Math 201 lecture for Monday, Week 10

Diagonalization algorithm

Recall from last time: an *eigenvector* for a linear transformation $f\colon V \to V$ is a *nonzero* vector $v \in V$ such that

$$f(v) = \lambda v$$

for some $\lambda \in F$. In that case, λ is called an *eigenvalue* for f.

Definition. Let V be an n-dimensional vector space. A linear mapping $f\colon V \to V$ is *diagonalizable* if there exists an ordered basis α of V such that $[f]_\alpha^\alpha = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. A matrix $A \in M_{n\times n}(F)$ is *diagonalizable* if its corresponding linear mapping f_A is diagonalizable.

Proposition. A linear mapping $f\colon V \to V$ is diagonalizable if and only if V has a basis consisting solely of eigenvectors for f.

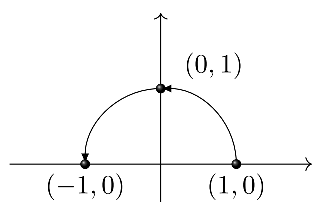
Proof. Let α be any ordered basis. Then $[f]_\alpha^\alpha$ is diagonal if and only if, for each $j = 1, \ldots, n$, the j-th column of $[f]_\alpha^\alpha$ has a single non-zero entry, in the j-th row. That j-th column is determined by

$$f(v_j) = 0 \cdot v_1 + \cdots + 0 \cdot v_{j-1} + \lambda_j \cdot v_j + 0 \cdot v_{j+1} + \cdots + 0 \cdot v_n,$$

for some scalar λ_j. However, the above condition is equivalent to $f(v_j) = \lambda_j v_j$ for $j = 1, \ldots, n$, i.e., to α being a basis of eigenvectors. $\square$

Example. Not all linear transformations of a vector space to itself are diagonalizable. For instance, consider the linear transformation $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ that is rotation of the plane by $90°$, having matrix

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



There is no point $0 \neq v \in \mathbb{R}^2$ such that $Av = \lambda v$ for some λ. (The matrix is diagonalizable over $\mathbb{C}$, though. Can you find two eigenvectors? Don't get your hopes up, though—there are matrices that are not diagonalizable over $\mathbb{C}$.)

Suppose $f\colon F^n \to F^n$ is a linear transformation, and let A be the matrix corresponding to f, i.e., the matrix whose j-th column is $f(e_j)$ for all j (i.e., the matrix for f with respect to the standard basis for F^n). Suppose we can find a basis $\alpha = \langle v_1, \ldots, v_n \rangle$ of eigenvectors for f with corresponding, not necessarily distinct, eigenvalues $\lambda_1, \ldots, \lambda_n$. Let P be the matrix with columns $v_1, \ldots, v_n$. Then, as we saw last time,

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Definition. Two $n \times n$ matrices A and B over F are *similar* or *conjugate* if there exists an invertible matrix P such that $A = P^{-1}BP$.

Exercise. The reader should verify that similarity is an equivalence relation.

Remark. Let $f\colon V \to V$ be a linear transformation of a finite-dimensional vector space, and let α and β be two ordered bases for V. Then we saw earlier in the semester that the matrices $A := [f]_\alpha^\alpha$ and $B := [f]_\beta^\beta$ are conjugate, i.e., the matrices for f with respect to any two bases for V are conjugate. The converse is also true: every matrix conjugate to A is the matrix representing f with respect to some basis.

Finding eigenvectors and eigenvalues. Let $A \in M_{n\times n}(F)$ with corresponding linear function

$$\begin{aligned} f_A\colon F^n &\to F^n \\ v &\mapsto Av. \end{aligned}$$

As mentioned last time, the following argument is of central importance in the story of eigenvectors and eigenvalues: We are looking for nonzero $v \in F^n$ and any $\lambda \in F$ such that $Av = \lambda v$. We have

$$Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I_n)v = 0 \quad \Leftrightarrow \quad v \in \ker(A - \lambda v).$$

This says that:

> $\lambda \in F$ is an eigenvalue for A if and only if $\ker(A - \lambda I_n) \neq \{0\}$.

So we would like to determine those λ for which the kernel of $A - \lambda I_n$ is nontrivial. The following is key:

$$\ker(A - \lambda I_n) \neq \{0\} \quad \Leftrightarrow \quad \operatorname{rank}(A - \lambda I_n) < n \quad \Leftrightarrow \quad \det(A - \lambda I_n) = 0.$$

Definition. The *characteristic polynomial* of A is

$$p_A(x) := \det(A - xI_n).$$

We have just seen that

> $\lambda \in F$ is an eigenvalue for A if and only if it is a zero of the characteristic polynomial for A, i.e., if and only if $p_A(\lambda) = 0$.

Example. Let

$$A = \begin{pmatrix} 2 & -7 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

The characteristic polynomial of A is

$$p_A(t) = \det\left(\begin{pmatrix} 2 & -7 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix} - x\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$$

$$= \det\left(\begin{pmatrix} 2 & -7 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix} - \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}\right)$$

$$= \det\begin{pmatrix} 2-x & -7 & 3 \\ 0 & -5-x & 3 \\ 0 & 0 & 2-x \end{pmatrix}$$

$$= (2-x)(-5-x)(2-x)$$

$$= -(x-2)^2(x+5).$$

Thus, $p_A(x) = 0$ if and only if $x \in \{2, -5\}$. So the eigenvalues of A are 2 (with *multiplicity* 2), and -5.

Recall that our goal is to diagonalize A by finding a basis of eigenvectors. That's not always possible, but we can try. The **first step** is to compute the zeros of the characteristic polynomial, $p_A(x)$. This tells us the eigenvalues for A. We then need to find the eigenvectors to go along with these eigenvalues. Recall that nonzero $v \in F^n$ is an eigenvector for A with eigenvalue λ if and only $v \in \ker(A - \lambda I_n)$.

Definition. Let λ be an eigenvalue of the $n \times n$ matrix A over F. Then the *eigenspace for* λ is

$$E_\lambda := E(A)_\lambda := \{v \in F^n : Av = \lambda v\} = \ker(A - \lambda I_n v).$$

The eigenspace, being the kernel of a matrix, is a linear subspace of F^n.

The **second step** in trying to diagonalize A is to compute a basis for each eigenspace E_λ.

Example. We have seen that the eigenvalues for

$$A = \begin{pmatrix} 2 & -7 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

are 2 (with multiplicity 2) and -5. Let's compute the corresponding eigenspaces in the case $F = \mathbb{R}$.

$\boxed{E_2}$

$$A - 2I_3 = \begin{pmatrix} 2 & -7 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -7 & 3 \\ 0 & -7 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 0 & 1 & -3/7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The first and third variables are free. Hence,

$$\ker(A - 2I_3) = \left\{(x, \tfrac{3}{7}z, z) : x, z \in \mathbb{R}\right\}.$$

For a basis we could take $\left\{(1,0,0), (0, \frac{3}{7}, 1)\right\}$, or easier, $\{(1,0,0), (0,3,7)\}$.

$\boxed{E_{-5}}$

$$A - (-5)I_3 = \begin{pmatrix} 2 & -7 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 7 & -7 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 7 \end{pmatrix}$$

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

Hence,

$$\ker(A + 5I_3) = \{(y, y, 0) : y \in \mathbb{R}\}.$$

For a basis we could take $(1,1,0)$.

Thus, we have found three eigenvectors $(1,0,0)$, $(0,3,7)$, and $(1,1,0)$. It turns out that eigenvectors for distinct eigenvalues are always linearly independent (we'll see this later). Hence, we have found a basis of eigenvectors. Thus, A is diagonalizable, and if we use these eigenvectors as the columns for a matrix:

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 7 & 0 \end{pmatrix},$$

then one may check that

$$P^{-1}AP = \operatorname{diag}(2, 2, -5).$$

Example. Now consider a matrix that is just slightly different from A:

$$B = \begin{pmatrix} 2 & 1 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

The characteristic polynomial for A and for B are the same:

$$\det(B - xI_3) = \det \begin{pmatrix} 2-x & 1 & 3 \\ 0 & -5-x & 3 \\ 0 & 0 & 2-x \end{pmatrix} = -(x-2)^2(x+5).$$

Thus, A and B have the same eigenvalues. Let's compute the eigenspaces for B over $\mathbb{R}$.

$\boxed{E_2}$

$$
\begin{aligned}
B - 2I_3 &= \begin{pmatrix} 2 & 1 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 3 \\ 0 & -7 & 3 \\ 0 & 0 & 0 \end{pmatrix} \\
&\longrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Thus, $\ker(B - 2I_3)$ has basis $\{(1,0,0)\}$. It is only one-dimensional. Recall that $\ker(A - 2I_3)$ was two-dimensional. This is a crucial difference.

$\boxed{E_{-5}}$

$$
\begin{aligned}
A - (-5)I_3 &= \begin{pmatrix} 2 & 1 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} \\
&= \begin{pmatrix} 7 & 1 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 7 \end{pmatrix} \\
&\longrightarrow \begin{pmatrix} 1 & 1/7 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Hence,

$$\ker(A - 3I_3) = \{(y, -y/7, 0) : y \in \mathbb{R}\}.$$

For a basis we could take $(-7, 1, 0)$.

Our calculations prove that, at most, we can find two linearly independent vectors that are eigenvectors for B. Thus, there is no basis for $\mathbb{R}^3$ consisting of eigenvectors for B. Therefore, B is not diagonalizable.

Diagonalizing Algorithm Let $A \in M_{n\times n}(F)$.

(a) Find the eigenvalues of A as the zeros of its characteristic polynomial,

$$p_A(x) = \det(A - xI_n).$$

(b) For each eigenvalue λ, compute a basis for the eigenspace $E_\lambda = \ker A - \lambda I_n$.

(c) **The matrix A is diagonalizable if and only if of the total number of eigenvectors in the bases found in the previous step is n. i.e., if and only if the sum of the dimensions of the eigenspaces is n.** If so, the union of these vectors is a basis for F^n. Create a matrix P whose columns are these vectors. Then $P^{-1}AP = D$, where D is a diagonal matrix with the eigenvalues along the diagonal, and we get a corresponding commutative diagram:

$$\begin{tikzcd} F^n \arrow[r, "A"] \arrow[d, "P^{-1}"', "\wr"] & F^n \arrow[d, "P^{-1}", "\wr"'] \\ F^n \arrow[r, "D"] & F^n. \end{tikzcd}$$

The matrix P^{-1}, considered as a linear function, takes coordinates with respect to the basis of eigenvalues.

Remark. An $n\times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. Step (c) of the diagonalization algorithm depends on a fact we will prove next time: eigenvectors with distinct eigenvalues are linearly independent. (We compute bases for each eigenspace, and of course the elements in a basis are linearly independent. But when we combine the bases for all of the eigenspaces, why is the resulting set independent?)