Math 201 lecture for Wednesday, Week 10

Eigenspaces

Before getting started, we make an observation which should have probably already been mentioned:

Proposition. Let A, B be $n \times n$ matrices representing a linear function $f\colon V \to V$ with respect to different bases. Then their characteristic polynomials are the same: $p_A(x) = p_B(x)$.

Proof. We have $A = P^{-1}BP$ for some $n \times n$ matrix P. Then

$$\begin{aligned}
p_A(x) &= \det(A - xI_n) \\
&= \det(P^{-1}BP - xI_n) \\
&= \det(P^{-1}BP - xP^{-1}I_nP) \\
&= \det(P^{-1}BP - P^{-1}(xI_n)P) \qquad (x \text{ is a scalar}) \\
&= \det(P^{-1}(B - xI_n)P) \\
&= \det(P^{-1})\det(B - xI_n)\det(P) \\
&= \det(B - xI_n).
\end{aligned}$$

For the last step, recall that $\det(P^{-1}) = \det(P)^{-1}$, which follows from multiplicativity of the determinant:

$$1 = \det(I_n) = \det(P^{-1}P) = \det(P^{-1})\det(P).$$

□

Thus, it makes sense to talk about the **characteristic polynomial of a linear transformation**: it the characteristic polynomial of any matrix representing the transformation.

Last time, we discussed the following algorithm that determines whether a matrix is diagonalizable and, if it is, shows how to diagonalize it.

Diagonalization 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.** In other words, A is diagonalizable if and only if $\sum_\lambda \dim E_\lambda = n$ where the sum is over all eigenvalues λ of A. If so, then 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.

As mentioned last time, Step (c) of the diagonalization algorithm depends on the fact that eigenvectors with distinct eigenvalues are linearly independent. (Thus, when we combine the bases for all of the eigenspaces, we end up with a linearly independent set.) We now prove this.

Proposition. Let V be any vector space, and let $f\colon V \to V$ be a linear transformation. Let $v_1, \ldots, v_k \in V$ be eigenvectors for f with corresponding eigenvalues λ_i:

$$f(v_i) = \lambda_i v_i$$

for $i = 1, \ldots, k$. Suppose $\lambda_1, \ldots, \lambda_k$ are *distinct*. Then $v_1, \ldots, v_k$ are linearly independent.

Proof. We will prove this by induction on k. The case $k = 1$ is OK since, by definition, an eigenvector is a nonzero vector. Suppose $v_1, \ldots, v_{k-1}$ are linearly independent for some $k > 1$ and that

$$a_1 v_1 + \cdots + a_k v_k = 0$$

for some $a_i \in F$. Let id_V be the identity transformation defined by $\mathrm{id}_V(v) = v$ for all $v \in V$. Apply the linear transformation $f - \lambda_k \mathrm{id}_V$ to the above dependence relation to get

$$\begin{aligned}
&(f - \lambda_k \mathrm{id}_V)(a_1 v_1 + \cdots + a_k v_k) = (f - \lambda_k \mathrm{id}_V)(0) = 0 \\
\Rightarrow\quad & f(a_1 v_1 + \cdots + a_k v_k) - \lambda_k \mathrm{id}_V(a_1 v_1 + \cdots + a_k v_k) = 0 \\
\Rightarrow\quad & (a_1 \lambda_1 v_1 + \cdots + a_k \lambda_k v_k) - (a_1 \lambda_k v_1 + \cdots + a_k \lambda_k v_k) = 0 \\
\Rightarrow\quad & a_1(\lambda_1 - \lambda_k) v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} + a_k(\lambda_k - \lambda_k) v_k = 0 \\
\Rightarrow\quad & a_1(\lambda_1 - \lambda_k) v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} = 0
\end{aligned}$$

Since $v_1, \ldots, v_{k-1}$ are linearly independent, all the coefficients are zero:

$$a_1(\lambda_1 - \lambda_k) = \cdots = a_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Since the λ_i are distinct, this implies $a_1 = \cdots = a_{k-1} = 0$. Therefore, the original equation, $a_1 v_1 + \cdots + a_k v_k = 0$ becomes $a_k v_k = 0$. Since v_k is an eigenvector, it is nonzero. Hence, $a_k = 0$, as well. □

Corollary. Suppose $\dim V = n$ and $f\colon V \to V$ is a linear transformation. Then if f has n distinct eigenvalues, it is diagonalizable.

Proof. Each eigenvalue has at least one corresponding eigenvector. From the above proposition, if f has n distinct eigenvalues, then it has n linearly independent eigenvectors. Since V has dimension n, these eigenvectors form a basis for V. Let α be an ordered basis consisting of those eigenvectors. Then $[f]_\alpha^\alpha$ is diagonal. □

Remark. The Proposition implies that the union of bases for the eigenspaces of A forms a linearly independent sets. For instance, for convenience, suppose that A has three (distinct) eigenvalues λ_1, λ_2, and λ_3, and suppose the corresponding eigenspaces have bases $\{u_1, \ldots, u_p\}$, $\{v_1, \ldots, v_q\}$, and $\{w_1, \dots, w_r\}$, respectively. We would like to show that the union of these sets is linearly dependent. So suppose we have a relation

$$a_1u_1 + \cdots + a_pu_p + b_1v_1 + \cdots + b_qv_q + c_1w_1 + \cdots + c_rw_r = 0.$$

Let $u = \sum_{i=1}^p a_iu_i$, $v = \sum_{i=1}^q b_iv_i$, and $w = \sum_{i=1}^r c_iw_i$. Then we have $u \in E_{\lambda_1}$, $v \in E_{\lambda_2}$, and $w \in E_{\lambda_3}$ and

$$u + v + w = 0.$$

By the Proposition, must have $u = v = w = 0$. Otherwise, this relation would be a nontrivial linear relation among eigenvectors with distinct eigenvalues. (Note that the only element of an eigenspace that is not an eigenvector is the zero vector.)

Warning. The converse to the corollary is not true. For instance, consider the identity function on F^n. Its matrix is I_n, which is already diagonal, and 1 is its only eigenvalue:

$$p_{I_n}(x) = \det(I_n - xI_n) = \det\left((1-x)I_n\right) = (1-x)^n \det(I_n) = (1-x)^n.$$

So I_n is diagonalizable (in fact, it's already diagonal) even though its eigenvalues are not distinct.

Cramer's rule

Definition. Let $A \in M_{n\times n}(F)$. For $i, j \in \{1, \dots, n\}$, let $A^{ij} \in M_{(n-1)\times(n-1)}(F)$ be the matrix formed by removing the i-th row and j-th column of A. The *i, j-th minor* of A is $\det(A^{ij})$, and the *i, j-th cofactor* of A is $(-1)^{i+j}\det(A^{ij})$. The *adjugate* of A is the matrix $\operatorname{adj}(A) \in M_{n\times n}(F)$ with i, j-th coordinate

$$\operatorname{adj}(A)_{ij} = (-1)^{i+j}\det(A^{ji}).$$

(Note we are using A^{ji}, not A^{ij}.)

Theorem (Cramer's rule). Let $A \in M_{n\times n}(F)$ be an invertible matrix, and let $b \in F^n$. Then the solution to the system of linear equations $Ax = b$ is given by

$$x_j = \frac{\det(M_j)}{\det(A)}$$

for $j = 1, \dots, n$ where $M_j \in M_{n\times n}(F)$ is the matrix formed by replacing the j-th column of A with b.

Corollary. If $A \in M_{n\times n}(F)$ is invertible, then

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$$

where $\operatorname{adj}(A)$ is the *adjugate* of A, defined by

Corollary. If $A \in M_{n\times n}(F)$ is invertible and $F = \mathbb{R}$ of $F = \mathbb{C}$, then

(a) the solution for the system of equations $Ax = b$ is a continuous function of the entries of A and b, and

(b) the entries of A^{-1} are continuous functions of the entries of A.

Proof. The entries in the determinant of a matrix B are polynomials in the entries of B. A quotient f/g of polynomials f and g is a continuous function wherever g is nonzero. $\square$

Example. Consider the matrix

$$A = \begin{pmatrix} 3 & -1 & 6 \\ -7 & 1 & 2 \\ 2 & 0 & 2 \end{pmatrix}.$$

The adjugate of A is

$$\operatorname{adj}(A) = \begin{pmatrix} 2 & 2 & -8 \\ 18 & -6 & -48 \\ -2 & -2 & -4 \end{pmatrix}.$$

For instance, to find the $1,2$-entry of $\operatorname{adj}(A)$ is

$$(-1)^{1+2}\det(A^{2,1}) = (-1)^3 \det \begin{pmatrix} -1 & 6 \\ 0 & 2 \end{pmatrix} = 2.$$

Using Cramer's rule to compute the inverse of A, we get

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A) = -\frac{1}{24}\begin{pmatrix} 2 & 2 & -8 \\ 18 & -6 & -48 \\ -2 & -2 & -4 \end{pmatrix} = \begin{pmatrix} -\frac{1}{12} & -\frac{1}{12} & \frac{1}{3} \\ -\frac{3}{4} & \frac{1}{4} & 2 \\ \frac{1}{12} & \frac{1}{12} & \frac{1}{6} \end{pmatrix}.$$