

Lecture 31: Diagonalization of Square Matrices

Math 3013
Oklahoma State University

April 15, 2022

Agenda

1. Diagonal Matrices
2. Diagonalizability
3. Some Matrix Multiplication Identities
4. Connection with Eigenvalue Multiplicities
5. Other Special Cases for Diagonalizability

Diagonal Matrices

Definition

An $n \times n$ matrix is **diagonal** if its only non-zero entries occur along its main diagonal; i.e.

$$(\mathbf{A})_{ij} = 0 \quad \text{if } i \neq j$$

In other words, a diagonal matrix is a matrix of the form

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Two Special Properties of Diagonal Matrices

Let \mathbf{D} , \mathbf{D}' be diagonal matrices :

Say

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}, \quad \mathbf{D}' = \begin{bmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mu_n \end{bmatrix}$$

Then

$$\mathbf{D}\mathbf{D}' = \begin{bmatrix} \lambda_1\mu_1 & 0 & \cdots & 0 \\ 0 & \lambda_2\mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n\mu_n \end{bmatrix} \quad (1)$$

Two Special Properties of Diagonal Matrices, Cont'd

Suppose $\mathbf{e}_i \in \mathbb{R}^n$ is the i^{th} standard basis vector. Then

$$\mathbf{D}\mathbf{e}_i = \lambda_i \mathbf{e}_i \quad (*)$$

i.e., \mathbf{e}_i is an eigenvector of \mathbf{D} with eigenvalue λ_i .

N.B. Property $(*)$ makes it easy to compute the action of diagonal matrices on more general vectors. If $\mathbf{v} = [v_1, v_2, \dots, v_n] \in \mathbb{R}^n$

$$\begin{aligned} \mathbf{D}\mathbf{v} &= \mathbf{D}(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n) \\ &= v_1(\mathbf{D}\mathbf{e}_1) + v_2(\mathbf{D}\mathbf{e}_2) + \cdots + v_n(\mathbf{D}\mathbf{e}_n) \\ &= (v_1\lambda_1)\mathbf{e}_1 + (v_2\lambda_2)\mathbf{e}_2 + \cdots + (v_n\lambda_n)\mathbf{e}_n \end{aligned}$$

Thus, even if \mathbf{v} is not an eigenvector of \mathbf{D} , the action of \mathbf{D} on \mathbf{v} is relatively simple to compute (and understand).

Moral: Dealing with a matrix \mathbf{A} is much easier when the matrix is diagonal

Diagonalizability

Definition

An $n \times n$ matrix **A** is said to be **diagonalizable** if there is an **invertible** $n \times n$ matrix **C** and a diagonal matrix **D** such that

$$\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{D}$$

Today, I'll explain how to find the matrix **C** that will convert an $n \times n$ matrix **A** into a diagonal matrix **A**.

In particular, I show how the “diagonalizing matrix” **C** can be constructed from the eigenvectors of **A** and how the corresponding diagonal matrix **D** is constructed from the eigenvalues of **A**.

Digression: Some Matrix Multiplication Identities

Lemma

Let \mathbf{A} and \mathbf{C} are $n \times n$ matrices, and suppose $\mathbf{c}_1, \dots, \mathbf{c}_n$ are column vectors of \mathbf{C} . then

$$\mathbf{AC} = \left[\begin{array}{ccc} \uparrow & & \uparrow \\ (\mathbf{Ac}_1) & \cdots & (\mathbf{Ac}_n) \\ \downarrow & & \downarrow \end{array} \right]$$

Example

$$\begin{aligned}\mathbf{AC} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}c_{11} + a_{12}c_{21} & a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{11} + a_{22}c_{21} & a_{21}c_{12} + a_{22}c_{22} \end{bmatrix}\end{aligned}$$

Meanwhile

$$\begin{aligned}\mathbf{Ac}_1 &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix} = \begin{bmatrix} a_{11}c_{11} + a_{12}c_{21} \\ a_{21}c_{11} + a_{22}c_{21} \end{bmatrix} \\ \mathbf{Ac}_2 &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_{12} \\ c_{22} \end{bmatrix} = \begin{bmatrix} a_{11}c_{12} + a_{12}c_{22} \\ a_{21}c_{12} + a_{22}c_{22} \end{bmatrix}\end{aligned}$$

Digression: Some Matrix Multiplication Identities, Cont'd

Lemma

Suppose \mathbf{C} is an $n \times n$ matrix with columns $\mathbf{c}_1, \dots, \mathbf{c}_n$ and \mathbf{D} is a diagonal $n \times n$ matrix

$$\mathbf{C} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{c}_1 & \cdots & \mathbf{c}_n \\ \downarrow & & \downarrow \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Then

$$\mathbf{CD} = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ (\lambda_1 \mathbf{c}_1) & \cdots & (\lambda_n \mathbf{c}_n) \\ \downarrow & \cdots & \downarrow \end{bmatrix}$$

Proof of Lemma

Write

$$(D)_{ij} = \lambda_i \delta_{ij} \quad , \quad \text{where } \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Then

$$\begin{aligned} (\mathbf{CD})_{ij} &= \sum_{k=1}^n c_{ik} d_{kj} \\ &= \sum_{k=1}^n (\mathbf{c}_k)_i (\lambda_k \delta_{jk}) \\ &= (\mathbf{c}_j)_i \lambda_j \end{aligned}$$

So the j^{th} column of \mathbf{CD} is

$$Col_j(\mathbf{CD}) = \begin{bmatrix} \uparrow \\ \lambda_j \mathbf{c}_j \\ \downarrow \end{bmatrix}$$

A Necessary Condition for Diagonalizability

Theorem

Suppose \mathbf{A} is a diagonalizable $n \times n$ matrix; so there is an $n \times n$ matrix \mathbf{C} and a $n \times n$ diagonal matrix \mathbf{D} such that $\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{D}$. Write

$$\mathbf{C} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{c}_1 & \cdots & \mathbf{c}_n \\ \downarrow & & \downarrow \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Then for each $i = 1, \dots, n$, the i^{th} column, \mathbf{c}_i , of \mathbf{C} must be an eigenvector of \mathbf{A} with eigenvalue $\lambda_i = i^{\text{th}}$ diagonal entry of \mathbf{D} .

Proof of Theorem

Multiplying both sides of the equation $\mathbf{C}^{-1}\mathbf{AC} = \mathbf{D}$ from the left by \mathbf{C} we get

$$\mathbf{AC} = \mathbf{CD}$$

Let $\mathbf{c}_1, \dots, \mathbf{c}_n$ be the columns of \mathbf{C} . From our matrix multiplication identities

$$\mathbf{AC} = \left[\begin{array}{ccc} \uparrow & & \uparrow \\ \mathbf{Ac}_1 & \cdots & \mathbf{Ac}_n \\ \downarrow & & \downarrow \end{array} \right], \quad \mathbf{CD} = \left[\begin{array}{ccc} \uparrow & & \uparrow \\ \lambda_1 \mathbf{c}_1 & \cdots & \lambda_n \mathbf{c}_n \\ \downarrow & & \downarrow \end{array} \right]$$

and the Theorem follows by comparing \mathbf{AC} and \mathbf{CD} column by column. □

The Diagonalization Algorithm

The theorem suggests how we might find a matrix \mathbf{C} that diagonalizes \mathbf{A} :

- (i) Find n eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ of \mathbf{A} and use them as the columns of a matrix \mathbf{C} (in the same order)
- (ii) Use the eigenvalue of \mathbf{v}_i as the i^{th} entry of a diagonal matrix \mathbf{D}

We then have

$$\mathbf{C} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

and so, by the theorem,

$$\mathbf{AC} = \mathbf{CD}$$

If \mathbf{C} is invertible, we can then multiply both sides of this last equation by \mathbf{C}^{-1} to get

$$\mathbf{C}^{-1}\mathbf{AC} = \mathbf{D}$$

thus diagonalizing \mathbf{A} .

Example

Find a matrix **C** that diagonalizes $\mathbf{A} = \begin{bmatrix} 2 & 6 \\ 0 & -1 \end{bmatrix}$

Step 1: We first need to find the eigenvalues of **A**.

$$\begin{aligned} 0 &= p_{\mathbf{A}}(\lambda) = \det \begin{pmatrix} 2 - \lambda & 6 \\ 0 & -1 - \lambda \end{pmatrix} = (2 - \lambda)(-1 - \lambda) \\ \Rightarrow \quad \lambda &= 2, -1 \end{aligned}$$

E_2 : 2-eigenspace

Step 2: Find the eigenvectors of **A**

$$\begin{aligned} E_2 &= NullSp \left(\begin{bmatrix} 2-2 & 6 \\ 0 & -1-2 \end{bmatrix} \right) \\ &= NullSp \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ \Rightarrow x_2 &= 0, x_1 \text{ is free} \\ \Rightarrow \mathbf{x} &= \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \Rightarrow \mathbf{v}_{\lambda=2} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

So

$$E_2 = span \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$$

$E_{-1} : (-1)$ -eigenspace

$$\begin{aligned} E_{-1} &= \text{NullSp} \left(\begin{bmatrix} 2 - (-1) & 6 \\ 0 & -1 - (-1) \end{bmatrix} \right) \\ &= \text{NullSp} \left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \right) \\ \Rightarrow \quad x_1 &= -2x_2, \quad x_2 \text{ is free} \\ \Rightarrow \quad \mathbf{x} &= \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix} \\ \mathbf{v}_{\lambda=-1} &= \begin{bmatrix} -2 \\ 1 \end{bmatrix} \end{aligned}$$

Example, Cont'd

We thus have two linearly independent eigenvectors and so we can form an invertible matrix \mathbf{C} using $\mathbf{v}_{\lambda=2}$ and $\mathbf{v}_{\lambda=-1}$ as columns

$$\mathbf{C} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

Using the cofactor formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

we find

$$\mathbf{C}^{-1} = \frac{1}{1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Example, Cont'd

And so

$$\begin{aligned}\mathbf{C}^{-1}\mathbf{A}\mathbf{C} &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 4 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}\end{aligned}$$

Note that the diagonal entries of the diagonal matrix are just the eigenvalues of \mathbf{A} (the first eigenvalue 2 corresponding to the first eigenvector $\mathbf{v}_{\lambda=2}$ and the second eigenvalue -1 corresponding to the second eigenvector $\mathbf{v}_{\lambda=-1}$).

How to ensure that \mathbf{C} is invertible

We have just seen that we can diagonalize \mathbf{A} if we can form an **invertible** matrix \mathbf{C} using the eigenvectors of \mathbf{A} as columns. Recall an $n \times n$ matrix \mathbf{C} is invertible if and only if each of the following statements are true

- (i) \mathbf{C} is row reducible to the identity matrix
- (ii) $\mathbf{C}\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{R}^n$
- (iii) The only solution of $\mathbf{C}\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$
- (iv) $\text{rank}(\mathbf{C}) = n$
- (v) $\det(\mathbf{C}) \neq 0$

Since $\text{rank}(\mathbf{C}) = \dim(\text{ColSp}(\mathbf{C}))$, (iv) in turn requires that the columns of \mathbf{C} are linearly independent.

Thus,

Corollary

*An $n \times n$ matrix **A** is diagonalizable **if and only if** it has n linearly independent eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. In this situation, if we set*

$$\mathbf{C} = \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

with

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

*Then **C** is invertible and*

$$\mathbf{AC} = \mathbf{CD} \implies \mathbf{C}^{-1}\mathbf{AC} = \mathbf{D}$$

Lemma

Suppose \mathbf{v}_1 and \mathbf{v}_2 are eigenvectors of a matrix \mathbf{A} with different eigenvalues:

$$\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$$

$$\mathbf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2$$

with $\lambda_1 \neq \lambda_2$. Then $\{\mathbf{v}_1, \mathbf{v}_2\}$ are linearly independent.

Proof

Suppose $\{\mathbf{v}_1, \mathbf{v}_2\}$ are not linearly independent. Then there are non-zero numbers x_1 and x_2 such that

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 = \mathbf{0} \quad \Rightarrow \quad \mathbf{v}_1 = -\frac{x_2}{x_1} \mathbf{v}_2$$

But then

$$\begin{aligned} \lambda_1 \mathbf{v}_1 &= \mathbf{A} \mathbf{v}_1 \\ &= \mathbf{A} \left(-\frac{x_2}{x_1} \mathbf{v}_2 \right) \\ &= \left(-\frac{x_2}{x_1} \right) \mathbf{A} \mathbf{v}_2 \\ &= \left(-\frac{x_2}{x_1} \right) (\lambda_2 \mathbf{v}_2) \\ &= \lambda_2 \left(-\frac{x_2}{x_1} \mathbf{v}_2 \right) \\ &= \lambda_2 \mathbf{v}_1 \end{aligned}$$

More generally

which implies

$$(\lambda_1 - \lambda_2) \mathbf{v}_2 = \mathbf{0}$$

which can't happen since $\lambda_1 \neq \lambda_2$ and \mathbf{v}_2 is a non-zero vector.

Theorem

If $\mathbf{v}_1, \dots, \mathbf{v}_k$ are eigenvectors of a matrix \mathbf{A} with different eigenvalues, then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ are linearly independent.

So if an $n \times n$ matrix \mathbf{A} has n distinct eigenvalues, it will be diagonalizable.

Diagonalizability and Multiplicities of Eigenvalues

We have seen that an $n \times n$ matrix is diagonalizable if and only if it has n linearly independent eigenvectors.

I'll now show you another criterion for diagonalizability based on the notions of algebraic and geometric multiplicities.

Recall that the **algebraic multiplicity** of an eigenvalue r of an $n \times n$ matrix \mathbf{A} is the number of factors of $(\lambda - r)$ that occur in the characteristic polynomial $p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$ of \mathbf{A} . If

$$p_{\mathbf{A}}(\lambda) = a(\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$$

then

$$Mult_{alg}(r_i) = m_i$$

Note that since $\deg(p_{\mathbf{A}}(\lambda)) = n$, we always have

$$n = m_1 + m_2 + \cdots + m_k$$

Diagonalizability and Multiplicities of Eigenvalues, Cont'd

The **geometric multiplicity** μ of an eigenvalue r of \mathbf{A} is the dimension of the r -eigenspace of \mathbf{A} :

$$\mu = \text{Mult}_{\text{geom}}(r) = \dim(\text{NullSp}(\mathbf{A} - r\mathbf{I}))$$

The geometric multiplicities effectively counts the number of linearly independent eigenvectors in each eigenspace. If we sum over the geometric multiplicities μ_i of each eigenvalue r_i , we get the total number of linearly independent eigenvectors. Thus,

$$\# \text{ linearly independent eigenvectors} = \mu_1 + \mu_2 + \cdots + \mu_k$$

Thus,

Theorem

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

$$n = \mu_1 + \cdots + \mu_k \quad \Longleftrightarrow \quad \mathbf{A} \text{ is diagonalizable}$$

Numerics for Algebraic and Geometric Multiplicities

Theorem

Let r_1, \dots, r_k be the eigenvalues of an $n \times n$ matrix \mathbf{A} , let m_1, \dots, m_k be the corresponding algebraic multiplicities, and let μ_1, \dots, μ_k be the corresponding geometric multiplicities. Then

$$1 \leq \mu_i \leq m_i \text{ for } i = 1, \dots, k$$

Corollary

Let r_1, \dots, r_k be the eigenvalues of an $n \times n$ matrix \mathbf{A} . Then

\mathbf{A} is diagonalizable $\iff Mult_{geom}(r_i) = Mult_{alg}(r_i)$ for all i

Put another way, \mathbf{A} is **not diagonalizable** if there is an eigenvalue r_i such that

$$Mult_{geom}(r_i) < Mult_{alg}(r_i)$$

For if $m_i = Mult_{alg}(r_i)$ and $\mu_i = Mult_{geom}(r_i)$, then

$$\begin{aligned} \# \text{ lin. indep. eigenvectors} &= \mu_1 + \dots + \mu_k \\ &< m_1 + \dots + m_k = n \end{aligned}$$

Example

Determine if the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is diagonalizable.

We have

$$0 = p_{\mathbf{A}}(\lambda) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2$$

and so we have one eigenvalue $\lambda = 1$ with algebraic multiplicity 2.

$$E_1 = \text{NullSp} \begin{pmatrix} 1 - 1 & 1 \\ 0 & 1 - 1 \end{pmatrix} = \text{NullSp} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{span} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$$

Since E_1 is 1-dimensional, the geometric multiplicity of $\lambda = 1$ is 1.

Since

$$1 = \text{Geo. Mult.} (\lambda = 1) < \text{Alg. Mult.} (\lambda = 1) = 2$$

The matrix \mathbf{A} is **not diagonalizable**.

Other Criteria for Diagonalizability

Theorem

Let \mathbf{A} be an $n \times n$ matrix.

- (i) \mathbf{A} has n distinct eigenvalues, then \mathbf{A} is diagonalizable.
- (ii) If $\mu_i = m_i$ for each eigenvalue r_i , \mathbf{A} is diagonalizable.
- (iii) If \mathbf{A} has an eigenvalue r_i for which the corresponding geometric multiplicity μ_i is strictly less than the corresponding algebraic multiplicity m_i , then \mathbf{A} is not diagonalizable.

Other Tests for Diagonalizability, Cont'd

Lastly, we have the following theorem which is frequently applicable in physical problems.

Theorem

If \mathbf{A} is a symmetric matrix (i.e., $\mathbf{A}^t = \mathbf{A}$), then

- (i) All the eigenvalues of \mathbf{A} are real numbers.*
- (ii) \mathbf{A} is diagonalizable.*