Section 5.4

Diagonalization

Many real-word linear algebra problems have the form:

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

This is called a **difference equation**.

Our toy example about rabbit populations had this form.

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

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

Powers of Diagonal Matrices

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

$$D = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \quad D^2 = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad D^3 = \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix}, \quad \ldots \quad D^n = \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix}.$$

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

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

Powers of Matrices that are Similar to Diagonal Ones

What if A is not diagonal?

Example

Let $A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix}$. Compute A^n, using

$$A = CDC^{-1} \quad \text{for} \quad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

We compute:

$$\begin{aligned} A^2 &= (CDC^{-1})(CDC^{-1}) = CD(C^{-1}C)DC^{-1} = CDIDC^{-1} = CD^2C^{-1} \\ A^3 &= (CDC^{-1})(CD^2C^{-1}) = CD(C^{-1}C)D^2C^{-1} = CDID^2C^{-1} = CD^3C^{-1} \\ &\vdots \\ A^n &= CD^nC^{-1} \end{aligned}$$

Closed formula in terms of n:
easy to compute

Therefore

$$A^n = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix} \frac{1}{-2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2^n + (-1)^n & 2^n + (-1)^{n+1} \\ 2^n + (-1)^{n+1} & 2^n + (-1)^n \end{pmatrix}.$$

Similar Matrices

Definition

Two $n \times n$ matrices are **similar** if there exists an invertible $n \times n$ matrix C such that $A = CBC^{-1}$.

Fact: if two matrices are similar then so are their powers:

$$A = CBC^{-1} \quad \Longrightarrow \quad A^n = CB^nC^{-1}.$$

Fact: if A is similar to B and B is similar to D, then A is similar to D.

$$A = CBC^{-1}, \quad B = EDE^{-1} \quad \Longrightarrow \quad A = C(EDE^{-1})C^{-1} = (CE)D(CE)^{-1}.$$

Diagonalizable Matrices

Definition

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

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

Important

If $A = CDC^{-1}$ 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

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

So diagonalizable matrices are easy to raise to any power.

The Diagonalization Theorem

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

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

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

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.

Diagonalization

The Diagonalization Theorem

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

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

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

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).

Note that the decomposition is not unique: you can reorder the eigenvalues and eigenvectors.

$$A = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix}^{-1} = \begin{pmatrix} | & | \\ v_2 & v_1 \\ | & | \end{pmatrix} \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} \begin{pmatrix} | & | \\ v_2 & v_1 \\ | & | \end{pmatrix}^{-1}$$

Question: What does the Diagonalization Theorem say about the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}?$$

This is a triangular matrix, so the eigenvalues are the diagonal entries $1, 2, 3$.

A diagonal matrix just scales the coordinates by the diagonal entries, so we can take our eigenvectors to be the unit coordinate vectors e_1, e_2, e_3. Hence the Diagonalization Theorem says

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

It doesn't give us anything new because the matrix was already diagonal!

A diagonal matrix D is diagonalizable! It is similar to itself:

$$D = I_n D I_n^{-1}.$$

Problem: Diagonalize $A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix}$.

The characteristic polynomial is

$$f(\lambda) = \lambda^2 - \operatorname{Tr}(A)\,\lambda + \det(A) = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2).$$

Therefore the eigenvalues are -1 and 2. Let's compute some eigenvectors:

$$(A + 1I)x = 0 \iff \begin{pmatrix} 3/2 & 3/2 \\ 3/2 & 3/2 \end{pmatrix} x = 0 \overset{\text{rref}}{\rightsquigarrow} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} x = 0$$

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

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

The parametric form is $x = y$, so $v_2 = \binom{1}{1}$ is an eigenvector with eigenvalue 2.

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

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

Problem: Diagonalize $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

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 the 1-eigenspace:

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

The parametric vector form is

$$\begin{aligned} x &= y \\ y &= y \\ z &= \quad\; z \end{aligned} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Hence a basis for the 1-eigenspace is

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

Problem: Diagonalize $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

Now let's compute the 2-eigenspace:

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

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

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

Note: In this case, there are three linearly independent eigenvectors, but only two distinct eigenvalues.

Problem: 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 $\left\{\binom{1}{0}\right\}$. 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

Poll

Which of the following matrices are diagonalizable, and why?

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 $A = CDC^{-1}$ for

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

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 $A = CDC^{-1}$, 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 C. They are linearly independent because C is invertible. So $Ce_i = v_i$, hence $C^{-1}v_i = e_i$.

$$Av_i = CDC^{-1}v_i = CDe_i = C(\lambda_i e_i) = \lambda_i Ce_i = \lambda_i v_i.$$

Hence v_i is an eigenvector of A with eigenvalue λ_i. So the columns of C 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 C be the invertible matrix with columns $v_1, v_2, \ldots, v_n$. Let $D = C^{-1}AC$.

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

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

Algebraic Multiplicity

Definition

The **(algebraic) multiplicity** of an eigenvalue λ is its multiplicity as a root of the characteristic polynomial.

This is not a very interesting notion *yet*. It will become interesting when we also define *geometric* multiplicity later.

Example

In the rabbit population matrix, $f(\lambda) = -(\lambda - 2)(\lambda + 1)^2$, so the algebraic multiplicity of the eigenvalue 2 is 1, and the algebraic multiplicity of the eigenvalue -1 is 2.

Example

In the matrix $\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$, $f(\lambda) = (\lambda - (3 - 2\sqrt{2}))(\lambda - (3 + 2\sqrt{2}))$, so the algebraic multiplicity of $3 + 2\sqrt{2}$ is 1, and the algebraic multiplicity of $3 - 2\sqrt{2}$ is 1.

Non-Distinct Eigenvalues

Definition

Let λ be an eigenvalue of a square matrix A. The **geometric multiplicity** of λ is the dimension of the λ-eigenspace.

Theorem

Let λ be an eigenvalue of a square matrix A. Then

$$1 \leq (\text{the geometric multiplicity of } \lambda) \leq (\text{the algebraic multiplicity of } \lambda).$$

The proof is beyond the scope of this course.

Corollary

Let λ be an eigenvalue of a square matrix A. If the algebraic multiplicity of λ is 1, then the geometric multiplicity is also 1: the eigenspace is a *line*.

The Diagonalization Theorem (Alternate Form)

Let A be an $n \times n$ matrix. The following are equivalent:

1. A is diagonalizable.
2. The sum of the geometric multiplicities of the eigenvalues of A equals n.
3. The sum of the algebraic multiplicities of the eigenvalues of A equals n, and for each eigenvalue, *the geometric multiplicity equals the algebraic multiplicity*.

Example

If A has n distinct eigenvalues, then the algebraic multiplicity of each equals 1, hence so does the geometric multiplicity, and therefore A is diagonalizable.

For example, $A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix}$ has eigenvalues -1 and 2, so it is diagonalizable.

Example

The matrix $A = \begin{pmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ has characteristic polynomial

$$f(\lambda) = -(\lambda - 1)^2(\lambda - 2).$$

The algebraic multiplicities of 1 and 2 are 2 and 1, respectively. They sum to 3.

We showed before that the geometric multiplicity of 1 is 2 (the 1-eigenspace has dimension 2). The eigenvalue 2 automatically has geometric multiplicity 1.

Hence the geometric multiplicities add up to 3, so A is diagonalizable.

Example

The matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has characteristic polynomial $f(\lambda) = (\lambda - 1)^2$.

It has one eigenvalue 1 of algebraic multiplicity 2.

We showed before that the geometric multiplicity of 1 is 1 (the 1-eigenspace has dimension 1).

Since the geometric multiplicity is smaller than the algebraic multiplicity, the matrix is *not* diagonalizable.

Summary

- A matrix A is **diagonalizable** if it is similar to a diagonal matrix D: $A = CDC^{-1}$.
- It is easy to take powers of diagonalizable matrices: $A^r = CD^rC^{-1}$.
- An $n \times n$ matrix is diagonalizable if and only if it has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$, in which case $A = CDC^{-1}$ for

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

- If A has n distinct eigenvalues, then it is diagonalizable.
- The **geometric multiplicity** of an eigenvalue λ is the dimension of the λ-eigenspace.
- $1 \leq$ (geometric multiplicity) $\leq$ (algebraic multiplicity).
- An $n \times n$ matrix is diagonalizable if and only if the sum of the geometric multiplicities is n.