Announcements

October 31

- WeBWorK assignment 5.2 is due on Friday at 6am.
- Your midterms will be returned to you in recitation on Friday.
  - The solutions are posted online.
  - The score breakdown is on Piazza.
- Midterm 2 will take place in recitation on Friday, 11/18.
- Office hours: Wednesday 1–2pm, Thursday 3:30–4:30pm, and by appointment, in Skiles 221.
  - As always, TAs' office hours are posted on the website.
  - Math Lab is also a good place to visit.

Section 5.3

Diagonalization

Motivation

Many real-word linear algebra problems have the form:

$$v_1 = Av_0 \quad v_2 = Av_1 = A^2 v_0 \quad v_3 = Av_2 = A^3 v_0 \quad \ldots \quad v_n = Av_{n-1} = A^n v_0.$$

The question is, what happens to v_n as $n \to \infty$?

Our toy example about rabbit populations had this form.

- Taking powers of diagonal matrices is easy!
- Taking powers of *diagonalizable* matrices is still easy!
- Figuring out if a matrix is diagonalizable is an eigenvalue problem.

Powers of Diagonal Matrices

If A is diagonal, then A^n is also diagonal; its diagonal entries are the nth powers of the diagonal entries of A:

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad D^2 = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \quad D^3 = \begin{pmatrix} 8 & 0 \\ 0 & 27 \end{pmatrix} \quad \cdots \quad D^n = \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix}.$$

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \quad D^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{9} \end{pmatrix} \quad D^3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{27} \end{pmatrix}$$

$$\cdots \quad D^n = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & 0 & \frac{1}{3^n} \end{pmatrix}$$

Powers of Matrices that are Similar to Diagonal Ones

What if A is not diagonal?

Example

Let $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$. Compute A^n.

In §5.2 lecture we saw that A was similar to a diagonal matrix:

$$A = PDP^{-1} \quad \text{where} \quad P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Then

$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PD^2P^{-1} \\ A^3 &= (PDP^{-1})(PD^2P^{-1}) = PD(P^{-1}P)D^2P^{-1} = PDID^2P^{-1} = PD^3P^{-1} \\ &\vdots \\ A^n &= PD^nP^{-1} \end{aligned}$$

Therefore

$$A^n = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2^{n+1} - 3^n & -2^{n+1} + 2 \cdot 3^n \\ 2^n - 3^n & -2^n + 2 \cdot 3^n \end{pmatrix}.$$

Diagonalizable Matrices

Definition

An $n \times n$ matrix A is **diagonalizable** if it is similar to a diagonal matrix:

$$A = PDP^{-1} \quad \text{for } D \text{ diagonal.}$$

Important

If $A = PDP^{-1}$ for $D = \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}$ then

$$A^k = PD^K P^{-1} = P \begin{pmatrix} d_{11}^k & 0 & \cdots & 0 \\ 0 & d_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn}^k \end{pmatrix} P^{-1}.$$

The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In this case, $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the corresponding eigenvalues (in the same order).

Corollary

← a theorem that follows easily from another theorem

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

The Corollary is true because eigenvectors with distinct eigenvalues are always linearly independent. We will see later that a diagonalizable matrix need not have n distinct eigenvalues though.

Let $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$. The characteristic polynomial is

$$f(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{pmatrix} = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).$$

Therefore the eigenvalues are 2 and 3. Let's compute some eigenvectors:

$$(A - 2I)x = 0 \iff \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = 2y$, so $v_1 = \binom{2}{1}$ is an eigenvector with eigenvalue 2.

$$(A - 3I)x = 0 \iff \begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = y$, so $v_2 = \binom{1}{1}$ is an eigenvector with eigenvalue 3.

The eigenvectors v_1, v_2 are linearly independent, so the Diagonalization Theorem says

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Let $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$. The characteristic polynomial is

$$f(\lambda) = \det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda - 1)^2(\lambda - 2).$$

Therefore the eigenvalues are 1 and 2, with respective multiplicities 2 and 1. Let's compute some eigenvectors:

$$(A - I)x = 0 \iff \begin{pmatrix} 3 & -3 & 0 \\ 2 & -2 & 0 \\ 1 & -1 & 0 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = 0$$

The parametric vector form is

$$\begin{cases} x = y \\ y = y \\ z = \qquad z \end{cases} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Hence a basis for the 1-eigenspace is

$$\mathcal{B}_1 = \{v_1, v_2\} \quad \text{where} \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now let's compute an eigenvector with eigenvalue 2:

$$(A-2I)x=0 \iff \begin{pmatrix} 2 & -3 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & -1 \end{pmatrix} x = 0 \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} x = 0$$

The parametric form is $x = 3z, y = 2z$, so an eigenvector with eigenvalue 2 is

$$v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

The eigenvectors v_1, v_2, v_3 are linearly independent: v_1, v_2 form a basis for the 1-eigenspace, and v_3 is not contained in the 1-eigenspace. Therefore the Diagonalization Theorem says

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Note that in this case, there were three linearly independent eigenvectors, but only two distinct eigenvalues.

Example

Show that $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.

This is an upper-triangular matrix, so the only eigenvalue is 1. Let's compute the 1-eigenspace:

$$(A - I)x = 0 \iff \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x = 0.$$

This is row reduced, but has only one free variable x; a basis for the 1-eigenspace is $\{\binom{1}{0}\}$. So *all eigenvectors* of A are multiples of $\binom{1}{0}$.

Conclusion: A has only one linearly independent eigenvector, so by the "only if" part of the diagonalization theorem, A is not diagonalizable.

Poll

Which of the following matrices are diagonalizable?

A. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ **B.** $\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$ **C.** $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ **D.** $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

Matrix **A** is not diagonalizable: its only eigenvalue is 1, and its 1-eigenspace is spanned by $\binom{1}{0}$. Similarly, matrix **C** is not diagonalizable.

Matrix **B** is diagonalizable because it is a 2×2 matrix with distinct eigenvalues. Matrix **D** is already diagonal!

How to diagonalize a matrix A:

1. Find the eigenvalues of A using the characteristic polynomial.
2. For each eigenvalue λ of A, compute a basis $\mathcal{B}_\lambda$ for the λ-eigenspace.
3. If there are fewer than n total vectors in the union of all of the eigenspace bases $\mathcal{B}_\lambda$, then the matrix is not diagonalizable.
4. Otherwise, the n vectors $v_1, v_2, \ldots, v_n$ in your eigenspace bases are linearly independent, and $A = PDP^{-1}$ for

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_i is the eigenvalue for v_i.

Why is the Diagonalization Theorem true?

A diagonalizable implies A has n linearly independent eigenvectors: Suppose $A = PDP^{-1}$, where D is diagonal with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $v_1, v_2, \ldots, v_n$ be the columns of P. They are linearly independent because P is invertible. So $Pe_i = v_i$, hence $P^{-1}v_i = e_i$.

$$Av_i = PDP^{-1}v_i = PDe_i = P(\lambda_i e_i) = \lambda_i Pe_i = \lambda_i v_i.$$

Hence v_i is an eigenvector of A with eigenvalue λ_i. So the columns of P form n linearly independent eigenvectors of A, and the diagonal entries of D are the eigenvalues.

A has n linearly independent eigenvectors implies A is diagonalizable: Suppose A has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let P be the invertible matrix with columns $v_1, v_2, \ldots, v_n$. Let $D = P^{-1}AP$.

$$De_i = P^{-1}APe_i = P^{-1}Av_i = P^{-1}(\lambda_i v_i) = \lambda_i P^{-1}v_i = \lambda_i e_i.$$

Hence D is diagonal, with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$. Solving $D = P^{-1}AP$ for A gives $A = PDP^{-1}$.

Non-Distinct Eigenvalues

Theorem

Let A be an $n \times n$ matrix. Let:

- $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of A,
- $a_1, a_2, \ldots, a_k$ be the algebraic multiplicities of the eigenvalues, and
- $d_1, d_2, \ldots, d_k$ be the dimensions of the eigenspaces.

Then:

1. $1 \leq d_i \leq a_i$ for all i, and
2. A is diagonalizable if and only if $d_1 + d_2 + \cdots + d_k = n$; equivalently,
3. A is diagonalizable if and only if $d_i = e_i$ for all i, and the characteristic polynomial for A has all real roots.

The number d_i, i.e. the dimension of the λ_i-eigenspace, is called the **geometric multiplicity** of the eigenvalue λ_i. So assertion 1 states that the geometric multiplicity is bounded by the algebraic multiplicity.

Assertion 2 was used in the diagonalization procedure before: it says that there are n vectors in the bases of all the eigenspaces.

Assertion 3 says that, if A is diagonalizable, then *the geometric multiplicities must equal the algebraic multiplicities.*

If A has n distinct eigenvalues, then $k = n$ and $d_i = a_i = 1$ for all i, so the matrix is diagonalizable. For example, $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$ has eigenvalues 2 and 3, hence is diagonalizable.

The matrix $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ has characteristic polynomial $f(\lambda) = -(\lambda - 1)^2(\lambda - 2)$. If $\lambda_1 = 1$ and $\lambda_2 = 2$, then $a_1 = 2$ and $a_2 = 1$. Since the 1-eigenspace has dimension 2, we know $a_1 = d_1 = 2$. Automatically $a_2 = d_2 = 1$. So $d_1 = a_1$ and $d_2 = a_2$, and $f(\lambda)$ has all real roots, so A is diagonalizable.

The matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has characteristic polynomial $f(\lambda) = (\lambda - 1)^2$. Hence there is one eigenvalue $\lambda_1 = 1$, and $a_1 = 2$. But the 1-eigenspace is spanned by $\binom{1}{0}$, so $d_1 = 1 < 2 = a_1$. Hence A is not diagonalizable.