Eigenvalues and Eigenvectors
§5.2 Diagonalization

Satya Mandal, KU

Summer 2017
Suppose $A$ is square matrix of order $n$.

- Provide necessary and sufficient condition when there is an invertible matrix $P$ such that $P^{-1}AP$ is a diagonal matrix.
Definitions

- Two square matrices $A, B$ are said to be **similar**, if there is an invertible matrix $P$, such that $A = P^{-1}BP$.

- A square matrix $A$ said to be **diagonalizable**, if there is an invertible matrix $P$, such that $P^{-1}AP$ is a diagonal matrix. That means, if $A$ is similar to a diagonal matrix, we say that $A$ is **diagonalizable**.
Theorem 5.2.1

Suppose $A, B$ are two similar matrices. Then, $A$ and $B$ have same eigenvalues.

**Proof.** Write $A = P^{-1}BP$. Then

$$
|\lambda I - A| = |\lambda I - P^{-1}BP| = |\lambda(P^{-1}P) - P^{-1}BP| = |P^{-1}(\lambda I - BP)|
$$

$$
= |P^{-1}||\lambda I - B||P| = |P|^{-1}|\lambda I - B||P| = |\lambda I - B|
$$

So, $A$ and $B$ has same characteristic polynomials. So, they have same eigenvalues. The proof is complete.
Theorem 5.2.2: Diagonalizability

We ask, when a square matrix is diagonalizable?

Theorem 5.2.2 A square matrix $A$, of order $n$, is diagonalizable if and only if $A$ has $n$ linearly independent eigenvectors.

**Proof.** There are two statements to prove. First, suppose $A$ is diagonalizable.

Then $P^{-1}AP = D$, and hence $AP = PD$

where $P$ is an invertible matrix and $D$ is a diagonal matrix.

Write, $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$, $P = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \end{pmatrix}$
Since \( AP = PA \), we have

\[
A \begin{pmatrix} p_1 & p_2 & \cdots & p_n \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.
\]

Or

\[
\begin{pmatrix} Ap_1 & Ap_2 & \cdots & Ap_n \end{pmatrix} = \begin{pmatrix} \lambda_1 p_1 & \lambda_2 p_2 & \cdots & \lambda_n p_n \end{pmatrix}.
\]
So, 

\[ A p_i = \lambda_i p_i \quad \text{for} \quad i = 1, 2, \ldots, n \]

Since \( P \) is invertible, \( p_i \neq 0 \) and hence \( p_i \) is an eigenvector of \( A \), for \( \lambda \).

Also, \( \text{rank}(P) = n \). So, its columns \( \{p_1, p_2, \ldots, p_n\} \) are linearly independent.

So, it is established that if \( A \) is diagonalizable, then \( A \) has \( n \) linearly independent eigenvectors.
Now, we prove the converse. So, we assume \( A \) bas has \( n \) linearly independent eigenvectors:

\[
\{ p_1, p_2, \ldots, p_n \}
\]

So,

\[
A p_1 = \lambda_1 p_1, \quad A p_2 = \lambda_2 p_2, \quad \cdots, \quad A p_n = \lambda_n p_n \quad \text{for some} \quad \lambda_i.
\]
Write,

\[ P = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \end{pmatrix} \text{ and } D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \]

It follows from the equations \( A p_i = \lambda_i p_i \) that

\[ AP = PD. \text{ So, } P^{-1}AP = D \text{ is diagonal.} \]

The proof is complete.
Steps for Diagonalizing

Suppose $A$ is a square matrix of order $n$.

- If $A$ does not have $n$ linearly independent eigenvectors, then $A$ is not diagonalizable.
- When possible, find $n$ linearly independent eigenvectors $p_1, p_2, \ldots, p_n$ for $A$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.
- Then, write

$$P = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

- We have $D = P^{-1}AP$ is a diagonal matrix.
Corollary 4.4.3

Suppose $V$ is a vectors space and $x_1, x_2, \ldots, x_n$ be vectors in $V$. Then, $x_1, x_2, \ldots, x_n$ are linearly dependent if and only if there is an integer $m \leq n$ such that (1) $x_1, x_2, \ldots, x_m$ are linearly dependent and (2) $x_m \in \text{span}(x_1, x_2, \ldots, x_{m-1})$.

**Proof.** Suppose $x_1, x_2, \ldots, x_n$ are linearly dependent. By Theorem 4.4.2, one of these vectors is a linear combination of the rest. By relabeling, we can assume $x_n$ is a linear combination of $x_1, x_2, \ldots, x_{n-1}$. Let

$$m = \min\{ k : x_k \in \text{span}(x_1, x_2, \ldots, x_{k-1}) \}$$
If \( x_1, x_2, \ldots, x_{m-1} \) are linearly dependent, then we could apply Theorem 4.4.2 again, which would lead to a contradiction, that \( m \) is minimum. So, \( x_1, x_2, \ldots, x_{m-1} \) are linearly independent. This establishes one way implication. Conversely, suppose there is an \( m \leq n \) such that (1) and (2) holds. Then,

\[
x_m = c_1x_1 + \cdots + c_{m-1}x_{m-1}
\]

for some \( c_1, \ldots, c_{m-1} \in \mathbb{R} \). So,

\[
c_1x_1 + \cdots + c_{m-1}x_{m-1} + (-1)x_m = 0
\]

which is a nontrivial linear combination. So, \( x_1, x_2, \ldots, x_m, \ldots, x_n \) are linearly dependent.
Theorem 5.2.3: With Distinct Eigenvalues

Let $A$ be a square matrix $A$, of order $n$. Suppose $A$ has $n$ distinct eigenvalues. Then

- the corresponding eigenvectors are linearly independent
- and $A$ is diagonalizable.

Proof.

- The second statement follows from the first, by theorem 5.2.2. So, we prove the first statement only.
- Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct eigenvalues of $A$.
- So, for $i = 1, 2, \ldots, n$ we have

$$Ax_i = \lambda_i x_i \quad \text{where} \quad x_i \neq 0 \quad \text{are eigenvectors.}$$
We need to prove that $x_1, x_2, \ldots, x_n$ are linearly independent. We prove by contra-positive argument.

- So, assume they are linearly dependent.
- By Corollary 4.4.3 there is an $m < n$ such that $x_1, x_2, \ldots, x_m$ are mutually linearly independent and $x_{m+1}$ is in can be written as a linear combination of \{ $x_1, x_2, \ldots, x_m$ \}. So,

$$x_{m+1} = c_1 x_1 + c_2 x_2 + \cdots + c_m x_m \quad (1)$$

Here, at least one $c_i \neq 0$. Re-labeling $x_i$, if needed, we can assume $c_1 \neq 0$. 
Multiply (1) by $A$ on the left:

$$Ax_{m+1} = c_1Ax_1 + c_2Ax_2 + \cdots + c_mAxA_m$$  \hspace{1cm} (2)

Now, use $Ax_i = \lambda_i x_i$,

$$\lambda_{m+1}x_{m+1} = \lambda_1c_1x_1 + \lambda_2c_2x_2 + \cdots + \lambda_m c_mA_m$$  \hspace{1cm} (3)

Also, multiply (1) by $\lambda_{m+1}$, we have

$$\lambda_{m+1}x_{m+1} = \lambda_{m+1}c_1x_1 + \lambda_{m+1}c_2x_2 + \cdots + \lambda_{m+1}c_mA_m$$  \hspace{1cm} (4)
Subtract (3) from (4):

$$(\lambda_{m+1} - \lambda_1)c_1 x_1 + (\lambda_{m+1} - \lambda_2)c_2 x_2 + \cdots + (\lambda_{m+1} - \lambda_m)c_m x_m = 0.$$ 

Since these vectors are linearly independent, and hence

$$(\lambda_{m+1} - \lambda_i)c_i = 0 \quad \text{for} \quad i = 1, 2, \ldots, m.$$ 

Since $c_1 \neq 0$ we get $\lambda_{m+1} - \lambda_1 = 0$ or $\lambda_{m+1} = \lambda_1$. This contradicts that $\lambda_i$s are distinct. So, we conclude that $x_1, x_2, \ldots, x_n$ are linearly independent. The proof is complete.
Example 5.2.2

Let \( A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 3 \end{pmatrix} \) and \( P = \begin{pmatrix} 1 & 1 & 5 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \).

Verify that \( A \) is diagonalizable, by computing \( P^{-1}AP \).

**Solution:** We do it in a two steps.

1. Use TI to compute

\[
P^{-1} = \begin{pmatrix} 1 & 1 & -3 \\ 0 & -1 & 0.5 \\ 0 & 0 & 0.5 \end{pmatrix} .
\]

So, \( P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \).

So, it is verified that \( P^{-1}AP \) is a diagonal matrix.
Example 5.2.3

Let \( A = \begin{pmatrix} 3 & 1 \\ -9 & -3 \end{pmatrix} \).

Show that \( A \) is not diagonalizable.

**Solution:** Use Theorem 5.2.2 and show that \( A \) does not have 2 linearly independent eigenvectors. To do this, we have find and count the dimensions of all the eigenspaces \( E(\lambda) \). We do it in a few steps.

1. First, find all the eigenvalues. To do this, we solve

\[
\det(\lambda I - A) = \begin{vmatrix} \lambda - 3 & -1 \\ 9 & \lambda + 3 \end{vmatrix} = \lambda^2 = 0.
\]

So, \( \lambda = 0 \) is the only eigenvalue of \( A \).
Now we compute the eigenspace $E(0)$ of the eigenvalue $\lambda = 0$. We have $E(0)$ is solution space of

\[(0I - A) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ or } \begin{pmatrix} -3 & -1 \\ 9 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\]

Using TI (or by hand), a parametric solution of this system is given by $x = -0.5t$, $y = t$.

So, $E(0) = \{(t, -3t) : t \in \mathbb{R}\} = \mathbb{R}1, -3)$. So, the (sum of) dimension(s) of the eigenspace(s)

\[= \dim E(0) = 1 < 2.\]

Therefore $A$ is not diagonalizable.
Example 5.2.3

Let  
\[ A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{pmatrix}. \]

Show that \( A \) is not diagonalizable.

**Solution:** Use Theorem 5.2.2 and show that \( A \) does not have 3 linearly independent eigenvectors.

- To find the eigenvalues, we solve

\[
\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -1 & -1 \\ 0 & \lambda + 3 & -1 \\ 0 & 0 & \lambda + 3 \end{vmatrix} = (\lambda - 1)(\lambda + 3)^2 = 0.
\]

So, \( \lambda = 1, -3 \) are the only eigenvalues of \( A \).
We have $E(1)$ is solution space of

\[(I - A) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \]

Or

\[
\begin{pmatrix} 0 & -1 & -1 \\ 0 & 4 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\]

(As an alternative approach, avoid solving this system.)

The (column) rank of the coefficient matrix is 2. So,

\[\dim(E(1)) = \text{nullity} = 3 - \text{rank} = 3 - 2 = 1.\]
Now we compute the dimension dim $E(-3)$. $E(-3)$ is the solution space of

$$(-3I - A) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -4 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The rank of the coefficient matrix is 2 (use TI, if you need). So,

$$\dim(E(-3)) = \text{nullity} = 3 - \text{rank} = 3 - 2 = 1.$$
So, the sum of dimensions of the eigenspaces

\[ \dim E(1) + \dim E(-3) = 2 < 3. \]

Therefore $A$ is not diagonalizable.
Example 5.2.4

Let \( A = \begin{pmatrix} 17 & 113 & -2 \\ 0 & \sqrt{2} & 1 \\ 0 & 0 & \pi \end{pmatrix} \) Find its eigenvalues and determine (use Theorem 5.2.3), if \( A \) is diagonalizable. If yes, write down a an invertible matrix \( P \) so that \( P^{-1}AP \) is a diagonal matrix.

**Solution:** To find eigenvalues solve

\[
\det(\lambda I - A) = \begin{vmatrix} \lambda - 17 & -113 & 2 \\ 0 & \lambda - \sqrt{2} & -1 \\ 0 & 0 & \lambda - \pi \end{vmatrix} = (\lambda - 17)(\lambda - \sqrt{2})(\lambda - \pi) = 0.
\]
So, $A$ has three distinct eigenvalues $\lambda = 17, \sqrt{2}, \pi$. Since $A$ is a $3 \times 3$ matrix, by Theorem 5.2.3, $A$ is diagonalizable. We will proceed to compute the matrix $P$, by computing bases of $E(17)$, $E(\sqrt{2})$ and $E(\pi)$. 

\[\text{\hfill }\]
To compute $E(17)$, we solve: $(17I_3 - A)x = 0$, which is

$$
\begin{pmatrix}
0 & -113 & 2 \\
0 & 17 - \sqrt{2} & -1 \\
0 & 0 & 17 - \pi
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
$$

So, $z = y = 0$ and $x = t$, for any $t \in \mathbb{R}$. So,

$$E(17) = \left\{ \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix} : t \in \mathbb{R} \right\}$$
with $t = 1$ a basis of $E(17)$ is \[
\begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}
\]
To compute $E(\sqrt{2})$, we solve: $(\sqrt{2}I_3 - A)x = 0$, which is

$$
\begin{pmatrix}
\sqrt{2} - 17 & -113 & 2 \\
0 & 0 & -1 \\
0 & 0 & \sqrt{2} - \pi
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix} =
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
$$

So, $z = 0$ and $x = t$ and $y = \frac{\sqrt{2} - 17}{113}t$ for any $t \in \mathbb{R}$. So,

$$E(\sqrt{2}) = \left\{ \begin{pmatrix} t \\ \frac{\sqrt{2} - 17}{113}t \\ 0 \end{pmatrix} : t \in \mathbb{R} \right\}$$
with $t = 113$ a basis of $E(\sqrt{2})$ is

$$\left\{ \begin{pmatrix} 113 \\ \sqrt{2} - 17 \\ 0 \end{pmatrix} \right\}$$
To compute $E(\pi)$, we solve: $(\pi I_3 - A)x = 0$, which is

$$
\begin{pmatrix}
\pi - 17 & -113 & 2 \\
0 & \pi - \sqrt{2} & -1 \\
0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
$$

$$
\begin{cases}
z = t \\
y = \frac{1}{\pi - \sqrt{2}}z = \frac{1}{\pi - \sqrt{2}}t \\
x = \frac{113}{\pi - 17}y - \frac{2}{\pi - 17}z = \frac{113 + 2\sqrt{2} - 2\pi}{(\pi - 17)(\pi - \sqrt{2})}t
\end{cases}
$$

$$
E(\pi) = \left\{ \begin{pmatrix}
\frac{113 + 2\sqrt{2} - 2\pi}{(\pi - 17)(\pi - \sqrt{2})}t \\
\frac{1}{\pi - \sqrt{2}}t \\
t
\end{pmatrix} : t \in \mathbb{R} \right\}
$$
With $t = (\pi - 17)(\pi - \sqrt{2})$ a basis of $E(\pi)$ is

\[
\begin{pmatrix}
113 + 2\sqrt{2} - 2\pi \\
\pi - 17 \\
(\pi - 17)(\pi - \sqrt{2})
\end{pmatrix}
\]
We form the matrix of the eigenvectors.

\[
P = \begin{pmatrix}
1 & 113 & 113 + 2\sqrt{2} - 2\pi \\
0 & \sqrt{2} - 17 & \pi - 17 \\
0 & 0 & (\pi - 17)(\pi - \sqrt{2})
\end{pmatrix}.
\]

We check

\[
P^{-1}AP = \begin{pmatrix}
17 & 0 & 0 \\
0 & \sqrt{2} & 0 \\
0 & 0 & \pi
\end{pmatrix}
\]

Or

\[
AP = P \begin{pmatrix}
17 & 0 & 0 \\
0 & \sqrt{2} & 0 \\
0 & 0 & \pi
\end{pmatrix}
\]
Continued

We have

\[ AP = \begin{pmatrix} 17 & 113 & -2 \\ 0 & \sqrt{2} & 1 \\ 0 & 0 & \pi \end{pmatrix} \begin{pmatrix} 1 & 113 & 113 + 2\sqrt{2} - 2\pi \\ 0 & \sqrt{2} - 17 & \pi - 17 \\ 0 & 0 & (\pi - 17)(\pi - \sqrt{2}) \end{pmatrix} \]

\[ = \begin{pmatrix} 17 & 113\sqrt{2} & \pi(113 + 2\sqrt{2} - 2\pi) \\ 0 & \sqrt{2}(\sqrt{2} - 17) & \pi(\pi - 17) \\ 0 & 0 & \pi(\pi - 17)(\pi - \sqrt{2}) \end{pmatrix} \]

\[ = P \begin{pmatrix} 17 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \pi \end{pmatrix} \]