21 Lecture 21: Diagonalization of $n \times n$ matrices

21.1 Recalls

21.1.1 Some remarks about Eigenvalues, eigenvectors, and eigenspace

- Eigenvalues of an $n \times n$ matrix $\mathbf{A}$ are roots of the characteristic polynomial

$$p_{\mathbf{A}}(\lambda) = |\mathbf{A} - \lambda \mathbf{I}_n|$$

- An integer $p \geq 1$ is called the multiplicity of an eigenvalue λ_1 if the characteristic polynomial of $\mathbf{A}$ contains the factor $(\lambda - \lambda_0)^p$ and p is highest number possible.

  For example, if the characteristic polynomial of a matrix $\mathbf{A}$ is factorized as $p_{\mathbf{A}}(\lambda) = (\lambda - 1)(\lambda - 3)^2$, then $\mathbf{A}$ has three eigenvalues, but only two distinct eigenvalues, which are $\lambda = 1$ (multiplicity 1) and $\lambda = 3$ (multiplicity 2).

- An $n \times n$ matrix $\mathbf{A}$ has exactly n eigenvalues counting complex eigenvalues and multiplicity.

- An eigenvector of an $n \times n$ matrix $\mathbf{A}$ associated with an eigenvalue λ is the solution of the homogeneous linear system
$$(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} = \mathbf{0}$$

- The eigenspace of an $n \times n$ matrix $\mathbf{A}$ associated with an eigenvalue λ is the null space of the matrix $\mathbf{A} - \lambda \mathbf{I}_n$.

- A set of linearly independe eigenvector of an $n \times n$ matrix $\mathbf{A}$ associated with an eigenvalue λ is a basis of the null space of the matrix $\mathbf{A} - \lambda \mathbf{I}_n$.

21.1.2 Definitions of a diagonalizable matrix and the diagonalization

An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n **linearly independent eigenvectors**. In the case $]A$ is diagonalizable, the diagonalization of $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1},$$

where columns of $\mathbf{S}$ are the n eigenvectors and the diagonal matrix $\mathbf{D}$ is formed eigenvalues of $\mathbf{A}$ that correspond to the eigenvector in $\mathbf{S}$. That is,

$$\mathbf{S} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

where $\mathbf{A}\mathbf{v}_\ell = \lambda_\ell \mathbf{v}_\ell, \quad \ell = 1, 2, \dots, n.$

21.2 Diagonalizable matrices and examples

Theorem 10. *Eigenvectors associated with different eigenvalues are linearly independent.*

21.2.1 Matrices whose all eigenvalues are distinct

Theorem 11. *An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.*

Example 21.1. *Diagonalize the matrix* $\mathbf{A}$ *below, if possible*

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

Solution: The characteristic polynomial

$$p_{\mathbf{A}}(\lambda) = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{vmatrix} = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

Thus, $\mathbf{A}$ has three distinct eigenvalues $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$. We now find the sets of independent eigenvectors for each λ_1, λ_2, and λ_3.

For $\lambda_1 = 1$,

$$\mathbf{A} - \lambda_1\mathbf{I} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{A} - \lambda_1\mathbf{I} & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which implies, x_1 is free, $x_2 = 0$, and $x_3 = 0$. The solution set in parametric vector form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \text{ Choose } \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

For $\lambda_2 = 2$,

$$\mathbf{A} - \lambda_2\mathbf{I} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{A} - \lambda_1\mathbf{I} & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which implies, $x_1 = x_2$, x_2 is free, and $x_3 = 0$. The solution set in parametric vector form

$$\mathbf{x} = \begin{bmatrix} x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}. \text{ Choose } \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

For $\lambda_3 = 3$,

$$\mathbf{A} - \lambda_3\mathbf{I} = \begin{bmatrix} -2 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -3/2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{A} - \lambda_1\mathbf{I} & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -3/2 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which implies, $x_1 = 3/2x_3$, $x_2 = 2x_3$, and x_3 is free. The solution set in parametric vector form

$$\mathbf{x} = \begin{bmatrix} 3/2x_3 \\ 2x_3 \\ x_3 \end{bmatrix} = \frac{x_3}{2}\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}. \text{ Choose } \mathbf{v}_3 = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}.$$

Matrix $\mathbf{A}$ is diagonalizable and a diagonalization of $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{SDS}^{-1},$$

where

$$\mathbf{S} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

21.2.2 Matrices whose eigenvalues are not distinct

Assume that an $n \times n$ matrix $\mathbf{A}$ has only p distinct eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_p$ with $p < n$. Then,

- For $1 \leq \ell \leq p$, the dimension of the eigenspace associated with the eigenvalue λ_ℓ is less than or equal to the multiplicity of the eigenvalue λ_ℓ.
- Matrix $\mathbf{A}$ is diagonalizable (counting complex eigenvalues and eigenvectors) if and only if the dimension of the eigenspace associated with the eigenvalue λ_ℓ is equal to the multiplicity of the eigenvalue λ_ℓ, for all $\ell = 1, 2, \ldots, p$.

Example 21.2. *Diagonalize the following matrix, if possible.*

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

Solution: The characteristic polynomial

$$p_{\mathbf{A}}(\lambda) = |\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = (\lambda - 1)^2.$$

Thus, $\mathbf{A}$ have an eigenvalue $\lambda = 1$ of multiplicity 2. We now find the set of independent eigenvectors associated with the eigenvalue $\lambda = 1$.

$$\mathbf{A} - \mathbf{I} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

Thus the solution of the homogeneous linear system $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$ is x_1 is free and $x_2 = 0$; or in the parametric vector form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The eigenspace have only one vector, thus it has dimension one, while the multiplicity of the eigenvalue is two. Therefore, matrix $\mathbf{A}$ is not diagonalizable.

Example 21.3. *Diagonalize the following matrix, if possible.*

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solution: The characteristic polynomial of $\mathbf{A}$ is

$$p_{\mathbf{A}}(\lambda) = |\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} 2-\lambda & 4 & 3 \\ -4 & -6-\lambda & 3 \\ 3 & 3 & 1-\lambda \end{vmatrix} = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda-1)(\lambda+2)^2.$$

Thus, eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = -2$ (multiplicity 2).

For $\lambda_1 = 1$, solve equation $(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{x} = \mathbf{0}$:

$$\mathbf{A} - \lambda_1 \mathbf{I} = \begin{bmatrix} 1 & 4 & 3 \\ -4 & -7 & -3 \\ 3 & 3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{A} - \lambda_1 \mathbf{I} & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, $x_1 = x_3$, $x_3 = -x_3$, and x_3 is free. So,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}. \text{ Choose } \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = -2$, solve equation $(\mathbf{A} - \lambda_2 \mathbf{I})\mathbf{x} = \mathbf{0}$:

$$\mathbf{A} - \lambda_2 \mathbf{I} = \begin{bmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{A} - \lambda_1 \mathbf{I} & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, $x_1 = -x_2$, x_2 is free, $x_3 = 0$. So,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

The eigenspace for $\lambda = -2$ has dimension 1, whose a basis is $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, while the multiplicity of the eigenvalue $\lambda = -2$ is 2. Thus, matrix $\mathbf{A}$ is not diagonalizable.

21.2.3 Application of diagonalization

We here consider an application of the diagonalization in evaluating $p(\mathbf{A})$, where

$$p(\mathbf{A}) = a_k\mathbf{A}^k + a_{k-1}\mathbf{A}^{k-1} + \ldots + a_1\mathbf{A} + a_0,$$

where $a_0, a_1, \ldots, a_k \in \mathbf{R}$.

Assume that an $n \times n$ matrix $\mathbf{A}$ is diogonalizable with the diagonalization

$$\mathbf{A} = \mathbf{SDS}^{-1},$$

where $\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$ a diagonal matrix. Then, for all $\ell = 1, 2, \ldots, k$:

- For any positive integer ℓ

  $$\mathbf{A}^\ell = (\mathbf{SDS}^{-1})(\mathbf{SDS}^{-1})\ldots(\mathbf{SDS}^{-1}) = \mathbf{SD}(\mathbf{S}^{-1}\mathbf{S})\mathbf{D}(\mathbf{S}^{-1}\ldots\mathbf{S})\mathbf{DS}^{-1} = \mathbf{SD}^p\mathbf{S}^{-1}.$$

  Note that,

  $$\mathbf{D}^\ell = \begin{bmatrix} \lambda_1^\ell & 0 & \ldots & 0 \\ 0 & \lambda_2^\ell & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \lambda_n^\ell \end{bmatrix}$$

  is a diagonal matrix.

  From the diagonalization of $\mathbf{A}$ and $\mathbf{A}^\ell$ we have that: if λ_ℓ is an eigenvalue of $\mathbf{A}$ associated with an eigenvector $\mathbf{v}_j$, then λ_j^ℓ is an eigenvalue of the matrix $\mathbf{A}^\ell$ associated with the same eigenvector $\mathbf{v}_j$.

- In addition, $a_\ell\mathbf{A}^\ell = a_\ell\mathbf{SD}^\ell\mathbf{S}^{-1} = \mathbf{S}(a_\ell\mathbf{D}^\ell)\mathbf{S}^{-1}$.

Thus,

$$p(\mathbf{A}) = \mathbf{S}\,p(\mathbf{D})\,\mathbf{S}^{-1} = \mathbf{S}\begin{bmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & p(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & p(\lambda_n) \end{bmatrix}\mathbf{S}^{-1}$$

21.2.4 More about 2×2 real matrix with complex eigenvalues

Assume that $\lambda = \alpha + i\beta$ is a complex eigenvalue of 2×2 real matrix $\mathbf{A}$, and $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ is an associated eigenvector. Then

- A have a diagonalization with complex-valued matrices as

$$\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}, \quad \text{where} \quad \mathbf{S} = \begin{bmatrix} \mathbf{v} & \overline{\mathbf{v}} \end{bmatrix}, \ \mathbf{D} = \begin{bmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{bmatrix}$$

- **A** can be factorized as a "diagonalization-like factorization" of real matrices as

$$\mathbf{A} = \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}^{-1}.$$

In addition, $\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} = r \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, where $(\alpha, -\beta) = r(\cos\theta, \sin\theta)$.