Math 1600B Lecture 30, Section 2, 21 Mar 2014

Announcements:

Today we finish 4.4. **Read** Markov chains part of Section 4.6 for next class. Not covering Section 4.5, or rest of 4.6 (which contains many interesting applications!) Work through recommended homework questions.

Next class: **course evaluations**.

Tutorials: Quiz 9 (the last one) will cover Section 4.4.

Help Centers: Monday-Friday 2:30-6:30 in MC 106.

Office hour: Monday, 1:30-2:30, MC103B.

The **final exam** will take place on Tuesday, April 22, 2-5pm. All students write in AH201 (Alumni Hall). See the course home page for final exam **conflict** policy. You should **immediately** notify the registrar or your Dean (and your instructor) of any conflicts!

True/false: Every polynomial of degree n has exactly n distinct roots over $\mathbb{C}$. **False**. For example, $(x-1)^2$ has only the root 1. A polynomial of degree n has exactly n complex roots if you count them with multiplicity. Over $\mathbb{R}$, the sum of the multiplicities is *at most* n.

True/false: The complex eigenvalues of a matrix always come in conjugate pairs. **False**. This is true if the matrix has only real entries, but $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$ has i as an eigenvalue, but not $\bar{i}$.

True/false: If λ is an eigenvalue of A and $k \geqslant 0$, then λ^k is an eigenvalue of A^k. **True**, since if $A\vec{x} = \lambda\vec{x}$, then $A^k\vec{x} = A^{k-1}\lambda\vec{x} = \lambda^2 A^{k-2}\vec{x} = \cdots = \lambda^k\vec{x}$.

Review of Section 4.4

Definition: Let A and B be $n \times n$ matrices. We say that A is **similar** to B if there is an invertible matrix P such that $P^{-1}AP = B$. When this is the case, we write $A \sim B$.

It is equivalent to say that $AP = PB$ or $A = PBP^{-1}$.

It is tricky in general to find such a P when it exists. We'll learn a method that works in a certain situation in this section.

Theorem 4.21: Let A, B and C be $n \times n$ matrices. Then:
a. $A \sim A$.
b. If $A \sim B$ then $B \sim A$.
c. If $A \sim B$ and $B \sim C$, then $A \sim C$.

Similar matrices have a lot of properties in common.

Theorem 4.22: Let A and B be similar matrices. Then:
a. $\det A = \det B$
b. A is invertible iff B is invertible.
c. A and B have the same rank.
d. A and B have the same characteristic polynomial.
e. A and B have the same eigenvalues.

Question: Are $\begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$ similar?

True/false: The identity matrix is similar to every matrix.
False. Since $PIP^{-1} = I$ for any invertible P, the identity matrix is only similar to itself.

True/False: If A and B have the same eigenvalues, then A and B are similar.

False. For example, I and $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ have the same eigenvalues, but aren't similar.

Diagonalization

Definition: A is **diagonalizable** if it is similar to some diagonal matrix.

If A is similar to a diagonal matrix D, then D must have the eigenvalues of A on the diagonal. But how to find P?

Theorem 4.23: Let A be an $n \times n$ matrix. If P is an $n \times n$ matrix whose columns are linearly independent eigenvectors of A, then $P^{-1}AP$ is a diagonal matrix D with the corresponding eigenvalues of A on the diagonal.

On the other hand, if P is any invertible matrix such that $P^{-1}AP$ is diagonal, then the columns of P are linearly independent eigenvectors of A.

It follows that A is diagonalizable if and only if it has n linearly independent eigenvectors.

This theorem is one of the main reasons we want to be able to find eigenvectors of a matrix. Moreover, the more eigenvectors the better, so this motivates allowing complex eigenvectors.

New material: Section 4.4 continued

Proof of Theorem 4.23: Suppose $\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n$ are n linearly independent eigenvectors of A, and let $P = [\, \vec{p}_1 \, \vec{p}_2 \, \cdots \, \vec{p}_n \,]$. Write λ_i for the ith eigenvalue, so $A\vec{p}_i = \lambda_i \vec{p}_i$ for each i, and let D be the diagonal matrix with the λ_i's on the diagonal. Then

$$AP = A[\, \vec{p}_1 \, \vec{p}_2 \, \cdots \, \vec{p}_n \,] = [\, \lambda_1 \vec{p}_1 \;\; \lambda_2 \vec{p}_2 \;\; \cdots \;\; \lambda_n \vec{p}_n \,]$$

Also,

$$\begin{aligned} PD &= [\, \vec{p}_1 \, \vec{p}_2 \, \cdots \, \vec{p}_n \,] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\ &= [\, \lambda_1 \vec{p}_1 \;\; \lambda_2 \vec{p}_2 \;\; \cdots \;\; \lambda_n \vec{p}_n \,] \end{aligned}$$

so $P^{-1}AP = D$, as required. (Why is P invertible?)

On the other hand, if $P^{-1}AP = D$ and D is diagonal, then $AP = PD$, and if follows from an argument like the one above that the columns of P are eigenvectors of A, and the eigenvalues are the diagonal entries of D. $\square$

So we'd like to be able to find enough linearly independent eigenvectors of a matrix. Recall that in Section 4.3, we saw:

Theorem 4.20: If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, then $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ are linearly independent.

Example: Is the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ diagonalizable?

Yes. The eigenvalues are 1, 4 and 6, and for each one there is at least one eigenvector. These are linearly independent (by Theorem 4.20), and there are three of them, so A is diagonalizable (by Theorem 4.23).

To **find** the matrix P explicitly, we need to solve the three systems to find the eigenvectors.

Theorem 4.25: If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.

Example 4.25: Is $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$ diagonalizable? If so, find a matrix P that diagonalizes it.

Solution: In Example 4.18 we found that the eigenvalues are $\lambda = 1$ (with algebraic multiplicity 2) and $\lambda = 2$ (with algebraic multiplicity 1). A basis for E_1 is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and a basis for E_2 is $\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$. Since every eigenvector is a scalar multiple of one of these, it is not possible to find three linearly independent eigenvectors. So A is not diagonalizable.

Example 4.26: Is $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$ diagonalizable? If so, find a matrix P that diagonalizes it.

Solution: In Example 4.19 (done mostly on board, but also in text) we found that the eigenvalues are $\lambda = 0$ (with algebraic multiplicity 2) and $\lambda = -2$ (with algebraic multiplicity 1). A basis for E_0 is $\vec{p}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{p}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. A basis for E_{-2} is $\vec{p}_3 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$. These are linearly independent (see below). Thus

$$P = [\,\vec{p}_1\,\vec{p}_2\,\vec{p}_3\,] = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{bmatrix}$$

is invertible, and by the theorem, we must have

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} = D$$

(Note that to check an answer like this, it is usually easiest to check that $AP = PD$. Do so!)

Note: Different orders of eigenvectors/values work too.

Theorem 4.24: If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of A and, for each i, $\mathcal{B}_i$ is a basis for the eigenspace E_{λ_i}, then the union of the $\mathcal{B}_i$'s is a linearly independent set.

The proof of this is similar to the proof of Theorem 4.20, where we had only one non-zero vector in each eigenspace.

Combining Theorems 4.23 and 4.24 gives the following important consequence:

Theorem: An $n \times n$ matrix is diagonalizable if and only if the sum of the geometric multiplicities of the eigenvalues is n.

Look at Examples 4.25 and 4.26 again.

So it is important to understand the geometric multiplicities better. Here is a helpful result:

Lemma 4.26: If λ_1 is an eigenvalue of an $n \times n$ matrix A, then

$$\text{geometric multiplicity of } \lambda_1 \leqslant \text{algebraic multiplicity of } \lambda_1$$

We'll prove this in a minute. First, let's look at what it implies:

Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Let their geometric multiplicities be $g_1, g_2, \ldots, g_k$ and their algebraic multiplicities be $a_1, a_2, \ldots, a_k$. We know

$$g_i \leqslant a_i \quad \text{for each } i$$

and so

$$g_1 + \cdots + g_k \leqslant a_1 + \cdots + a_k \leqslant n$$

So the only way to have $g_1 + \cdots + g_k = n$ is to have $g_i = a_i$ for each i and $a_1 + \cdots + a_k = n$.

This gives the **main theorem** of the section:

Theorem 4.27 (The Diagonalization Theorem): Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Let their geometric multiplicities be $g_1, g_2, \ldots, g_k$ and their algebraic multiplicities be $a_1, a_2, \ldots, a_k$. Then the following are equivalent:
a. A is diagonalizable.
b. $g_1 + \cdots + g_k = n$.
c. $g_i = a_i$ for each i and $a_1 + \cdots + a_k = n$.

Note: This is stated incorrectly in the text. The red part must be added unless you are working over $\mathbb{C}$, in which case it is automatic that $a_1 + \cdots + a_k = n$. With the way I have stated it, it is correct over $\mathbb{R}$ or over $\mathbb{C}$.

Example: Is $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ diagonalizable?

It depends. If we are working over $\mathbb{R}$, there are no eigenvalues and no eigenvectors, so no, it is not diagonalizable, and (a), (b) and (c) all fail.

If we are working over $\mathbb{C}$, then i and $-i$ are eigenvalues, and are distinct, so A is diagonalizable, and (b) and (c) hold too. Note that in this case, P will be complex: To find P, we first find that corresponding eigenvectors are $\vec{p}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$ and $\vec{p}_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$. So if we take

$$P = [\,\vec{p}_1\,\vec{p}_2\,] = \begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}$$

we find that

$$P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = D$$

Summary of diagonalization: Given an $n \times n$ matrix A, we would like to determine whether A is diagonalizable, and if it is, find the invertible matrix P and the diagonal matrix D such that $P^{-1}AP = D$. The result may depend upon whether you are working over $\mathbb{R}$ or $\mathbb{C}$.

Steps:

1. Compute the characteristic polynomial $\det(A - \lambda I)$ of A.
2. Find the roots of the characteristic polynomial and their algebraic multiplicities by factoring.
3. If the algebraic multiplicities don't add up to n, then A is not diagonalizable, and you can stop. (If you are working over $\mathbb{C}$, this can't happen.)
4. For each eigenvalue λ, compute the dimension of the eigenspace E_λ. This is the geometric multiplicity of λ, and if it is less than the algebraic multiplicity, then A is not diagonalizable, and you can stop.
5. Compute a basis for the eigenspace E_λ.
6. If for each eigenvalue the geometric multiplicity equals the algebraic multiplicity, then you take the n eigenvectors you found and put them in the columns of a matrix P. Put the eigenvalues in the same order on the diagonal of a matrix D.
7. **Check** that $AP = PD$.

Note that step 4 only requires you to find the row echelon form of $A - \lambda I$, as the number of free variables here is the geometric multiplicity. In step 5, you solve the system.

I still owe you a proof of:

Lemma 4.26: If λ_1 is an eigenvalue of an $n \times n$ matrix A, then

$$\text{geometric multiplicity of } \lambda_1 \leqslant \text{algebraic multiplicity of } \lambda_1$$

Proof (more direct than in text): Suppose that λ_1 is an eigenvalue of A with geometric multiplicity g, and let $\vec{v}_1, \ldots, \vec{v}_g$ be a basis for E_{λ_1}, so

$$A\vec{v}_i = \lambda_1 \vec{v}_i \quad \text{for each } i.$$

Let Q be an invertible matrix whose first g columns are $\vec{v}_1, \ldots, \vec{v}_g$:

$$Q = [\,\vec{v}_1 \ \cdots \ \vec{v}_g \quad \text{other vectors} \quad]$$

Since $Q^{-1}Q = I$, we know that $Q^{-1}\vec{v}_i = \vec{e}_i$ for $1 \leqslant i \leqslant g$. Also, the first g columns of AQ are $\lambda_1\vec{v}_1, \ldots, \lambda_1\vec{v}_g$. So the first g columns of $Q^{-1}AQ$ are $\lambda_1\vec{e}_1, \ldots, \lambda_1\vec{e}_g$. Therefore the matrix $Q^{-1}AQ$ has λ_1 as an eigenvalue with algebraic multiplicity at least g. But $Q^{-1}AQ$ has the same characteristic polynomial as A, so λ_1 must also have algebraic multiplicity at least g for A. $\square$.