Section 5.1: Eigenvalues and Eigenvectors

Definition. Let A be a matrix (or linear transformation). A number λ is called an eigenvalue of A if there exists a non-zero vector $\vec{u}$ such that

$$A(\vec{u}) = \lambda\vec{u}, \text{ or } A\vec{u} - \lambda\vec{u} = \vec{0} \,. \tag{1}$$

The vector $\vec{u}$ is called an eigenvector associated with this eigenvalue λ. The set of all eigenvectors associated with λ forms a subspace, and is called the eigenspace associated with λ.

Remarks. (1) $\vec{u} \neq \vec{0}$ is crucial, since $\vec{u} = \vec{0}$ always satisfies Equ (1). (2) If $\vec{u}$ is an eigenvector for λ, then so is $c\vec{u}$ for any constant c. (3) Geometrically, in 3D, eigenvectors of A are directions that are unchanged under transformation A.

We observe from Equ (1) that λ is an eigenvalue iff Equ (1) has a non-trivial solution, i.e.,

$$(A - \lambda I)\vec{u} = A\vec{u} - \lambda\vec{u} = \vec{0} \tag{2}$$

has a non-trivial solution. Therefire

$$eigenspace = Null\,(A - \lambda I)\,.$$

By the inverse matrix theorem, we conclude that Equ (2) has a non-trivial solution iff

$$\det(A - \lambda I) = 0. \tag{3}$$

Equ (3) is often referred as to "Characteristic Equation".

Finding eigenvalues and eigenvectors: Step #1. Solve Characteristic Equ (3) for λ. Step #2. For each λ, find a basis for the eigenspace $Null\,(A - \lambda I)$ (i.e., solution set of Equ (2)).

Example 5.1.1. Find all eigenvalues and their eigenspaces for

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

Solution:

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

Write characteristic equation

$$\det(A - \lambda I) = (3 - \lambda)(-\lambda) - (-2) = 0.$$

We find

$$\lambda^2 - 3\lambda + 2 = 0 \quad \Longrightarrow \quad (\lambda - 1)(\lambda - 2) = 0 \quad \Longrightarrow \lambda = 1, \; \lambda = 2.$$

There are two eigenvalues $\lambda_1 = 1,\ \lambda_2 = 2$. We next find eigenvectors associated with each eigenvalue. For $\lambda_1 = 1$,

$$\vec{0} = (A - \lambda_1 I)\,\vec{x} = \begin{bmatrix} 3-1 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

or

$$x_1 = x_2.$$

The parametric vector form of solution set for $(A - \lambda_1 I)\,\vec{x} = \vec{0}$:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ basis: } \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

This is only (linearly independent) eigenvector for $\lambda_1 = 1$.

The last step can be done slightly differently as follows. From solutions (for $(A - \lambda_1 I)\,\vec{x} = \vec{0}$)

$$x_1 = x_2,$$

we know there is only one free variable x_2. Therefore, there is only one generator in any basis. To find it, we take x_2 to be any nonzero number, for instance, $x_2 = 1$, and compute $x_1 = x_2 = 1$. We obtain

$$\lambda_1 = 1, \quad \vec{u}_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = 2$,

$$\vec{0} = (A - \lambda_2 I)\,\vec{x} = \begin{bmatrix} 3-2 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

or

$$x_1 = 2x_2.$$

To find a basis, we take $x_2 = 1$, we have $x_1 = 2$, and

$$\lambda_2 = 2, \quad \vec{u}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Example 5.1.2. Given that 2 is an eigenvalue for

$$A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}.$$

Find a basis of its eigenspace.

Solution:

$$A - 2I = \begin{bmatrix} 4-2 & -1 & 6 \\ 2 & 1-2 & 6 \\ 2 & -1 & 8-2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, $(A-2I)\vec{x}=\vec{0}$ becomes

$$2x_1 - x_2 + 6x_3 = 0, \text{ or } x_2 = 2x_1 + 6x_3, \tag{4}$$

where we select x_1 and x_3 as free variables only to avoid fractions. Solutions in parametric form:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 + 6x_3 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix}.$$

A basis for the eigenspace:

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \text{ and } \vec{u}_2 = \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix}.$$

Another way of solving Equ (4) may be a little easy. From Equ (4), we know that x_1 an x_3 are free variables. We choose $(x_1,\ x_3) = (1,0)$ and $(0,1)$, respectively,

$$\begin{aligned} x_1 = 1,\ x_3 = 0 &\Longrightarrow x_2 = 2 \Longrightarrow \vec{u}_1 \\ x_1 = 0,\ x_3 = 1 &\Longrightarrow x_2 = 6 \Longrightarrow \vec{u}_2. \end{aligned}$$

Example 5.1.3. Find eigenvalues: (a)

$$A = \begin{bmatrix} 3 & -1 & 6 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}, \quad A - \lambda I = \begin{bmatrix} 3-\lambda & -1 & 6 \\ 0 & -\lambda & 6 \\ 0 & 0 & 2-\lambda \end{bmatrix}.$$

$$\det(A - \lambda I) = (3-\lambda)(-\lambda)(2-\lambda) = 0$$

The eigenvalues are 3, 0, 2, exactly the diagonal elements. (b)

$$B = \begin{bmatrix} 4 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 4 \end{bmatrix}, \quad B - \lambda I = \begin{bmatrix} 4-\lambda & 0 & 0 \\ 2 & 1-\lambda & 0 \\ 1 & 0 & 4-\lambda \end{bmatrix}$$

$$\det(B - \lambda I) = (4-\lambda)^2 (1-\lambda) = 0.$$

The eigenvalues are 4, 1, 4 (4 is a double root), exactly the diagonal elements.

Theorem. (1) The eigenvalues of a triangle matrix are its diagonal elements.

(2) Eigenvectors for different eigenvalues are linearly independent. More precisely, suppose that λ_1, λ_2, ..., λ_p are p different eigenvalues of a matrix A. Then, a set of

$$\text{a basis of } Null(A - \lambda_1 I),\ \text{a basis of } Null(A - \lambda_2 I), \ldots,\ \text{a basis of } Null(A - \lambda_p I)$$

is linearly independent.

Proof. (2) For simplicity, we assume $p = 2 : \lambda_1 \neq \lambda_2$ are two different eigenvalues. Suppose that $\vec{u}_1$ is an eigenvector of λ_1 and $\vec{u}_2$ is an eigenvector of λ_2 To show independence, we need to show that the only solution to

$$x_1\vec{u}_1 + x_2\vec{u}_2 = \vec{0}$$

is $x_1 = x_2 = 0$. Indeed, if $x_1 \neq 0$, then

$$\vec{u}_1 = \frac{x_2}{x_1}\vec{u}_2. \tag{5}$$

We now apply A to the above equation. It leads to

$$A\vec{u}_1 = \frac{x_2}{x_1}A\vec{u}_2 \Longrightarrow \lambda_1\vec{u}_1 = \frac{x_2}{x_1}\lambda_2\vec{u}_2. \tag{6}$$

Equ (5) and Equ (6) are contradictory to each other: by Equ (5),

$$\begin{aligned}\text{Equ (5)} &\Longrightarrow \lambda_1\vec{u}_1 = \frac{x_2}{x_1}\lambda_1\vec{u}_2 \\ \text{Equ (6)} &\Longrightarrow \lambda_1\vec{u}_1 = \frac{x_2}{x_1}\lambda_2\vec{u}_2,\end{aligned}$$

or

$$\frac{x_2}{x_1}\lambda_1\vec{u}_2 = \lambda_1\vec{u}_1 = \frac{x_2}{x_1}\lambda_2\vec{u}_2.$$

Therefor $\lambda_1 = \lambda_2$, a contradiction to the assumption that they are different eigenvalues. ■

Section 5.2: Characteristic Equations

As we discussed in the previous section, the key to find eigenvalues and eigenvectors is to solve the Characteristic Equation (3)

$$\det(A - \lambda I) = 0.$$

For 2×2 matrix,

$$\begin{aligned}A - \lambda I &= \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix}, \\ \det(A - \lambda I) &= (a-\lambda)(d-\lambda) - bc \\ &= \lambda^2 + (-a-d)\lambda + (ad - bc)\end{aligned}$$

is a guadratic function (i.e., a polynomial of degree 2). In general, for any $n \times n$ matrix A,

$$A - \lambda I = \begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix},$$

$$\det(A-\lambda I) = (a_{11}-\lambda)\det\begin{bmatrix} a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix} + \ldots$$

Therefore, $\det(A-\lambda I)$ is a polynomial of degree n, and is often called the **characteristic polynomial** of A. Consequently, there are total of n (the number of rows in the matrix A) eigenvalues (real or complex, after taking account for multiplicity). Finding roots for higher order polynomials may be very difficult.

Example 5.2.1. Find all eigenvalues for

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Solution:

$$A-\lambda I = \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 1 & 1-\lambda \end{bmatrix},$$

$$\begin{aligned} \det(A-\lambda I) &= (5-\lambda)\det\begin{bmatrix} 3-\lambda & -8 & 0 \\ 0 & 5-\lambda & 4 \\ 0 & 1 & 1-\lambda \end{bmatrix} \\ &= (5-\lambda)(3-\lambda)\det\begin{bmatrix} 5-\lambda & 4 \\ 1 & 1-\lambda \end{bmatrix} \\ &= (5-\lambda)(3-\lambda)\left[(5-\lambda)(1-\lambda)-4\right] = 0. \end{aligned}$$

There are 4 roots:

$$\begin{aligned} (5-\lambda) = 0 &\implies \lambda = 5 \\ (3-\lambda) = 0 &\implies \lambda = 3 \\ (5-\lambda)(1-\lambda)-4 = 0 &\implies \lambda^2 - 6\lambda + 1 = 0 \\ &\implies \lambda = \frac{6 \pm \sqrt{36-4}}{2} = 3 \pm 2\sqrt{2}. \end{aligned}$$

Question: Suppose that B is obtained from A by elementary row operations. Do A and B has the same eigenvalues? (Ans: No)

Example 5.2.2.

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \overset{R_2+R_1\to R_2}{\longrightarrow} \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} = B$$

A has eigenvalues 1 and 2. For B, the characteristic equation is

$$\begin{aligned} \det(B-\lambda I) &= \begin{bmatrix} 1-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix} \\ &= (1-\lambda)(3-\lambda) - 1 = \lambda^2 - 4\lambda + 2. \end{aligned}$$

The eigenvalues are

$$\lambda = \frac{4 \pm \sqrt{16-8}}{2} = \frac{4 \pm \sqrt{8}}{2} = 2 \pm \sqrt{2}.$$

Definition. Two $n \times n$ matrices A and B are called similar, and is denoted as $A \sim B$, if there exists an invertible matrix P such that $A = PBP^{-1}$.

Claim. If A and B are similar, then they have exact the same characteristic polynomial and consequently the same eigenvalues.

Indeed, if $A = PBP^{-1}$, then $P(B - \lambda I)P^{-1} = PBP^{-1} - \lambda PIP^{-1} = (A - \lambda I)$. Therefore,

$$\det(A - \lambda I) = \det\left(P(B - \lambda I)P^{-1}\right) = \det(P)\det(B - \lambda I)\det\left(P^{-1}\right) = \det(B - \lambda I).$$

Caution: If $A \sim B$, and if λ_0 is an eigenvalue, then an corresponding eigenvector for A may not be an eigenvector for B. In other words, A and B have the same eigenvalues but different eigenvectors.

Example 5.2.3. Though row operation alone will not perserve eigenvalues, a pair of row and column operation do maintain similarity. We first observe that if P is a type 1 (row) elementary matrix,

$$P = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \overset{R_1 + aR_2 \to R_2}{\longleftarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

then its inverse P^{-1} is a type 1 (column) elementary matrix obtained from the identity matrix by an elementary column operation that is of the same kind with "opposite sign" to the previous row operation, i.e.,

$$P^{-1} = \begin{bmatrix} 1 & 0 \\ -a & 1 \end{bmatrix} \overset{C_1 - aC_2 \to C_1}{\longleftarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We call the column operation

$$C_1 - aC_2 \to C_1$$

is "inverse" to the row operation

$$R_1 + aR_2 \to R_2.$$

Now we perform a row operation on A followed immediately by the column operation inverse to the row operation

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \overset{R_1 + R_2 \to R_2}{\to} \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \text{ (left multiply by } P\text{)}$$

$$\overset{C_1 - C_2 \to C_1}{\to} \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} = B \text{ (right multiply by } P^{-1}.\text{)}$$

We can verify that A and B are similar through P (with $a = 1$)

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

Now, $\lambda_1 = 1$ is an eigenvalue. Then,

$$\begin{aligned}(A-1)\,\vec{u} &= \begin{bmatrix}1-1 & 1\\ 0 & 2-1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}\\ &= \begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}\\ &\Longrightarrow \vec{u} = \begin{bmatrix}1\\0\end{bmatrix} \text{ is an eigenvector for } A.\end{aligned}$$

But

$$\begin{aligned}(B-1)\,\vec{u} &= \begin{bmatrix}0-1 & 1\\ -2 & 3-1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}\\ &= \begin{bmatrix}-1 & 1\\ -2 & 2\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}-1\\-2\end{bmatrix}\\ &\Longrightarrow \vec{u} = \begin{bmatrix}1\\0\end{bmatrix} \text{ is NOT an eigenvector for } B.\end{aligned}$$

In fact,

$$\begin{aligned}(B-1)\,\vec{v} &= \begin{bmatrix}-1 & 1\\ -2 & 2\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.\\ \text{So, } \vec{v} &= \begin{bmatrix}1\\-1\end{bmatrix} \text{ is an eigenvector for } B.\end{aligned}$$

Example 5.2.4. Find eigenvalues of A if

$$A \sim B = \begin{bmatrix}5 & -2 & 6 & -1\\ 0 & 3 & -8 & 0\\ 0 & 0 & 5 & 4\\ 0 & 0 & 0 & 4\end{bmatrix}.$$

Solution: Eigenvalues of B are $\lambda = 5, 3, 5, 4$. These are also the eigenvalues of A.

Section 5.3: Diagonalization

Diagonal matrix: only diagonal entries are non-zero

$$D = \begin{bmatrix}a_1 & 0 & \cdots & 0\\ 0 & a_2 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & a_n\end{bmatrix}. \tag{7}$$

Obviously,

$$D\vec{e}_1 = \begin{bmatrix}a_1 & 0 & \cdots & 0\\ 0 & a_2 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & a_n\end{bmatrix}\begin{bmatrix}1\\0\\\cdots\\0\end{bmatrix} = \begin{bmatrix}a_1\\0\\\cdots\\0\end{bmatrix} = a_1\vec{e}_1.$$

$$D\vec{e}_1 = a_1\vec{e}_1, \quad \text{i.e., } \vec{e}_1 \text{ is an eigenvector associated with } a_1.$$

In general,

$$D\vec{e}_i = a_i\vec{e}_i.$$

For any diagonal matrix, eigenvalues are all diagonal entries, and $\vec{e}_i$ is an eigenvector associated with a_i (ith entry). For the purpose of calculating eigenvalues and eigenvectors, diagonal matrices are easiest. Diagonalization is a process to find a diagonal matrix that is similar to a given non-diagonal matrix.

Example 5.3.1. Consider

$$A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}.$$

(a) Verify $A = PDP^{-1}$; (b) Find D^k and A^k; (c) Find eigenvalues and eigenvectors for A.

Solution: (a) It suffices to show that $AP = PD$ and that P is invertible. Direct calculations lead to

$$\det P = -1 \neq 0 \Longrightarrow P \text{ is invertible}$$

$$AP = \begin{bmatrix} 5 & 3 \\ -5 & -6 \end{bmatrix}, \quad PD = \begin{bmatrix} 5 & 3 \\ -5 & -6 \end{bmatrix}.$$

(b)

$$D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}, \quad D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}.$$

$$A^2 = PDP^{-1}\left(PDP^{-1}\right) = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$
$$A^k = PD^kP^{-1}$$

(d) Eigenvalues of A = Eigenvalues of $D: \lambda_1 = 5$, $\lambda_2 = 3$. For D,

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ is an eigenvectors for} \lambda_1 = 5: \; D\vec{e}_1 = \lambda_1\vec{e}_1 \tag{8}$$

$$\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ is an eigenvectors for} \lambda_2 = 3: \; D\vec{e}_2 = \lambda_2\vec{e}_2. \tag{9}$$

Since $AP = PD$, from (7), we see $AP\vec{e}_i = PD\vec{e}_i = a_iP\vec{e}_i$. In particular,

$$A\left(P\vec{e}_1\right) = PD\vec{e}_1 \overset{\text{by (8)}}{=} P\left(\lambda_1\vec{e}_1\right) = \lambda_1\left(P\vec{e}_1\right) \Longrightarrow P\vec{e}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ is an eigenvector}$$

$$A\left(P\vec{e}_2\right) = PD\vec{e}_2 \overset{\text{by (9)}}{=} P\left(\lambda_2\vec{e}_2\right) = \lambda_2\left(P\vec{e}_2\right) \Longrightarrow P\vec{e}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \text{ is an eigenvector.}$$

Conclusion: In general, if D is a diagonal matrix with diagonal entries $a_1, a_2, ..., a_n$ (see (7)) and if $AP = PD$, then

$$P\vec{e}_i \text{ is an eigenvector associated with } a_i.$$

Definition. An $n \times n$ matrix A is called diagonalizable if A is similar to a diagonal matrix D.

Theorem (Diagonalization). Let A be an $n \times n$ matrix. Suppose that A has n linearly independent eigenvectors $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$. Then, A is diagonalizable and $AP = PD$, where

$$P = [\vec{v}_1, \vec{v}_2, ..., \vec{v}_n], \ D = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} = [a_1\vec{e}_1, a_2\vec{e}_2, ..., a_n\vec{e}_n]$$

a_i is the eigenvalue associated with $\vec{v}_i$, i.e., $A\vec{v}_i = a_i\vec{v}_i$.

Proof. We only need to verify that $AP = PD$ as follows:

$$AP = A[\vec{v}_1, \vec{v}_2, ..., \vec{v}_n] = [A\vec{v}_1, A\vec{v}_2, ..., A\vec{v}_n] = [a_1\vec{v}_1, a_2\vec{v}_2, ..., a_n\vec{v}_n]$$
$$PD = P[a_1\vec{e}_1, a_2\vec{e}_2, ..., a_n\vec{e}_n] = [a_1P\vec{e}_1, a_2P\vec{e}_2, ..., a_nP\vec{e}_n].$$

Now,

$$P\vec{e}_1 = [\vec{v}_1, \vec{v}_2, ..., \vec{v}_n]\begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} = \vec{v}_1, \quad P\vec{e}_2 = [\vec{v}_1, \vec{v}_2, ..., \vec{v}_n]\begin{bmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{bmatrix} = \vec{v}_2, ...$$

This shows $AP = PD$. ■

Example 5.3.2. Diagonalize

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

Solution: Step 1. Find all eigenvalues.

$$\det(A - \lambda I) = \det\begin{bmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} = -\lambda^3 - 3\lambda^2 + 4$$
$$= -\left(\lambda^3 + 3\lambda^2 - 4\right) = -\left[\lambda^3 - \lambda^2 + 4\lambda^2 - 4\right]$$
$$= -\left[\lambda^2(\lambda - 1) + 4\left(\lambda^2 - 1\right)\right] = -\left[\lambda^2(\lambda - 1) + 4(\lambda + 1)(\lambda - 1)\right]$$
$$= -(\lambda - 1)\left[\lambda^2 + 4(\lambda + 1)\right] = -(\lambda - 1)\left[\lambda^2 + 4\lambda + 4\right] = -(\lambda - 1)(\lambda + 2)^2.$$

Eigenvalues are $\lambda_1 = 1$, $\lambda_2 = -2$ (this is a double root).

Step 2. Find all eigenvalues – find a basis for each eigenspace $Null(A - \lambda I_i)$. For $\lambda_1 = 1$,

$$A - \lambda_1 I = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \overset{R_1 \to R_2}{\longrightarrow} \begin{bmatrix} 3 & 3 & 0 \\ -3 & -6 & -3 \\ 0 & 3 & 3 \end{bmatrix} \overset{R_2 + R_1 \to R_2}{\longrightarrow} \begin{bmatrix} 3 & 3 & 0 \\ 0 & -3 & -3 \\ 0 & 3 & 3 \end{bmatrix}$$
$$\to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow x_1 = x_3, \ x_2 = -x_3$$

x_3 is the free variable. Choose $x_3 = 1$, we obtain an eigenvector

$$\vec{x} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = -2$,

$$A - \lambda_2 I = \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow x_1 + x_2 + x_3 = 0.$$

It follows that x_2 and x_3 are free variables. As we did before, we need to select (x_2, x_3) to be $(1, 0)$ and $(0, 1)$. Choose $x_2 = 1$, $x_3 = 0 \Longrightarrow x_1 = -x_2 - x_3 = -1$; choose $x_2 = 0$, $x_3 = 1 \Longrightarrow x_1 = -x_2 - x_3 = -1$. We thus got two independent eigenvectors for $\lambda_2 = -2$:

$$\vec{v}_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Step 3. Assemble orderly D and P as follows: there are several choices to pair D and P.

$$\begin{aligned}
&Choice\#1: \; D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad P = [\vec{v}_1, \vec{v}_2, \vec{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\
&Choice\#2: \; D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad P = [\vec{v}_1, \vec{v}_3, \vec{v}_2] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\
&Choice\#3: \; D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P = [\vec{v}_2, \vec{v}_3, \vec{v}_1] = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}.
\end{aligned}$$

Remark. Not every matrix is diagonalizable. For instance,

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \det(A - \lambda I) = (\lambda - 1)^2.$$

The only eigenvalue is $\lambda = 1$; it has the multiplicity $m = 2$. From

$$A - I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

we see that Null$(A - I)$ has dimension 1, and the basis consists of one vector

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

In general, if λ_0 is an eigenvalue of multiplicity m (i.e., the characteristic polynomial $\det(A-\lambda I) = (\lambda-\lambda_0)^m Q(\lambda)$), then

$$\dim(Null(A-\lambda I)) \le m.$$

Theorem. Let A be an $n \times n$ matrix with distinct real eigenvalues $\lambda_1, \lambda_2, ..., \lambda_p$ with multiplicity $m_1, m_2, ..., m_p$, respectively. Then,

1. $n_i = \dim(Null(A-\lambda_i I)) \le m_i$ and $m_1 + m_2 + ... + m_p \le n$.

2. A is diagonalizable iff $n_i = m_i$ and

$$m_1 + m_2 + ... + m_p = n.$$

   In this case, let $\mathcal{B}_i$ be a basis of $Null(A-\lambda_i I)$ for each i. Then

$$P = [\mathcal{B}_1, ..., \mathcal{B}_p], \quad D = \begin{bmatrix} \lambda_1 I_{m_1} & ... & 0 \\ ... & ... & ... \\ 0 & ... & \lambda_p I_{m_p} \end{bmatrix}, \ I_{m_i} = (m_i \times m_i) \text{ identity}$$

   i.e., the first m_1 columns of P are $\mathcal{B}_1$, the eigenvectors for λ_1, the next m_2 columns of P are $\mathcal{B}_2$, then $\mathcal{B}_3$, etc. The last m_p columns of P are $\mathcal{B}_p$; the first m_1 diagonal entries of D are λ_1, the next m_2 diagonal entries of D are λ_2, and so on.

3. In particular, if A has n distinct eigenvalues, then A is diagonalizable.

Note that there are multiple choices for assembling P. For instance, if A is 5×5, and A has two eigenvalues $\lambda_1 = 1, \lambda_2 = 2$ with a basis $\{\vec{a}_1, \vec{a}_2\}$ for $Null(A-I)$ and a basis $\{\vec{b}_1, \vec{b}_2, \vec{b}_2\}$ for $Null(A-2I)$, respectively, then, we have several choices to select pairs of (P, D) :

$$choice\#1 : P = \left[\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2, \vec{b}_2\right], \ D = \begin{bmatrix} I_2 & 0 \\ 0 & 2I_3 \end{bmatrix}$$

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

$$choice\#2 : P = \left[\vec{a}_1, \vec{b}_1, \vec{b}_2, \vec{a}_2, \vec{b}_2\right], \ D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

Example 5.3.3. Diagonalize A

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}.$$

Solution: Eigenvalues are $\lambda_1 = 5$, $m_1 = 2$, $\lambda_2 = -3$, $m_2 = 2$. For $\lambda_1 = 5$,

$$A - 5I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 4 & -8 & 0 \\ -1 & -2 & 0 & -8 \end{bmatrix} \overset{R_4+R_3\to R_4}{\longrightarrow} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 4 & -8 & 0 \\ 0 & 2 & -8 & -8 \end{bmatrix}$$
$$\to \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 4 & -8 & 0 \\ 0 & 1 & -4 & -4 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 8 & 16 \\ 0 & 1 & -4 & -4 \end{bmatrix}$$

Therefore, x_3 and x_4 are free variable, and

$$\begin{aligned} x_1 &= -8x_3 - 16x_4 \\ x_2 &= 4x_3 + 4x_4. \end{aligned}$$

Choose $(x_3, x_4) = (1, 0) \Longrightarrow x_1 = -8, x_2 = 4$; Choose $(x_3, x_4) = (0, 1) \Longrightarrow x_1 = -16,\ x_2 = 4$. We obtain two independent eigenvectors

$$\begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix} \quad (\text{for } \lambda_1 = 5).$$

For $\lambda_2 = -3$,

$$A - (-3)\,I = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ -1 & -2 & 0 & 0 \end{bmatrix} \overset{R_4+R_3\to R_4}{\longrightarrow} \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}$$
$$\overset{\substack{R_3 - 2R_4 \to R_3 \\ R_2 - 4R_4 \to R_2}}{\longrightarrow} \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix} \overset{R_1-8R_3\to R_1}{\longrightarrow} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}.$$

Hence

$$x_1 = 0,\ x_2 = 0.$$

Choose $(x_3, x_4) = (1, 0)$ and $(x_3, x_4) = (0, 1)$, respectively, we have eigenvectors

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{for } \lambda_2 = -3).$$

Assemble pairs (P, D) :

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

or

$$P = \begin{bmatrix} -8 & 0 & -16 & 0 \\ 4 & 0 & 4 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$