Lecture 32: Diagonalization of Square Matrices, Cont'd

Math 3013
Oklahoma State University

April 18, 2022

Agenda

1. Diagonalizability
2. Diagonalizability Criteria
3. Examples

Diagonalizability

Definition

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

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

Two Matrix Multiplication Identities

Lemma

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

$$\mathbf{AC} = \begin{bmatrix} \uparrow & & \uparrow \\ (\mathbf{Ac}_1) & \cdots & (\mathbf{Ac}_n) \\ \downarrow & & \downarrow \end{bmatrix}$$

Lemma

Suppose $\mathbf{C}$ *is an* $n \times n$ *matrix with columns* $\mathbf{c}_1, \ldots, \mathbf{c}_n$ *and* $\mathbf{D}$ *is a diagonal* $n \times n$ *matrix with diagonal entries* $\lambda_1, \ldots, \lambda_n$*. Then*

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

Necessary Conditions 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* $\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{D}$*. Then*

(i) *Each column of* $\mathbf{C}$ *must be an eigenvector of* $\mathbf{A}$*.*

(ii) *For each column index* j *the* j^{th} *diagonal entry* λ_j *of* $\mathbf{D}$ *coincides with the eigenvalue of the eigenvector* $Col_j(\mathbf{C})$*.*

.

Proof:

$$\begin{aligned}
\mathbf{C}^{-1}\mathbf{A}\mathbf{C} &= \mathbf{D} \\
\Longrightarrow &\quad \mathbf{C}\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{C}\mathbf{D} \\
\Longrightarrow &\quad \mathbf{A}\mathbf{C} = \mathbf{C}\mathbf{D} \\
\Longrightarrow &\quad \begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{A}\mathbf{c}_1 & \cdots & \mathbf{A}\mathbf{c}_n \\ \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow \\ \lambda_1\mathbf{c}_1 & \lambda_n\mathbf{c}_n \\ \downarrow & \downarrow \end{bmatrix}
\end{aligned}$$

Constructing $\mathbf{C}$ and $\mathbf{D}$

.

The Theorem suggests we can solve the diagonalize an $n \times n$ matrix $\mathbf{A}$ by solving the Eigenvector/Eigenvalue Problem for $\mathbf{A}$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors of $\mathbf{A}$ and let $\lambda_1, \ldots, \lambda_n$ be the corresponding list of eigenvalues (so $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for each i). Set

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

Then the identity

$$\mathbf{AC} = \mathbf{DC}$$

is automatically satisfied.

Caveat

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

but

$$\mathbf{A}\mathbf{C} = \mathbf{D}\mathbf{C} \quad \text{does not imply} \quad \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{D}$$

because $\mathbf{C}^{-1}$ may not exist.

In today's lecture, we'll discuss additional conditions on the eigenvectors of $\mathbf{A}$ that will guarantee that we can construct an **invertible** matrix $\mathbf{C}$ from the eigenvectors of $\mathbf{A}$.

Invertibility Criteria

Lemma

If a matrix $\mathbf{C}$ *is invertible, then its column vectors must be linearly independent.*

Proof: Suppose $\mathbf{C}$ is invertible. Then

- $\mathbf{C}$ can be row reduced to the identity matrix $\mathbf{I}$.
- Since $\mathbf{I}$ is a R.E.F. of $\mathbf{C}$, and every column of $\mathbf{I}$ contains a pivot, every column of $\mathbf{C}$ is a basis vector for $ColSp(\mathbf{C})$.
- Since basis vectors are linearly independent, the columns of $\mathbf{C}$ must be linearly independent.

$\square$

Theorem

A $n \times n$ *matrix* $\mathbf{A}$ *is diagonalizable* **if and only if** *it has* n *linearly independent eigenvectors.*

So how do we know if a matrix $\mathbf{A}$ has n linearly independent eigenvectors?

Answer 1: Connection with Geometric Multiplicities

Recall

Definition

Suppose $r_1, \ldots, r_k$ are the eigenvalues of an $n \times n$ matrix $\mathbf{A}$.

- The **algebraic multiplicity** of the eigenvalue r_i is the number m_i of factors of $(\lambda - r)$ that occur in the characteristic polynomial

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = (\lambda - r_1)^{m_1} (\lambda - r_2)^{m_2} \cdots (\lambda - r_k)^{m_k}$$

- The **geometric multiplicity of** and eigenvalue r_i is the dimension of the r_i-eigenspace of $\mathbf{A}$:

$$\mu = \dim(E_{r_i}) = \dim(NullSp(\mathbf{A} - r_i \mathbf{I}))$$

Answer 1: Connection with Geometric Multiplicities, Cont'd

Theorem

Suppose $\mathbf{A}$ *is an* $n \times n$ *matrix with eigenvalues* $\lambda_1, \ldots, \lambda_k$*. Let* m_i *and* μ_i *denote, respectively, the corresponding algebraic and geometric multiplicities. Then*

(i) $1 \leq \mu_i \leq m_i \qquad i = 1, \ldots, k$

(ii) $\sum_{i=1}^{k} m_i = n$

Answer 1: Connection with Geometric Multiplicities, Cont'd

Corollary

Suppose $\mathbf{A}$ *is an* $n \times n$ *matrix with eigenvalues* $\lambda_1, \ldots, \lambda_k$. *Let* m_i *and* μ_i *denote, respectively, the corresponding algebraic and geometric multiplicities. Then if any* $\mu_i < m_i$, *then* $\mathbf{A}$ *is not diagonalizable.*

This is because the geometric multiplicities effectively count the number of linearly independent (basis) vectors in each eigenspace. So when we sum over the available eigenspaces, the total number of linearly independent eigenvectors we can obtain is

$$\sum_{i=1}^{k} \mu_i = \text{maximal number of linearly independent eigenvectors}$$

In of the preceding Theorem, if any one of the μ_i is less then the corresponding m_i, then the μ_i will fail to add up to n. Since an $n \times n$ matrix needs n linearly independent eigenvectors in order to be diagonalizable, the conclusion follows.

Answer 2: The Case of Distinct Eigenvalues

Theorem

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ *are eigenvectors of a matrix* $\mathbf{A}$ *whose eigenvalues are all distinct (* $\lambda_i \neq \lambda_j$ *if* $i \neq j$*). Then these eigenvectors are linearly independent.*

Proof of $k = 2$ case

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

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

which implies

$$\lambda_1\mathbf{v}_1 = \lambda_2\mathbf{v}_1$$

which can't happen since $\lambda_1 \neq \lambda_2$ and $\mathbf{v}_1$ is a non-zero (eigen-) vector. (For $k > 2$, one uses a mathematical induction argument to complete the proof).

2nd Criterion for Diagonalizability

Theorem

If an $n \times n$ matrix $\mathbf{A}$ has n distinct eigenvalues, then $\mathbf{A}$ is diagonalizable

Proof. If $\mathbf{A}$ has n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, it has n linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
And so an **invertible**, diagonalizing matrix, $\mathbf{C}$ can be formed

$$\mathbf{C} = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & \cdots & \downarrow \end{bmatrix}$$

Example 1

If possible find a matrix $\mathbf{C}$ that diagonalizes $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

We have

$$0 = p_{\mathbf{A}}(\lambda) = \det\left(\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\right) = (1-\lambda)^2 \quad \Longrightarrow \quad \lambda = 1$$

Since we have only one eigenvalue, diagonalizability is not guaranteed.

Example 1, Cont'd

Indeed, when we calculate the $\lambda = 1$ eigenspace,

$$\begin{aligned} E_1 &= NullSp\left(\mathbf{A} - (1)\mathbf{I}\right) = NullSp\left(\begin{bmatrix} 1-1 & 1 \\ 0 & 1-1 \end{bmatrix}\right) \\ &= NullSp\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = span\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \end{aligned}$$

we see **A** has only 1 linearly independent eigenvector.
But we need 2 linearly independent eigenvectors to construct an invertible matrix **C** that would diagonalize **A**.
We conclude that **A** is **not diagonalizable**.

Example 2.

If possible find a matrix $\mathbf{C}$ that diagonalizes $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

We have

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det \begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} \\ &= \lambda^3 - \lambda^2 - 2\lambda = -\lambda(\lambda+1)(\lambda-2) \end{aligned}$$

and so $\mathbf{A}$ has three distinct eigenvalues: $\lambda = 0, -1, 2$. Since each of its eigenvalues are all distinct, the matrix $\mathbf{A}$ is diagonalizable.

Example 2, Cont'd

Let's now go ahead and diagonalize $\mathbf{A}$. We need to find 3 linearly independent eigenvectors

$\lambda = 0$ Eigenspace

$$\begin{aligned} NullSp\left(\mathbf{A} - (0)\,\mathbf{I}\right) &= NullSp\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = NullSp\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ &= span\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right) \quad \Rightarrow \quad \mathbf{v}_{\lambda=0} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \end{aligned}$$

Example 2, Cont'd

$\lambda = -1$ Eigenspace

$$\begin{aligned} NullSp\left(\mathbf{A} - (-1)\,\mathbf{I}\right) &= NullSp\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = NullSp\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ &= span\left(\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) \Rightarrow \quad \mathbf{v}_{\lambda=-1} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

Example 2, Cont'd

$\lambda = 2$ Eigenspace

$$\begin{aligned}
NullSp\left(\mathbf{A} - (2)\,\mathbf{I}\right) &= NullSp\begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \\
&= NullSp\begin{pmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix} \\
&= span\left(\begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}\right) = span\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \\
\Rightarrow \quad & \mathbf{v}_{\lambda=2} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}
\end{aligned}$$

Example 2, Cont'd

We thus have found three linearly independent eigenvectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \text{ with eigenvalue } \lambda_1 = 0$$

$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \text{ with eigenvalue } \lambda_2 = -1$$

$$\mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \text{ with eigenvalue } \lambda_3 = 2$$

and so we can use the 3 linearly independent eigenvectors to form an invertible matrix $\mathbf{C}$ that will diagonalize $\mathbf{A}$:

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 1 & 1 \\ 0 & -1 & 2 \end{bmatrix} \quad , \quad \mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{D}$$

Diagonalization: The Special Case of Symmetric Matrices

Recall that an $n \times n$ matrix $\mathbf{A}$ is **symmetric** if $\mathbf{A} = \mathbf{A}^t$; or, equivalently,

$$a_{ij} = a_{ji} \qquad \text{for all } i, j$$

Symmetric matrices arise frequently in physical applications. For this reason, the following theorem is very important in applications:

Theorem

If $\mathbf{A}$ is a symmetric matrix then

- *$\mathbf{A}$ is diagonalizable*
- *All the eigenvalues of $\mathbf{A}$ are real numbers.*