Chapter 5: Eigenvalues and eigenvectors

San Jose State University

Prof. Guangliang Chen

Fall 2022

Outline

Introduction

In this chapter we focus on square matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$.

We regard them as linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$:

$$\mathbf{x} \in \mathbb{R}^n \quad \longmapsto \quad \mathbf{A}\mathbf{x} \in \mathbb{R}^n$$

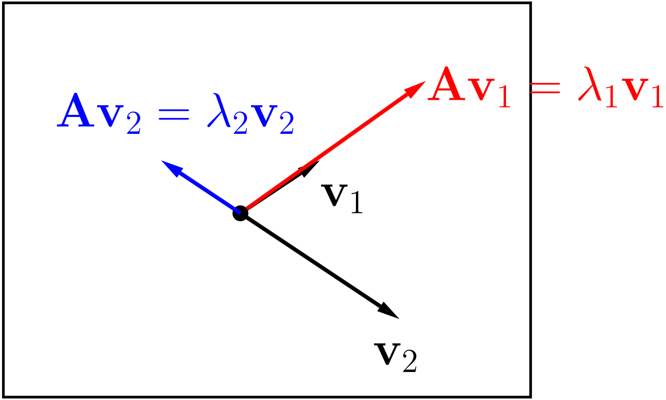
We will find special vectors $\mathbf{v} \in \mathbb{R}^n$ which are "stretched" by the matrix $\mathbf{A}$:

$$\underbrace{\mathbf{A} \cdot \mathbf{v}}_{\text{image of } \mathbf{v} \text{ under } \mathbf{A}} = \underbrace{\lambda \cdot \mathbf{v}}_{\text{a multiple of } \mathbf{v}}$$

and say that

- the scalar λ is an **eigenvalue** of $\mathbf{A}$, and
- the vector $\mathbf{v}$ is an **eigenvector** of $\mathbf{A}$ corresponding to the eigenvalue λ.

In the picture below, $\lambda_1 > 0$ and $\lambda_2 < 0$ are eigenvalues of $\mathbf{A}$ with associated eigenvectors $\mathbf{v}_1, \mathbf{v}_2$, respectively.



Example 0.1. Let

$$\mathbf{A} = \begin{bmatrix} \frac{5}{2} & -1 \\ 1 & 0 \end{bmatrix}, \ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Compute $\mathbf{A}\mathbf{v}_i$ for $i = 1, 2, 3$. Are they multiples of $\mathbf{v}_i$?

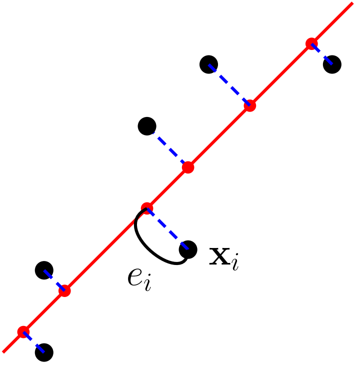
Application to orthogonal least squares fitting

Given a data set $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$, find the best-fit line that minimizes the total squared (orthogonal) fitting error

$$\sum_{i=1}^{n} e_i^2.$$

It turns out the optimal line is given by an eigenvector of some matrix.

We will introduce an algorithm later for finding such a line.



Defintion of eigenvalues and eigenvectors

Def 0.1. Let $\mathbf{A}$ be a square matrix. For any pair of scalar λ and nonzero vector $\mathbf{v}$ that satisfy the equation

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$$

we say that

- λ is an eigenvalue of $\mathbf{A}$, and
- $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ associated/corresponding to the eigenvalue λ

Remark. In the above definition, λ is allowed to be zero, so we may have zero eigenvalues (for which we have $\mathbf{Av} = \mathbf{0}$). $\longleftarrow$ We will revisit this later

Example 0.2. In the previous example where

$$\mathbf{A} = \begin{bmatrix} \frac{5}{2} & -1 \\ 1 & 0 \end{bmatrix}, \ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

we have shown that

$$\mathbf{A}\mathbf{v}_2 = 2\mathbf{v}_2, \quad \mathbf{A}\mathbf{v}_3 = \frac{1}{2}\mathbf{v}_3$$

Therefore,

- 2 is an eigenvalue of $\mathbf{A}$ and $\mathbf{v}_2$ is an eigenvector corresponding to it;
- $\frac{1}{2}$ is an eigenvalue of $\mathbf{A}$ and $\mathbf{v}_3$ is an eigenvector corresponding to it;
- $\mathbf{v}_1$ is not an eigenvector of $\mathbf{A}$ (since $\mathbf{A}\mathbf{v}_1$ is not a multiple of $\mathbf{v}_1$)

Example 0.3. Let

$$\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Determine if $\mathbf{v}$ is an eigenvector of $\mathbf{A}$. If yes, find the corresponding eigenvalue.

Example 0.4. Determine if -4 is an eigenvalue of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$. If yes, find all eigenvectors associated to it.

The previous example shows that for a fixed eigenvalue λ of a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, there are infinitely many eigenvectors associated to it. In fact, they form a subspace of $\mathbb{R}^n$.

Proof: We verify the three conditions directly:

Let $\mathbf{A}$ be an $n \times n$ matrix with an eigenvalue λ_0. Its associated eigenvectors all satisfy

$$\mathbf{A}\mathbf{v} = \lambda_0 \mathbf{v} \quad \text{i.e.} \quad (\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{v} = \mathbf{0}$$

This indicates that they comprise the null space of the $n \times n$ matrix $\mathbf{A} - \lambda_0 \mathbf{I}$, which is a subspace of $\mathbb{R}^n$.

Def 0.2. We call the subspace of all eigenvectors of $\mathbf{A}$ associated to the fixed eigenvalue λ_0, the **eigenspace** of $\mathbf{A}$ corresponding to λ_0, and denote it by

$$\mathrm{E}(\lambda_0) = \mathrm{Nul}(\mathbf{A} - \lambda_0 \mathbf{I})$$

Its dimension is called the **geometric multiplicity** of the eigenvalue λ_0:

$$g_0 = \dim \mathrm{Nul}(\mathbf{A} - \lambda_0 \mathbf{I})$$

Example 0.5. It is known that the following matrix

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

has two eigenvalues, $\lambda_1 = 1$ and $\lambda_2 = 3$. Find a basis for each of the two corresponding eigenspaces. What are the geometric multiplicities of the eigenvalues?

Existence of a zero eigenvalue $\Longleftrightarrow$ matrix is not invertible

Theorem 0.1. Let $\mathbf{A}$ be any square matrix. If $\mathbf{A}$ has a zero eigenvalue, then it is not invertible.

Proof. Suppose $\mathbf{A}$ has a zero eigenvalue with eigenvector $\mathbf{v} \neq \mathbf{0}$. That is,

$$\mathbf{A} \cdot \mathbf{v} = 0 \cdot \mathbf{v} = \mathbf{0}$$

This shows that the homogeneous equation $\mathbf{Ax} = \mathbf{0}$ has a nontrivial solution (i.e., the eigenvector $\mathbf{v}$). According to the Invertible Matrix Theorem, $\mathbf{A}$ is not invertible. □

Remark. The converse is also true, i.e., if $\mathbf{A}$ is not invertible, then 0 is an eigenvalue of $\mathbf{A}$.

Example 0.6. The matrix

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

is not invertible because 0 is an eigenvalue

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

How to find all the eigenvalues of $\mathbf{A}$

Theorem 0.2. Let $\mathbf{A}$ be any square matrix. λ is an eigenvalue of $\mathbf{A}$ if and only if $\mathbf{A} - \lambda\mathbf{I}$ is not invertible.

Proof. Suppose $\mathbf{A}$ has an eigenvalue λ with eigenvector $\mathbf{v} \neq \mathbf{0}$. That is,

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}, \quad \text{or equivalently, } (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}.$$

This shows that the homogeneous equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ has a nontrivial solution (i.e., the eigenvector $\mathbf{v}$). According to the Invertible Matrix Theorem, $\mathbf{A} - \lambda\mathbf{I}$ is not invertible.

The converse is also true by reversing the above process. □

The theorem implies that the eigenvalues λ of $\mathbf{A}$ all satisfy $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

Example 0.7. For each of the following matrices $\mathbf{A}$, find an expression in λ for $\det(\mathbf{A} - \lambda\mathbf{I})$:

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

Remark. In general, if $\mathbf{A}$ is an $n \times n$ matrix, then $\det(\mathbf{A} - \lambda\mathbf{I})$ is an nth order polynomial in λ.

The characteristic polynomial of $\mathbf{A}$

Def 0.3. For any square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$,

$$p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) \longleftarrow n\text{th order polynomial}$$

is called the **characteristic polynomial** of $\mathbf{A}$. Clearly, its roots are all and the only eigenvalues of $\mathbf{A}$ (and there are at most n of them).

If λ_0 is an eigenvalue of $\mathbf{A}$, then $\lambda - \lambda_0$ is a factor of $p(\lambda)$. Its exponent in $p(\lambda)$ is called the **algebraic multiplicity** of the eigenvalue λ_0.

Remark. It can be shown that for any eigenvalue λ_0, its geometric multiplicity never exceeds the algebraic multiplicity, i.e.,

$$1 \leq g_0 \leq a_0$$

Example 0.8. For each of the following matrices $\mathbf{A}$, find all of its eigenvalues, as well as the algebraic multiplicities. What are the geometric multiplicities?

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

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

Remark. Roots of a characteristic polynomial can be complex numbers sometimes, so a real, square matrix could have several complex eigenvalues.

Example 0.9. Find the eigenvalues of

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

Remark. The eigenvalues of a diagonal or lower/upper triangular matrix are the entries on its main diagonal.

Example 0.10. Determine the eigenvalues of the following matrix

$$\mathbf{A} = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & & \\ 4 & 2 & \\ 5 & 6 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 4 & 5 \\ & 2 & 6 \\ & & 3 \end{bmatrix}.$$

Example 0.11 (Practice question). Given the matrix $\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 5 & 1 & -1 \\ -2 & 2 & 4 \end{bmatrix}$, find its eigenvalues and associated eigenvectors. What are the algebraic and geometric multiplicities of the eigenvalues?

Eigenvectors corresponding to distinct eigenvalues must be linearly independent

Theorem 0.3. If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ of a square matrix $\mathbf{A}$, i.e.,

$$\mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1,\ \ldots,\ \mathbf{A}\mathbf{v}_k = \lambda_k \mathbf{v}_k \quad \text{(the } \lambda_i\text{'s are all different)}$$

then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ must be linearly independent.

Proof. We prove this result by using the so called Vandermonde matrix

$$\mathbf{M} = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{bmatrix}$$

which is invertible when all a_i's are dinstinct:

$$\det(\mathbf{M}) = \Pi_{1 \le i < j \le n} (a_j - a_i).$$

Detailed steps are shown in class. □

Similar matrices

Def 0.4. We say that two square matrices $\mathbf{A}, \mathbf{B}$ of the same size are **similar** if there exists an invertible matrix $\mathbf{P}$ such that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

Remark. An alternative, yet equivalent definition for $\mathbf{A}, \mathbf{B}$ to be similar is

$$\mathbf{B} = \mathbf{Q}\mathbf{A}\mathbf{Q}^{-1},$$

for an invertible matrix $\mathbf{Q}$.

Example 0.12. Verify that

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{\mathbf{B}} = \underbrace{\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}^{-1}}_{\mathbf{P}^{-1}} \underbrace{\begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{P}}$$

This shows that $\mathbf{A}, \mathbf{B}$ are similar to each other.

Remark. Similar matrices must have the same determinant. To see this, write

$$\det(\mathbf{B}) = \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \det(\mathbf{P}^{-1})\det(\mathbf{A})\det(\mathbf{P}) = \det(\mathbf{A}).$$

They are thus both invertible or non-invertible at the same time (in fact, they must have the same rank).

Similar matrices also have the same eigenvalues

Theorem 0.4. Let $\mathbf{A}, \mathbf{B}$ be two square matrices of the same size. If they are similar, then they have the same characteristic polynomial and thus the same eigenvalues (with the same algebraic multiplicities).

Proof: Suppose $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ for some invertible matrix $\mathbf{P}$ (of the same size). Then

$$\begin{aligned} p_{\mathbf{B}}(\lambda) &= \det(\mathbf{B} - \lambda\mathbf{I}) \\ &= \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{I}) \\ &= \det\left[\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}\right] \\ &= \det(\mathbf{A} - \lambda\mathbf{I}) \\ &= p_{\mathbf{A}}(\lambda) \end{aligned}$$

Remark. The converse of the theorem is not true, i.e., if $\mathbf{A}, \mathbf{B}$ have the same eigenvalues, then they are not necessarily similar to each other.

A counterexample is

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

They have the same characteristic polynomial and thus the same eigenvalues, but they are not similar. $\longleftarrow$ Why?

Diagonalizability of square matrices

Def 0.5. A square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is **diagonalizable** if it is similar to a diagonal matrix, i.e., there exist an invertible matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \qquad \text{or equivalently,} \quad \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}.$$

Remark. If we write $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_n]$ and $\mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, then the above equation can be rewritten as

$$\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D},$$

or in columns

$$\mathbf{A}[\mathbf{p}_1 \ldots \mathbf{p}_n] = [\mathbf{p}_1 \ldots \mathbf{p}_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

From this we get that

$$\mathbf{A}\mathbf{p}_i = \lambda_i \mathbf{p}_i, \, 1 \le i \le n.$$

This shows that $\mathbf{A}$ has n eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (not necessarily distinct) with corresponding eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^n$ which are linearly independent.

Thus, the above factorization of a diagonalizable matrix $\mathbf{A}$, i.e.,

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

is called the **eigenvalue decomposition**, or simply **eigendecomposition**, of $\mathbf{A}$.

Example 0.13. The matrix

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

is diagonalizable because

$$\begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 3 & \\ & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}^{-1} \longleftarrow \text{eigendecomposition}$$

but the matrix

$$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$$

is not (we will see why later).

Why are diagonalizable matrices important?

Every diagonalizable matrix is similar to a diagonal matrix (that consists of its eigenvalues), and is easy to deal with in a lot of ways.

For example, it can help compute **matrix powers** $(\mathbf{A}^k)$. To see this, suppose $\mathbf{A} \in \mathbb{R}^{n\times n}$ is diagonalizable, that is, $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ for some invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$. Then

$$\begin{aligned}
\mathbf{A}^2 &= \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}^2\mathbf{P}^{-1} \\
\mathbf{A}^3 &= \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}^3\mathbf{P}^{-1} \\
\mathbf{A}^k &= \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} \quad \text{(for any positive integer } k)
\end{aligned}$$

where $\mathbf{D}^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$.

Example 0.14. For the diagonalizable matrix in the preceding example,

$$\underbrace{\begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix}}_{\mathbf{A}} = \underbrace{\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}}_{\mathbf{P}} \underbrace{\begin{pmatrix} 3 & \\ & -1 \end{pmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}^{-1}}_{\mathbf{P}^{-1}}$$

the 10th power of $\mathbf{A}$ is

$$\mathbf{A}^{10} = \mathbf{P}\mathbf{D}^{10}\mathbf{P}^{-1} = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 3^{10} & \\ & (-1)^{10} \end{pmatrix} \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{3}{4} & -\frac{1}{4} \end{pmatrix} = \begin{pmatrix} 14763 & 14762 \\ 44286 & 44287 \end{pmatrix}.$$

Checking diagonalizability of a square matrix

Theorem 0.5. A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is diagonalizable if and only if it has n linearly independent eigenvectors (i.e., $\sum g_i = n$).

Proof. We have already proved this result earlier:

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \iff \mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D} \iff \mathbf{A}\mathbf{p}_i = \lambda_i \mathbf{p}_i, 1 \le i \le n$$

The $\mathbf{p}_i$'s are linearly independent if and only if $\mathbf{P}$ is invertible.

Example 0.15. The matrix $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$ is not diagonalizable because it has one distinct eigenvalue $\lambda_1 = 1$ with $a_1 = 2$ and $g_1 = 1$ (only one linearly independent eigenvector).

Example 0.16. Is the following matrix diagonalizable? If yes, find the eigendecomposition.

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

The previous theorem immediately implies the following results.

Corollary 0.6. Any matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ with n distinct eigenvalues, i.e.,

$$a_i = 1, \quad 1 \le i \le n,$$

is diagonalizable.

,

Example 0.17. Is the following matrix diagonalizable? If yes, find the eigendecomposition.

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

To be introduced later

Another important result is that symmetric matrices, i.e., $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{A}^T = \mathbf{A}$, are always diagonalizable.

We will learn this in Chapter 7.