_{\lambda_j}$ gives

$$\begin{aligned}(T - \lambda_j I)v &= Tv - \lambda_j v \\ &= \lambda v - \lambda_j v \\ &= (\lambda - \lambda_j)v\end{aligned}$$

and we see that

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

We now have that $\lambda = \lambda_j$ since $v \neq 0$.

$\square$