Lecture 14: eigenspaces and the main diagonalizability theorem

Last time we saw a sufficient condition for a linear transformation to be diagonalizable. If $T : V \to V$ and dim $V = n$ then T is diagonalizable if it has n distinct eigenvalues. But this is of course not necessary: consider T to be the identity operator. Then every nonzero vector is an eigenvector with eigenvalue 1. But of course T is diagonalizable since its matrix form (relative to any basis) is the identity, a diagonal matrix.

Today we look more into the necessary conditions. For this we define the eigenspace E_λ.

Definition 0.1. *If $\lambda \in \mathbb{F}$ then the eigenspace*

$$E_\lambda = \{v \in V : T(v) = \lambda v\} = N(\lambda I - T) \ .$$

Note that λ is an eigenvalue of T if and only if $E_\lambda \neq \{\vec{0}\}$.

The eigenspace E_λ is the set of all eigenvectors associated to λ, unioned with $\vec{0}$. Just as eigenvectors for distinct eigenvalues are linearly independent, so are the eigenspaces.

Theorem 0.2. *Let $T : V \to V$ be linear. If $\lambda_1, \ldots, \lambda_k$ are distinct (not necessarily eigenvalues) then $E_{\lambda_1}, \ldots, E_{\lambda_k}$ are independent:*

$$E_{\lambda_1} + \cdots + E_{\lambda_k} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \ .$$

Proof. In the homework, you showed that for subspaces $A_1, \ldots, A_k$, the following are equivalent.

1. $A_1, \ldots, A_k$ are independent.
2. Whenever B_i is a basis of A_i for $i = 1, \ldots, k$, the set $B = \cup_{i=1}^k B_i$ is a basis for $A_1 + \cdots + A_k$.
3. Whenever $v_i \in A_i$ for $i = 1, \ldots, k$ and $v_1 + \cdots + v_k = \vec{0}$, all v_i's must be $\vec{0}$.

So let $v_i \in E_{\lambda_i}$ for all i be such that $v_1 + \cdots + v_k = \vec{0}$. By way of contradiction, assume they are not all zero. This means that for some subset S of $\{1, \ldots, k\}$, the vectors v_i for $i \in S$ are nonzero (and are thus eigenvectors) and $\sum_{i \in S} v_i = \vec{0}$. But then we have a linear combination of eigenvectors for distinct eigenvectors equal to zero. Linear independence gives a contradiction. $\square$

We can now give the main diagonalizability theorem.

Theorem 0.3 (Main diagonalizability theorem). *Let $T : V \to V$ be linear with dim $V = n < \infty$. The following are equivalent.*

1. *T is diagonalizable.*
2. *$V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T.*
3. *We can write $c_T(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$, where $n_i =$ dim E_{λ_i}.*

Proof. Assume that T is diagonalizable. Then we can find a basis B for V consisting of eigenvectors for T. Each of these vectors is associated with a particular eigenvalue, so write $\lambda_1, \ldots, \lambda_k$ for the distinct ones. We can then group together the elements of B associated with λ_i, span them, and call the resulting subspace E_{λ_i}. It follows then that

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} = E_{\lambda_1} + \cdots + E_{\lambda_k} = \text{Span } B = V \ .$$

So 1 implies 2.

Now assume that 2 holds. Then build a basis B_i of size n_i of each E_{λ_i}. Since the λ_i's are eigenvalues, the B_i's consist of eigenvectors for eigenvalue λ_i and $n_i \geq 1$ for all i. Since we assumed item 2, the set $B = \cup_{i=1}^k B_i$ is a basis for V and $[T]_B^B$ is a diagonal matrix with distinct entries $\lambda_1, \ldots, \lambda_k$, with λ_i repeated n_i times. By computing the characteristic polynomial c_T through this matrix, we find item 3 holds. This proves 2 implies 3.

Last if 3 holds, then $n_1 + \cdots + n_k = n$, because c_T has degree n. Therefore

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} = n \ .$$

Because each λ_i is a root of c_T, it is an eigenvalue and therefore has an eigenvector. This means each E_{λ_i} has dimension at least 1. Let B_i be a basis of E_{λ_i}. As the λ_i's are distinct, the E_{λ_i}'s are independent and so

$$E_{\lambda_1} + \cdots + E_{\lambda_i} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} \ .$$

This means $B = \cup_{i=1}^k B_i$ is a basis for the sum and therefore is linearly independent. Since it has size n, it is a basis for V. But each vector of B is an eigenvector for T so T is diagonalizable. □

Let's finish by giving an example. We would like to check if a matrix is diagonalizable. Let A be the real matrix

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

So we compute the characteristic polynomial.

$$c_A(x) = \det \begin{pmatrix} x-1 & 1 & 1 \\ -1 & x-1 & -1 \\ 0 & 0 & x-1 \end{pmatrix} = (x-1) \cdot \det \begin{pmatrix} x-1 & 1 \\ -1 & x-1 \end{pmatrix} \ .$$

Here we have used the formula for the determinant of a block upper-triangular matrix. So it equals

$$(x-1)((x-1)^2+1) = (x-1)(x^2-2x+2) \ .$$

The last factor does not have any roots in $\mathbb{R}$. Therefore we cannot write c_A in the form $(x-\lambda_1)^{n_1} \cdots (x-\lambda_k)^{n_k}$ and A is not diagonalizable.

On the other hand, if we consider A as a complex matrix, it is diagonalizable. This is because the characteristic polynomials factors

$$(x-1)(x-a)(x-\bar{a}) \ ,$$

where $a = 1+i$ (where $i = \sqrt{-1}$) and $\bar{a}$ is the complex conjugate of a. Since A has 3 distinct eigenvalues and the dimension of $\mathbb{C}^3$ is 3, the matrix is diagonalizable.

If, however, the characteristic polynomial were

$$c_A(x) = (x-1)(x-i)^2 ,$$

then we would have to investigate further. For instance this would be the case if we had

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

In this case if we write $c_A(x) = (x-\lambda_1)^{n_1} \cdots (x-\lambda_k)^{n_k}$, we have $\lambda_1 = 1$, $\lambda_2 = i$ and $n_1 = 1$, $n_2 = 2$. If A were diagonalizable, then we would need (from the previous theorem) $1 = n_1 = \dim E_1$ and $2 = n_2 = \dim E_i$. However, note that

$$iI - A = \begin{pmatrix} i-1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} ,$$

which has nullity 1, not 2. This means that $\dim E_i = 1$ and A is not diagonalizable.

Jordan form

For a diagonalizable transformation, we have a very nice form. However it is not always true that a transformation is diagonalizable. For example, it may be that the roots of the characteristic polynomials are not in the field (as in the example above). Even if the roots are in the field, they may not have multiplicities equal to the dimensions of the eigenspaces. So we look for a more general form. Instead of looking for a diagonal form, we look for a *block diagonal* form. That is, we want to write

$$A = \begin{pmatrix} B & 0 & 0 & \cdots \\ 0 & C & 0 & \cdots \\ 0 & 0 & D & \cdots \\ & & & \cdots \end{pmatrix} ,$$

where $B, C, D, \ldots$ are square matrices.

For this purpose we define generalized eigenspaces.

Definition 0.4. *Let $T : V \to V$ be linear. Given $\lambda \in \mathbb{F}$, the generalized eigenspace $\hat{E}_\lambda$ is the subspace*

$$\hat{E}_\lambda = \{v \in V : (\lambda I - T)^k v = \vec{0} \text{ for some } k \in \mathbb{N}\} .$$

Note that $E_\lambda \subset \hat{E}_\lambda$ and we can write

$$\hat{E}_\lambda = \cup_{k=1}^{\infty} N(\lambda I - T)^k \ .$$

Although in general if A, B are subspaces of V then $A \cup B$ need not be a subspace, we know it is a subspace if $A \subset B$. Because

$$N(\lambda I - T) \subset N(\lambda I - T)^2 \subset N(\lambda I - T)^3 \subset \cdots \ ,$$

you can verify that $\hat{E}_\lambda$ is a subspace.

When $\dim V < \infty$ we know that $T : V \to V$ is diagonalizable if and only if V can be written as a direct sum of eigenspaces. Even if T is not diagonalizable, if the field $\mathbb{F}$ is algebraically closed, we can always write V as a direct sum of generalized eigenspaces.

Theorem 0.5 (Primary decomposition theorem)**.** *Let $T : V \to V$ be linear with dim $V < \infty$. If $\mathbb{F}$ is algebraically closed, then*

$$V = \hat{E}_{\lambda_1} \oplus \cdots \oplus \hat{E}_{\lambda_k} \ ,$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T.