Math 201 lecture for Friday, Week 10

Algebraic and geometric multiplicity. Jordan form.

When does a transformation fail to be diagonalizable? We now introduce a sequence of ideas that will allow us to answer this question.

Example. Earlier, we considered the linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ given by the matrix

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

Geometrically, it rotates the plane counterclockwise by 90° and, hence, has no eigenvectors: an eigenvector would not rotate—it would just be scaled. The characteristic polynomial of A is

$$p_A(x) = \det \begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} = x^2 + 1.$$

The equation $x^2 + 1 = 0$ has no solutions over $\mathbb{R}$, and hence, the transformation has no eigenvalues. Now consider the linear transformation $f\colon \mathbb{C}^2 \to \mathbb{C}^2$ given by the same matrix A. Over $\mathbb{C}$ we can solve $x^2 + 1 = 0$ to find two eigenvalues, $\pm i$. Each of these will have at least one eigenvector, and eigenvectors for distinct eigenvalues are linearly independent. Since $\mathbb{C}^2$ has dimension 2, that means we will get a basis of eigenvectors. Let's compute a basis for the eigenspace for i:

$$A - iI_2 = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix} \xrightarrow{r_2 \leftrightarrow r_2 + ir_1} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}.$$

So the kernel of $A - iI_2$ is $\{(iy, y) : y \in \mathbb{C}\}$, which has basis $\{(i, 1)\}$. Similarly, the eigenspace for $-i$ has basis $\{(-i, 1)\}$. Check:

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

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

Letting

$$P = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix},$$

we get

$$P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

This example illustrates one obstacle to diagonalization: the characteristic polynomial may not have enough roots in the field F.

Definition. A polynomial $p \in F[x]$ *splits over* F if there exist $c, \lambda_1, \dots, \lambda_n \in F$ such that

$$p(x) = c(x_1 - \lambda_1) \cdots (x - \lambda_n).$$

Equivalently, $p(x)$ had n roots (zeros), $\lambda_1, \dots, \lambda_n$, in F. These λ_i need not be distinct.

Remark. Let F be any field, and let $p(x)$ be a polynomial whose coefficients are in F, i.e., $p(x) \in F[x]$. It turns out that there exists a field $F \subseteq K$ such that $p(x)$ splits over K.

Example. The polynomial $p(x) = x^2 + 1$ splits over $\mathbb{C}$ but not over $\mathbb{R}$.

A useful fact from algebra:

Theorem. (Fundamental theorem of algebra) Every $p \in \mathbb{C}[x]$ splits over $\mathbb{C}$.

Proposition. Let V be a vector space over F with $\dim V = n$, and let $f\colon V \to V$ be a linear transformation. If f is diagonalizable, then its characteristic polynomial splits over F.

Proof. Let $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ be a diagonal matrix representing f. Then the characteristic polynomial for f (which, as we saw earlier, in the last lecture, does not depend on the choice of matrix representative) is

$$p_f(x) = p_D(x) = \det(D - xI_n) = (\lambda_1 - x) \cdots (\lambda_n - x) = (-1)^n (x_1 - \lambda_1) \cdots (x - \lambda_n).$$

$\square$

The converse of this proposition is not true:

Example. Let

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

The characteristic polynomial of A is

$$\begin{aligned} p_A(x) &= \det(A - xI_2) \\ &= \det \begin{pmatrix} 1-x & 1 \\ 0 & 1-x \end{pmatrix} \\ &= (x-1)^2. \end{aligned}$$

Thus, the characteristic polynomial splits over any field F. There is one eigenvalue, 1, which occurs with algebraic multiplicity 2 (the precise definition of *algebraic multiplicity* appears below). Let's proceed with the algorithm for diagonalization by computing a basis for the eigenspace for 1, i.e., for $\ker(A - I_2)$:

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

Therefore, $\ker(A - I_2) = \{(x, 0) : x \in F\}$. A basis is $\{(1, 0)\}$. Thus, there is no basis for F^2 consisting of eigenvectors: our theory says any eigenvector would have to have eigenvalue 1, and the space of eigenvectors with eigenvalue 1 is only one-dimensional!

Definition. Let $\dim V < \infty$. The *algebraic multiplicity* of an eigenvalue $\lambda \in F$ for a linear transformation $f\colon V \to V$ (or for any matrix representing f) is the largest number m such that $p_f(x) = (x - \lambda)^m q(x)$ for some polynomial $q(x) \in F[x]$.

The *geometric multiplicity* of λ is the dimension of the eigenspace $E_\lambda(f)$ for λ:

$$\dim E_\lambda(f) = \dim \ker(f - \lambda \operatorname{id}_V).$$

So if A is a matrix representing f, then the geometric multiplicity of $\lambda \in F$ is

$$\dim E_\lambda(A) = \dim \ker(A - \lambda\, I_n).$$

Remark. To rephrase something we already know: $A \in M_{n\times n}(F)$ **is diagonalizable if and only if the sum of its geometric multiplicities is n.** That's because this is the only case in which we have enough linearly independent eigenvectors to form a basis of eigenvectors.

Proposition. Let $\dim V < \infty$, and let λ be an eigenvalue of a linear transformation $f\colon V \to V$. Then the geometric multiplicity of λ is at most the algebraic multiplicity of λ.

Proof. Let $v_1, \ldots, v_k$ be a basis for $\ker(f - \lambda\, \mathrm{id}_V)$, and extend it to a basis $v_1, \ldots, v_n$ for all of V. We have $f(v_i) = \lambda v_i$ for $1 = 1, \ldots, k$. So with respect to our chosen basis, the matrix representing f has the form

$$A := \begin{pmatrix} \lambda I_k & B \\ 0 & C \end{pmatrix},$$

where B and C are $(n-k) \times (n-k)$ matrices. So the characteristic polynomial for f is

$$\begin{aligned} p_f(x) &= \det \begin{pmatrix} (\lambda - x)I_k & B \\ 0 & C - xI_{n-k} \end{pmatrix} \\ &= \det((\lambda - x)I_k)\det(C - xI_{n-k}) \\ &= (\lambda - x)^k \det(C - xI_{n-k}) \\ &= (\lambda - x)^k q(x), \end{aligned}$$

for some polynomial $q(x)$. (To see the second equality, above, expand the determinant in line 1 along the first *column*—there will only be one term, which will be $\lambda - x$ times a smaller determinant. Expand that determinant along its first column. Repeat k times, each time picking up a factor of $\lambda - x$.) This shows that the algebraic multiplicity of λ is at least k, the geometric multiplicity of λ. □

Corollary. Let $A \in M_{n\times n}(F)$. Then A is diagonalizable if and only if its characteristic polynomial splits over F and the geometric multiplicity and algebraic multiplicity of each eigenvalue are equal.

Proof. Suppose that A is diagonalizable. We saw earlier in this lecture that the characteristic polynomial for A then splits over F. So we can write

$$p_A(x) = (-1)^n \prod_{i-1}^{k} (x - \lambda_i)^{m_i}$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A. The degree of $p_A(x)$ is n (exercise!), from which it follows that $\sum_{i=1}^{k} m_i = n$, i.e, the sum of the algebraic multiplicities is n.

Let g_i be the geometric multiplicity of eigenvalue λ_i. Since A is diagonalizable, we know that the sum of its geometric multiplicities is also n.

Therefore, we have $n = \sum_{i=1}^k g_i = \sum_{i=1}^k m_i$, and by the Proposition, $g_i \le m_i$ for all i. Since the m_i are nonnegative, it follows that $m_i = g_i$ for all i.

Conversely, suppose that $p_A(x)$ splits and that the algebraic and geometric multiplicities of each eigenvalue are equal. Factor $p_A(x)$ as above and using same notation for algebraic and geometric multiplicities. As before, since the degree of $p_A(x) = n$, we have $n = \sum_{i=1}^k m_i$. By assumption, $m_i = g_i$ for all i. So it follows that the sum of the geometric multiplicities is n, and hence, A is diagonalizable. $\square$

Jordan form. What can we say when a linear transformation is not diagonalizable? Can we still choose a basis to make the matrix for the transformation simple in some sense? We give one answer here. First, we need a couple definitions. A *Jordan block of size k for $\lambda \in F$* is the $k \times k$ matrix with λs on the diagonal and 1s on the "superdiagonal":

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \cdots & 0 & 0 \\ & & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

For example,

$$J_4(3) = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Note the following example of the special case of a Jordan block of size 1:

$$J_1(5) = [5].$$

A matrix is in *Jordan form* if it is in block diagonal form with Jordan blocks for various λ along the diagonal:

$$\begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & J_{k_m}(\lambda_m) \end{pmatrix}$$

For example, here is a matrix in Jordan form:

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

It has two 2×2 Jordan blocks for 2, a 1×1 Jordan block for 5, and a 3×3 Jordan block for 4:

$$\begin{pmatrix} J_2(2) & 0 & 0 & 0 \\ 0 & J_2(2) & 0 & 0 \\ 0 & 0 & J_1(5) & 0 \\ 0 & 0 & 0 & J_3(4) \end{pmatrix}.$$

Theorem. Let $\dim V < \infty$. Suppose $f\colon V \to V$ is a linear transformation over F and that the characteristic polynomial for f splits, i.e., the field F contains all of the zeros of the characteristic polynomial. Then there exists an ordered basis for V such that the matrix representing f with respect to that basis is in Jordan form. The Jordan form is unique up to a permutation of the Jordan blocks.

So a matrix is diagonalizable if and only if its characteristic polynomial splits and all of its Jordan blocks have size 1. We also know that a matrix such as

$$\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

which is already in Jordan form but not diagonal, is not diagonalizable.