

Section 6.4

Diagonalization

Motivation

Difference equations

Many real-word linear algebra problems have the form:

$$v_1 = Av_0, \quad v_2 = Av_1 = A^2v_0, \quad v_3 = Av_2 = A^3v_0, \quad \dots \quad v_n = Av_{n-1} = A^nv_0.$$

This is called a **difference equation**.

Our toy example about rabbit populations had this form.

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

- ▶ 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 n th powers of the diagonal entries of D :

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:

$$A^2 =$$

$$A^3 =$$

$$\vdots$$

$$A^n =$$

Therefore

$$A^n =$$

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:

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

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

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

$$C = \left(\begin{array}{c|c|ccc|c} | & | & & & | \\ \hline v_1 & v_2 & \cdots & & v_n \\ \hline | & | & & & | \end{array} \right) \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where v_1, v_2, \dots, v_n are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \dots, \lambda_n$ 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.

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

$$C = \left(\begin{array}{c|c|c|c} | & | & \cdots & | \\ v_1 & v_2 & & v_n \\ | & | & & | \end{array} \right) \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

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

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

$$A = \left(\begin{array}{c|c} | & | \\ v_1 & v_2 \\ | & | \end{array} \right) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \left(\begin{array}{c|c} | & | \\ v_1 & v_2 \\ | & | \end{array} \right)^{-1} = \left(\begin{array}{c|c} | & | \\ v_2 & v_1 \\ | & | \end{array} \right) \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} \left(\begin{array}{c|c} | & | \\ v_2 & v_1 \\ | & | \end{array} \right)^{-1}$$

Diagonalization

Easy example

Question: What does the Diagonalization Theorem say about the matrix

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

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

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

Diagonalization

Example

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

Diagonalization

Another example

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

Diagonalization

Another example, continued

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

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

Diagonalization

A non-diagonalizable matrix

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

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

Diagonalization

Procedure

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}_λ for the λ -eigenspace.
3. If there are fewer than n total vectors in the union of all of the eigenspace bases \mathcal{B}_λ , then the matrix is not diagonalizable.
4. Otherwise, the n vectors v_1, v_2, \dots, v_n in your eigenspace bases are linearly independent, and $A = CDC^{-1}$ for

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

where λ_i is the eigenvalue for v_i .

Diagonalization

Proof

Why is the Diagonalization Theorem true?

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 *the geometric multiplicity equals the algebraic multiplicity* of each eigenvalue.

Non-Distinct Eigenvalues

Examples

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.

Non-Distinct Eigenvalues

Another example

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.

- ▶ 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, \dots, v_n , in which case $A = CDC^{-1}$ for

$$C = \left(\begin{array}{c|c|ccc|c} | & | & & & | \\ \hline v_1 & v_2 & \cdots & & v_n \\ \hline | & | & & & | \end{array} \right) \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

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