3 Spectral Theory V: Diagonalization

Prof.Dr.P.M.Bajracharya

School of Mathematical Sciences
T.U., Kirtipur

April 25, 2024

Summary

Diagonalization

Similar matrices

A square matrix B is **similar** to a matrix A if there exists an invertible matrix P such that $P^{-1}AP = B$.

We have proved in "3_Spectral theory II":

Theorem

Let A and B be similar matrices. Then the eigenvalues of these matrices are identical.

Diagonalizable matrix

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

Theorem

An $n \times n$ matrix A is **diagonalizable** iff it has n linearly independent eigenvectors.

Example

Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then A cannot be diagonalized.

Example

Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then A cannot be diagonalized.

Solution. A is diagonalizable implies A is similar to a diagonal matrix D with diagonal entries $\{d_1, d_2\} = \{0, 0\}$.

Example

Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then A cannot be diagonalized.

Solution. A is diagonalizable implies A is similar to a diagonal matrix D with diagonal entries $\{d_1, d_2\} = \{0, 0\}$. Hence

$$D = 0 \Rightarrow A = PDP^{-1} = 0.$$

This is a contradiction. $\square$

Example

Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Then A cannot be diagonalized.

Example

Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Then A cannot be diagonalized.

Solution. A is diagonalizable implies A is similar to a diagonal matrix D with diagonal entries $\{d_1, d_2, d_3\} = \{2, 2, 2\}$.

Example

Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Then A cannot be diagonalized.

Solution. A is diagonalizable implies A is similar to a diagonal matrix D with diagonal entries $\{d_1, d_2, d_3\} = \{2, 2, 2\}$. Hence

$$D = 2I \Rightarrow A = PDP^{-1} = 2I.$$

This is a contradiction. $\square$

Theorem

Let A be an $n \times n$ matrix. If P is an invertible matrix such that

$$PDP^{-1} = \operatorname{diag}(d_1, \dots, d_n),$$

then for $1 \leq i \leq n$, the i-th column of P is an eigenvector of A corresponding to d_i.

Proof. Let $P = [u_1, \ldots, u_n]$.

Proof. Let $P = [u_1, \ldots, u_n]$. Then

$$P^{-1}AP = D \Rightarrow AP = PD.$$

Proof. Let $P = [u_1, \ldots, u_n]$. Then
$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,
$$[Au_1, \ldots, Au_n] = A[u_1, \ldots, u_n]$$

Proof. Let $P = [u_1, \ldots, u_n]$. Then

$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,

$$\begin{aligned}[Au_1, \ldots, Au_n] &= A[u_1, \ldots, u_n] \\ &= AP = PD\end{aligned}$$

Proof. Let $P = [u_1, \ldots, u_n]$. Then

$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,

$$\begin{aligned}[Au_1, \ldots, Au_n] &= A[u_1, \ldots, u_n] \\ &= AP = PD \\ &= P\,\mathrm{diag}(d_1, \ldots, d_n)\end{aligned}$$

Proof. Let $P = [u_1, \ldots, u_n]$. Then

$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,

$$\begin{aligned}[Au_1, \ldots, Au_n] &= A[u_1, \ldots, u_n] \\ &= AP = PD \\ &= P \operatorname{diag}(d_1, \ldots, d_n) \\ &= [d_1u_1, \ldots, d_nu_n].\end{aligned}$$

Proof. Let $P = [u_1, \ldots, u_n]$. Then

$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,

$$\begin{aligned}[Au_1, \ldots, Au_n] &= A[u_1, \ldots, u_n] \\ &= AP = PD \\ &= P \operatorname{diag}(d_1, \ldots, d_n) \\ &= [d_1u_1, \ldots, d_nu_n].\end{aligned}$$

Or equivalently,

$$Au_i = d_iu_i, \text{ for} 1 \leq i \leq n.$$

Proof. Let $P = [u_1, \dots, u_n]$. Then
$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,
$$\begin{aligned}[Au_1, \dots, Au_n] &= A[u_1, \dots, u_n] \\ &= AP = PD \\ &= P\ \mathrm{diag}(d_1, \dots, d_n) \\ &= [d_1u_1, \dots, d_nu_n].\end{aligned}$$

Or equivalently,
$$Au_i = d_iu_i,\ \text{for} 1 \le i \le n.$$

As P is invertible, $u_1, \dots, u_n$ are linearly independent.

Proof. Let $P = [u_1, \ldots, u_n]$. Then

$$P^{-1}AP = D \Rightarrow AP = PD.$$

Thus,

$$\begin{aligned}[Au_1, \ldots, Au_n] &= A[u_1, \ldots, u_n] \\ &= AP = PD \\ &= P\,\mathrm{diag}(d_1, \ldots, d_n) \\ &= [d_1u_1, \ldots, d_nu_n].\end{aligned}$$

Or equivalently,

$$Au_i = d_iu_i, \ \text{for} 1 \le i \le n.$$

As P is invertible, $u_1, \ldots, u_n$ are linearly independent. Hence, (d_i, u_i), for $1 \le i \le n$, are eigen-pairs of A. This proves the theorem. $\square$

From the proof of the above theorem we obtain the following theorem

Theorem

Let A be an $n \times n$ matrix. If A is diagonalizable, then A has n linearly independent eigenvectors.

Its converse is also true.

Theorem

Let A be an $n \times n$ matrix. If A has n linearly independent eigenvectors, then A is diagonalizable.

Proof. Let $\{u_1, \ldots, u_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$.

Proof. Let $\{u_1, \ldots, u_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $V = [u_1, \ldots, u_n]$ is non-singular and

Proof. Let $\{u_1, \ldots, u_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $V = [u_1, \ldots, u_n]$ is non-singular and

$$AV = [Au_1, \ldots, Au_n]$$

Proof. Let $\{u_1, \ldots, u_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $V = [u_1, \ldots, u_n]$ is non-singular and

$$\begin{aligned} AV &= [Au_1, \ldots, Au_n] \\ &= [\lambda_1 u_1, \ldots, \lambda_n u_n] \end{aligned}$$

Proof. Let $\{u_1, \dots, u_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$. Then $V = [u_1, \dots, u_n]$ is non-singular and

$$\begin{aligned} AV &= [Au_1, \dots, Au_n] \\ &= [\lambda_1 u_1, \dots, \lambda_n u_n] \\ &= V\Lambda, \end{aligned}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Proof. Let $\{u_1, \dots, u_n\}$ be n linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$. Then $V = [u_1, \dots, u_n]$ is non-singular and

$$\begin{aligned} AV &= [Au_1, \dots, Au_n] \\ &= [\lambda_1 u_1, \dots, \lambda_n u_n] \\ &= V\Lambda, \end{aligned}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Therefore,

$$V^{-1}AV = \Lambda.$$

This implies that A is diagonalizable. $\square$

Example

Show that the following matrix is diagonalizable:

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

Solution: First, we shall find the eigenvalues of A.

Solution: First, we shall find the eigenvalues of A. The characteristic equation of A is given by:

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} = 0$$

Solution: First, we shall find the eigenvalues of A. The characteristic equation of A is given by:

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (1-\lambda)(2-\lambda)(3-\lambda) = 0$$

Solution: First, we shall find the eigenvalues of A. The characteristic equation of A is given by:

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (1-\lambda)(2-\lambda)(3-\lambda) = 0$$

$$\Rightarrow \lambda = 3, 2, 1.$$

$\underline{\underline{\lambda_1 = 3}}$

The eigenvector corresponding to $\lambda_1 = 3$ is the non-zero solution of the following matrix equation:

$$(A - 3I)X = 0$$

$\underline{\underline{\lambda_1 = 3}}$

The eigenvector corresponding to $\lambda_1 = 3$ is the non-zero solution of the following matrix equation:

$$(A - 3I)X = 0$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\underline{\underline{\lambda_1 = 3}}$

The eigenvector corresponding to $\lambda_1 = 3$ is the non-zero solution of the following matrix equation:

$$(A - 3I)X = 0$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -z \\ x + y + z \\ 2x + 2y + 2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This gives
$$\begin{cases} -z = 0 \\ x + y = 0 \end{cases}$$

This gives
$$\begin{cases} -z = 0 \\ x + y = 0 \end{cases}$$
If we take $x = 1 \Rightarrow y = -1$.

This gives
$$\begin{cases} -z = 0 \\ x + y = 0 \end{cases}$$
If we take $x = 1 \Rightarrow y = -1$.
Hence, the corresponding eigen-vector
$$u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Similarly, the eigenvector corresponding to $\lambda_2 = 2$ is given by:

$$\begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -x - z \\ x + z \\ 2x + 2y + z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Similarly, the eigenvector corresponding to $\lambda_2 = 2$ is given by:

$$\begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -x - z \\ x + z \\ 2x + 2y + z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} x + z & = 0 \\ 2x + 2y + z & = 0 \end{cases}.$$

Hence, the corresponding eigen-vector

$$u_1 = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$$

Hence, the corresponding eigen-vector

$$u_1 = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$$

Put

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

Now,

$$P^{-1}AP = \frac{1}{2}\begin{bmatrix} 0 & -2 & 1 \\ -2 & -2 & 0 \\ -2 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & -1 \\ 0 & 2 & -2 \end{bmatrix}.$$

Now,

$$P^{-1}AP = \frac{1}{2}\begin{bmatrix} 0 & -2 & 1 \\ -2 & -2 & 0 \\ -2 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & -1 \\ 0 & 2 & -2 \end{bmatrix}.$$

Then we see that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Now,

$$P^{-1}AP = \frac{1}{2}\begin{bmatrix} 0 & -2 & 1 \\ -2 & -2 & 0 \\ -2 & -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & -1 \\ 0 & 2 & -2 \end{bmatrix}.$$

Then we see that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Thus, A is diagonalizable.

Thus, from our work above, if there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A, then A is similar to a diagonal matrix D and so A is diagonalizable.

Thus, from our work above, if there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A, then A is similar to a diagonal matrix D and so A is diagonalizable.
On the other hand, if at least one of the eigenvalues of A is deficient, then A will not haven linearly independent eigenvectors.

Thus, from our work above, if there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A, then A is similar to a diagonal matrix D and so A is diagonalizable.
On the other hand, if at least one of the eigenvalues of A is deficient, then A will not haven linearly independent eigenvectors. Hence we will not be able to construct an invertible matrix P whose columns are eigenvectors of A. In this case, we say that A is not diagonalizable.