Linear Algebra 2

Lecture #21

Volume via determinants. An introduction to eigenvalues and eigenvectors

Irena Penev

April 16, 2025

- This lecture has three parts:

- This lecture has three parts:
    1. Volume via determinants

- This lecture has three parts:
    1. Volume via determinants
    2. Eigenvalues and eigenvectors of linear functions and square matrices

- This lecture has three parts:
  1. Volume via determinants
  2. Eigenvalues and eigenvectors of linear functions and square matrices
  3. The characteristic polynomial and spectrum

1. Determinants and volume

1. Determinants and volume

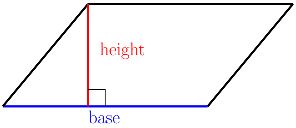
- In our study of determinants and volume, we assume throughout that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

1. Determinants and volume

- In our study of determinants and volume, we assume throughout that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

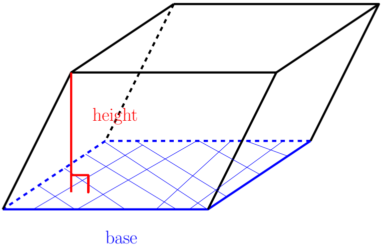
- For a parallelogram, we have the familiar formula

$$\begin{pmatrix} \text{area of} \\ \text{parallelogram} \end{pmatrix} = (\text{length of base}) \times (\text{height}).$$



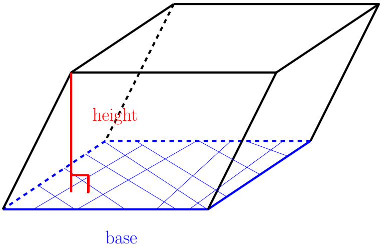
- We have a similar formula for the volume of a parallelepiped:

$$\begin{pmatrix} \text{volume of} \\ \text{parallelepiped} \end{pmatrix} = (\text{area of base}) \times (\text{height}).$$



- We have a similar formula for the volume of a parallelepiped:

$$\begin{pmatrix} \text{volume of} \\ \text{parallelepiped} \end{pmatrix} = (\text{area of base}) \times (\text{height}).$$



- We would now like to generalize this to arbitrary dimensions (next slide).

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

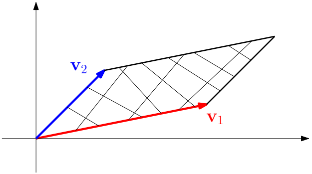
$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1 \right\}.$$

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \le c_1, \ldots, c_m \le 1 \right\}.$$

- For instance, given two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$, neither of which is a scalar multiple of the other, the 2-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2$ is just the usual parallelogram determined by these two vectors.

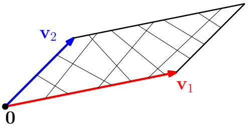


Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1 \right\}.$$

- For vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, neither of which is a scalar multiple of each other, the 2-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2$ is still a parallelogram, but this parallelogram lies in the plane (2-dimensional subspace) Span$(\mathbf{v}_1, \mathbf{v}_2)$ of $\mathbb{R}^n$.



Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\; 0 \le c_1, \ldots, c_m \le 1 \right\}.$$

- What happens if one of $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ is a scalar multiple of the other, say $\mathbf{v}_2 = \alpha\mathbf{v}_1$ for some scalar $\alpha \in \mathbb{R}$?

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1 \right\}.$$

- What happens if one of $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ is a scalar multiple of the other, say $\mathbf{v}_2 = \alpha\mathbf{v}_1$ for some scalar $\alpha \in \mathbb{R}$?
- Then the 2-parallelepiped determined by $\mathbf{v}_1$ and $\mathbf{v}_2$ is just set

$$\begin{aligned} & \left\{ c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \mid c_1, c_2 \in \mathbb{R},\ 0 \leq c_1, c_2 \leq 1 \right\} \\ = & \left\{ c_1\mathbf{v}_1 + c_2\alpha\mathbf{v}_1 \;\middle|\; c_1, c_2 \in \mathbb{R},\ 0 \leq c_1, c_2 \leq 1 \right\} \\ = & \left\{ (c_1 + c_2\alpha)\mathbf{v}_1 \;\middle|\; c_1, c_2 \in \mathbb{R},\ 0 \leq c_1, c_2 \leq 1 \right\} \\ = & \left\{ c(1+\alpha)\mathbf{v}_1 \;\middle|\; c \in \mathbb{R},\ 0 \leq c \leq 1 \right\}, \end{aligned}$$

which is 1-dimensional (a line segment) if $\mathbf{v}_1 \neq \mathbf{0}$, and is 0-dimensional (containing only the zero vector) if $\mathbf{v}_1 = \mathbf{0}$.

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1\right\}.$$

- What happens if one of $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ is a scalar multiple of the other, say $\mathbf{v}_2 = \alpha\mathbf{v}_1$ for some scalar $\alpha \in \mathbb{R}$?
- Then the 2-parallelepiped determined by $\mathbf{v}_1$ and $\mathbf{v}_2$ is just set

$$\begin{aligned} & \left\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \mid c_1, c_2 \in \mathbb{R},\ 0 \leq c_1, c_2 \leq 1\right\} \\ = & \left\{c_1\mathbf{v}_1 + c_2\alpha\mathbf{v}_1 \;\middle|\; c_1, c_2 \in \mathbb{R},\ 0 \leq c_1, c_2 \leq 1\right\} \\ = & \left\{(c_1 + c_2\alpha)\mathbf{v}_1 \;\middle|\; c_1, c_2 \in \mathbb{R},\ 0 \leq c_1, c_2 \leq 1\right\} \\ = & \left\{c(1+\alpha)\mathbf{v}_1 \;\middle|\; c \in \mathbb{R},\ 0 \leq c \leq 1\right\}, \end{aligned}$$

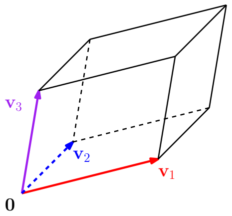
  which is 1-dimensional (a line segment) if $\mathbf{v}_1 \neq \mathbf{0}$, and is 0-dimensional (containing only the zero vector) if $\mathbf{v}_1 = \mathbf{0}$.
- We can think of these as "degenerate parallelograms."

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1 \right\}.$$

- Similarly, for three linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^n$, the 3-parallelepiped defined by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is just the usual parallelepiped whose edges are determined by these three vectors (see the picture below).



Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \le c_1, \ldots, c_m \le 1 \right\}.$$

- If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is not linearly independent, then the 3-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is either a parallelogram, or a line segment, or $\{\mathbf{0}\}$, depending on the dimension of $\text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.
  - Once again, we can think of these as "degenerate parallelepipeds."

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \le c_1, \ldots, c_m \le 1 \right\}.$$

- If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is not linearly independent, then the 3-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is either a parallelogram, or a line segment, or $\{\mathbf{0}\}$, depending on the dimension of $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.
  - Once again, we can think of these as "degenerate parallelepipeds."
- For more than three vectors, we get higher-dimensional generalizations.

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \le c_1, \ldots, c_m \le 1 \right\}.$$

- We would now like to define the "volume" (more precisely, the "m-volume") of an m-parallelepiped in $\mathbb{R}^n$.

Definition

Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1 \right\}.$$

- We would now like to define the "volume" (more precisely, the "m-volume") of an m-parallelepiped in $\mathbb{R}^n$.
- We do this recursively, as follows (next slide).

Definition

- The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$.
- For a positive integer m, the $(m+1)$-volume of the $(m+1)$-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1} \in \mathbb{R}^n$ is defined to be

$$V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) \quad := \quad V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m)\, ||\mathbf{v}_{m+1}^{\perp}||,$$

where $\mathbf{v}_{m+1}^{\perp} = \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$.[a]

---

[a] Equivalently (by Corollary 6.5.3): $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)}(\mathbf{v}_{m+1})$.

Definition

- The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$.
- For a positive integer m, the $(m+1)$-volume of the $(m+1)$-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1} \in \mathbb{R}^n$ is defined to be

$$V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) \quad := \quad V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \, ||\mathbf{v}_{m+1}^{\perp}||,$$

where $\mathbf{v}_{m+1}^{\perp} = \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$.[a]

[a] Equivalently (by Corollary 6.5.3): $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)}(\mathbf{v}_{m+1})$.

- In this recursive formula, the m-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is our "base" and $||\mathbf{v}_{m+1}^{\perp}||$ is our "height."

Definition

- The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$.
- For a positive integer m, the $(m+1)$-volume of the $(m+1)$-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1} \in \mathbb{R}^n$ is defined to be

$$V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) \quad := \quad V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \, ||\mathbf{v}_{m+1}^{\perp}||,$$

where $\mathbf{v}_{m+1}^{\perp} = \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$.[a]

[a] Equivalently (by Corollary 6.5.3): $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)}(\mathbf{v}_{m+1})$.

- In this recursive formula, the m-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is our "base" and $||\mathbf{v}_{m+1}^{\perp}||$ is our "height."
- So, we get the formula

$$\begin{pmatrix} (m+1)\text{-volume of} \\ (m+1)\text{-parallelepiped} \end{pmatrix} \quad = \quad (m\text{-volume of base}) \quad \times \quad (\text{height}).$$

Definition

- The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$.
- For a positive integer m, the $(m+1)$-volume of the $(m+1)$-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1} \in \mathbb{R}^n$ is defined to be

$$V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) \quad := \quad V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m)\, ||\mathbf{v}_{m+1}^{\perp}||,$$

where $\mathbf{v}_{m+1}^{\perp} = \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$.[a]

[a] Equivalently (by Corollary 6.5.3): $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)}(\mathbf{v}_{m+1})$.

- Note that 1-volume represents (1-dimensional) length, 2-volume represents (2-dimensional) area, and 3-volume represents (3-dimensional) volume.
- For $m \geq 4$, m-volume is an m-dimensional generalization of these concepts.

Proposition 7.10.1

Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$. Then $V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \geq 0$, and equality holds iff $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a linearly dependent set.

- Proof: Lecture Notes.
  - The fact that $V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \geq 0$ follows straight from the definition of m-volume (we keep computing lengths of vectors).
  - The second statement essentially states that the volume of an m-parallelepiped is zero iff that m-parallelepiped is "degenerate."

Definition

- The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$.
- For a positive integer m, the $(m+1)$-volume of the $(m+1)$-parallelepiped determined by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1} \in \mathbb{R}^n$ is defined to be

$$V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) \quad := \quad V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m)\, ||\mathbf{v}_{m+1}^{\perp}||,$$

where $\mathbf{v}_{m+1}^{\perp} = \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$.[a]

[a] Equivalently (by Corollary 6.5.3): $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)}(\mathbf{v}_{m+1})$.

- We will prove the following four results about m-volume (next two slides):

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) = \sqrt{\det(A^T A)}.$$

- Note that A is an $n \times m$ matrix. It is possible that $n \neq m$, and so $\det(A)$ is not necessarily defined.
- However, $A^T A$ is an $m \times m$ matrix, and so $\det(A^T A)$ is defined.

Theorem 7.10.2

Let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \dots, \mathbf{a}_m) = \sqrt{\det(A^T A)}.$$

- Note that A is an $n \times m$ matrix. It is possible that $n \neq m$, and so $\det(A)$ is not necessarily defined.
- However, $A^T A$ is an $m \times m$ matrix, and so $\det(A^T A)$ is defined.

Corollary 7.10.3

Let $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$. Then $V_n(\mathbf{a}_1, \dots, \mathbf{a}_n) = |\det(\begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix})|$.

- Note that we have n vectors in $\mathbb{R}^n$. So, $\begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ is an $n \times n$ matrix, and therefore, it has a determinant.

Corollary 7.10.4

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $\sigma \in S_m$. Then $V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) = V_m(\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(m)})$.

- So, merely permuting the vectors that determine an m-parallelepiped does not change the m-volume of that m-parallelepiped.

Corollary 7.10.4

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $\sigma \in S_m$. Then $V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) = V_m(\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(m)})$.

- So, merely permuting the vectors that determine an m-parallelepiped does not change the m-volume of that m-parallelepiped.

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) = |\det(A)|\, V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

- Here, it is important that we have n vectors in $\mathbb{R}^n$.
- If we have m vectors in $\mathbb{R}^n$, then this fails.
  - Counterexample: later!

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \quad = \quad \sqrt{\det(A^T A)}.$$

Proof.

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Proof. $\forall i \in \{1, \ldots, m\}$: $A_i := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{bmatrix}$.

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Proof. $\forall i \in \{1, \ldots, m\}$: $A_i := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{bmatrix}$. We will prove inductively that $\forall i \in \{1, \ldots, m\}$: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$.

For $i = 1$, we observe that $A_1^T A_1 = \begin{bmatrix} \mathbf{a}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{a}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 \end{bmatrix}$,

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Proof. $\forall i \in \{1, \ldots, m\}$: $A_i := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{bmatrix}$. We will prove inductively that $\forall i \in \{1, \ldots, m\}$: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. Obviously, this is enough, since $A_m = A$.

For $i = 1$, we observe that $A_1^T A_1 = \begin{bmatrix} \mathbf{a}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{a}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 \end{bmatrix}$, and consequently,

$$\sqrt{\det(A_1^T A_1)} \;=\; \sqrt{\mathbf{a}_1 \cdot \mathbf{a}_1} \;=\; ||\mathbf{a}_1|| \;=\; V_1(\mathbf{a}_1).$$

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Proof. $\forall i \in \{1, \ldots, m\}$: $A_i := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{bmatrix}$. We will prove inductively that $\forall i \in \{1, \ldots, m\}$: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. Obviously, this is enough, since $A_m = A$.

For $i = 1$, we observe that $A_1^T A_1 = \begin{bmatrix} \mathbf{a}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{a}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 \end{bmatrix}$, and consequently,

$$\sqrt{\det(A_1^T A_1)} \;=\; \sqrt{\mathbf{a}_1 \cdot \mathbf{a}_1} \;=\; ||\mathbf{a}_1|| \;=\; V_1(\mathbf{a}_1).$$

We may now assume that $m \geq 2$, for otherwise we are done by what we just showed.

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \quad = \quad \sqrt{\det(A^T A)}.$$

Proof. $\forall i \in \{1, \ldots, m\}$: $A_i := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{bmatrix}$. We will prove inductively that $\forall i \in \{1, \ldots, m\}$: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. Obviously, this is enough, since $A_m = A$.

For $i = 1$, we observe that $A_1^T A_1 = \begin{bmatrix} \mathbf{a}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{a}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 \end{bmatrix}$, and consequently,

$$\sqrt{\det(A_1^T A_1)} \quad = \quad \sqrt{\mathbf{a}_1 \cdot \mathbf{a}_1} \quad = \quad ||\mathbf{a}_1|| \quad = \quad V_1(\mathbf{a}_1).$$

We may now assume that $m \geq 2$, for otherwise we are done by what we just showed. Fix $i \in \{1, \ldots, m-1\}$, and assume inductively that $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$;
$A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$; WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

- $\mathbf{a}_{i+1}^{||} := \text{proj}_{\text{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_i)}(\mathbf{a}_{i+1})$;
- $\mathbf{a}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_i)^{\perp}}(\mathbf{a}_{i+1})$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$;
$A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

- $\mathbf{a}_{i+1}^{||} := \text{proj}_{\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)}(\mathbf{a}_{i+1})$;
- $\mathbf{a}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)^\perp}(\mathbf{a}_{i+1})$.

By Corollary 6.5.3, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

- $\mathbf{a}_{i+1}^{||} := \text{proj}_{\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)}(\mathbf{a}_{i+1})$;
- $\mathbf{a}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)^{\perp}}(\mathbf{a}_{i+1})$.

By Corollary 6.5.3, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$.

Since $\mathbf{a}_{i+1}^{||} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)$, $\exists c_1, \dots, c_i \in \mathbb{R}$ s.t. $\mathbf{a}_{i+1}^{||} = c_1\mathbf{a}_1 + \dots + c_i\mathbf{a}_i$, and consequently,

$$\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - \mathbf{a}_i^{||} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i.$$

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

- $\mathbf{a}_{i+1}^{||} := \text{proj}_{\text{Span}(\mathbf{a}_1,\dots,\mathbf{a}_i)}(\mathbf{a}_{i+1})$;
- $\mathbf{a}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{a}_1,\dots,\mathbf{a}_i)^{\perp}}(\mathbf{a}_{i+1})$.

By Corollary 6.5.3, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$.

Since $\mathbf{a}_{i+1}^{||} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)$, $\exists c_1, \dots, c_i \in \mathbb{R}$ s.t. $\mathbf{a}_{i+1}^{||} = c_1\mathbf{a}_1 + \dots + c_i\mathbf{a}_i$, and consequently,

$$\mathbf{a}_{i+1}^{\perp} \;=\; \mathbf{a}_{i+1} - \mathbf{a}_i^{||} \;=\; \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i.$$

Now, let B_{i+1} be the matrix obtained from A_{i+1} by replacing the rightmost column of A_{i+1} by $\mathbf{a}_{i+1}^{\perp}$, i.e.

$$B_{i+1} \;:=\; \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1}^{\perp} \end{bmatrix}.$$

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$;
$A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

- $\mathbf{a}_{i+1}^{||} := \text{proj}_{\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)}(\mathbf{a}_{i+1})$;
- $\mathbf{a}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)^{\perp}}(\mathbf{a}_{i+1})$.

By Corollary 6.5.3, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$.

Since $\mathbf{a}_{i+1}^{||} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)$, $\exists c_1, \dots, c_i \in \mathbb{R}$ s.t.
$\mathbf{a}_{i+1}^{||} = c_1\mathbf{a}_1 + \dots + c_i\mathbf{a}_i$, and consequently,

$$\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - \mathbf{a}_i^{||} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i.$$

Now, let B_{i+1} be the matrix obtained from A_{i+1} by replacing the rightmost column of A_{i+1} by $\mathbf{a}_{i+1}^{\perp}$, i.e.

$$B_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1}^{\perp} \end{bmatrix}.$$

WTS $\det(A_{i+1}^T A_{i+1}) = \det(B_{i+1}^T B_{i+1}) \overset{(*)}{=} V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1})^2$,
where for (*) we will use the ind. hyp. and the def. of volume.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$;
$A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i$.

$$B_{i+1}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ (\mathbf{a}_{i+1}^{\perp})^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ \mathbf{a}_{i+1}^T - c_1\mathbf{a}_1^T - \dots - c_i\mathbf{a}_i^T \end{bmatrix}.$$

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i$.

$$B_{i+1}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ (\mathbf{a}_{i+1}^{\perp})^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ \mathbf{a}_{i+1}^T - c_1\mathbf{a}_1^T - \dots - c_i\mathbf{a}_i^T \end{bmatrix}.$$

So, B_{i+1}^T can be obtained from A_{i+1}^T via the following sequence of i elementary row operations:

- $R_{i+1} \to R_{i+1} - c_1R_1$;

  $\vdots$

- $R_{i+1} \to R_{i+1} - c_iR_i$.

Let $E_1, \dots, E_i$ be the elementary matrices corresponding to these i elementary row operations,

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i$.

$$B_{i+1}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ (\mathbf{a}_{i+1}^{\perp})^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ \mathbf{a}_{i+1}^T - c_1\mathbf{a}_1^T - \dots - c_i\mathbf{a}_i^T \end{bmatrix}.$$

So, B_{i+1}^T can be obtained from A_{i+1}^T via the following sequence of i elementary row operations:

- $R_{i+1} \to R_{i+1} - c_1R_1$;

$\vdots$

- $R_{i+1} \to R_{i+1} - c_iR_i$.

Let $E_1, \dots, E_i$ be the elementary matrices corresponding to these i elementary row operations, so that $B_{i+1}^T = E_i \dots E_1A_{i+1}^T$, and consequently, $B_{i+1} = A_{i+1}E_1^T \dots E_i^T$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i$.

$$B_{i+1}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ (\mathbf{a}_{i+1}^{\perp})^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ \mathbf{a}_{i+1}^T - c_1\mathbf{a}_1^T - \dots - c_i\mathbf{a}_i^T \end{bmatrix}.$$

So, B_{i+1}^T can be obtained from A_{i+1}^T via the following sequence of i elementary row operations:

- $R_{i+1} \to R_{i+1} - c_1R_1$;

  $\vdots$

- $R_{i+1} \to R_{i+1} - c_iR_i$.

Let $E_1, \dots, E_i$ be the elementary matrices corresponding to these i elementary row operations, so that $B_{i+1}^T = E_i \dots E_1A_{i+1}^T$, and consequently, $B_{i+1} = A_{i+1}E_1^T \dots E_i^T$. By Theorem 7.3.2(c), we see that $\det(E_1) = \dots = \det(E_i) = 1$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \dots - c_i\mathbf{a}_i$.

$$B_{i+1}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ (\mathbf{a}_{i+1}^{\perp})^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \\ \mathbf{a}_{i+1}^T - c_1\mathbf{a}_1^T - \dots - c_i\mathbf{a}_i^T \end{bmatrix}.$$

So, B_{i+1}^T can be obtained from A_{i+1}^T via the following sequence of i elementary row operations:

- $R_{i+1} \to R_{i+1} - c_1R_1$;

  $\vdots$

- $R_{i+1} \to R_{i+1} - c_iR_i$.

Let $E_1, \dots, E_i$ be the elementary matrices corresponding to these i elementary row operations, so that $B_{i+1}^T = E_i \dots E_1 A_{i+1}^T$, and consequently, $B_{i+1} = A_{i+1}E_1^T \dots E_i^T$. By Theorem 7.3.2(c), we see that $\det(E_1) = \dots = \det(E_i) = 1$. We now compute (next slide):

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$;
$A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

$$
\begin{aligned}
\det(B_{i+1}^T B_{i+1}) &= \det\Big((E_i \dots E_1 A_{i+1}^T)(A_{i+1} E_1^T \dots E_i^T)\Big) \\
&\overset{(*)}{=} \det(E_i) \dots \det(E_1) \det(A_{i+1}^T A_{i+1}) \det(E_1^T) \dots \det(E_i^T) \\
&\overset{(**)}{=} \underbrace{\det(E_i)}_{=1} \dots \underbrace{\det(E_1)}_{=1} \det(A_{i+1}^T A_{i+1}) \underbrace{\det(E_1)}_{=1} \dots \underbrace{\det(E_i)}_{=1} \\
&= \det(A_{i+1}^T A_{i+1}),
\end{aligned}
$$

where (*) follows from Theorem 7.5.2, and (**) follows from Theorem 7.1.3.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

$$\begin{aligned}
\det(B_{i+1}^T B_{i+1}) &= \det\Big((E_i \dots E_1 A_{i+1}^T)(A_{i+1} E_1^T \dots E_i^T)\Big) \\
&\overset{(*)}{=} \det(E_i) \dots \det(E_1) \det(A_{i+1}^T A_{i+1}) \det(E_1^T) \dots \det(E_i^T) \\
&\overset{(**)}{=} \underbrace{\det(E_i)}_{=1} \dots \underbrace{\det(E_1)}_{=1} \det(A_{i+1}^T A_{i+1}) \underbrace{\det(E_1)}_{=1} \dots \underbrace{\det(E_i)}_{=1} \\
&= \det(A_{i+1}^T A_{i+1}),
\end{aligned}$$

where (*) follows from Theorem 7.5.2, and (**) follows from Theorem 7.1.3.

But note that $B_{i+1} = \left[\begin{array}{c:c} A_i & \mathbf{a}_{i+1}^{\perp} \end{array}\right]$, and so (next slide):

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$; WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

$$\begin{aligned} B_{i+1}^T B_{i+1} &= \left[\begin{array}{c} A_i^T \\ \hline (\mathbf{a}_{i+1}^\perp)^T \end{array}\right] \left[\begin{array}{c|c} A_i & \mathbf{a}_{i+1}^\perp \end{array}\right] \\ &= \left[\begin{array}{c|c} A_i^T A_i & A_i^T \mathbf{a}_{i+1}^\perp \\ \hline (\mathbf{a}_{i+1}^\perp)^T A_i & (\mathbf{a}_{i+1}^\perp)^T \mathbf{a}_{i+1}^\perp \end{array}\right] \\ &\overset{(*)}{=} \left[\begin{array}{c|c} A_i^T A_i & \mathbf{0} \\ \hline \mathbf{0}^T & ||\mathbf{a}_{i+1}^\perp||^2 \end{array}\right], \end{aligned}$$

where in (*), we used the fact that $\mathbf{a}_{i+1}^\perp$ is orthogonal to the columns of A, and so $A^T \mathbf{a}_{i+1}^\perp = \mathbf{0}$, and we also used the fact that $(\mathbf{a}_{i+1}^\perp)^T \mathbf{a}_{i+1}^\perp = \mathbf{a}_{i+1}^\perp \cdot \mathbf{a}_{i+1}^\perp = ||\mathbf{a}_{i+1}^\perp||^2$.

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$;
$A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

$$
\begin{aligned}
B_{i+1}^T B_{i+1} &= \left[\begin{array}{c} A_i^T \\ \hline (\mathbf{a}_{i+1}^\perp)^T \end{array}\right] \left[\begin{array}{c|c} A_i & \mathbf{a}_{i+1}^\perp \end{array}\right] \\
&= \left[\begin{array}{c|c} A_i^T A_i & A_i^T \mathbf{a}_{i+1}^\perp \\ \hline (\mathbf{a}_{i+1}^\perp)^T A_i & (\mathbf{a}_{i+1}^\perp)^T \mathbf{a}_{i+1}^\perp \end{array}\right] \\
&\overset{(*)}{=} \left[\begin{array}{c|c} A_i^T A_i & \mathbf{0} \\ \hline \mathbf{0}^T & ||\mathbf{a}_{i+1}^\perp||^2 \end{array}\right],
\end{aligned}
$$

where in (*), we used the fact that $\mathbf{a}_{i+1}^\perp$ is orthogonal to the columns of A, and so $A^T \mathbf{a}_{i+1}^\perp = \mathbf{0}$, and we also used the fact that $(\mathbf{a}_{i+1}^\perp)^T \mathbf{a}_{i+1}^\perp = \mathbf{a}_{i+1}^\perp \cdot \mathbf{a}_{i+1}^\perp = ||\mathbf{a}_{i+1}^\perp||^2$.

We now compute (next slide):

Proof (continued). Reminder: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$; $A_{i+1} := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i & \mathbf{a}_{i+1} \end{bmatrix}$; $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

$$\begin{aligned}
\det(A_{i+1}^T A_{i+1}) &= \det(B_{i+1}^T B_{i+1}) \\
&= \left| \begin{array}{c|c} A_i^T A_i & \mathbf{0} \\ \hline \mathbf{0}^T & ||\mathbf{a}_{i+1}^{\perp}||^2 \end{array} \right| \\
&\overset{(*)}{=} (-1)^{(i+1)+(i+1)} \, ||\mathbf{a}_{i+1}^{\perp}||^2 \det(A_i^T A_i) \\
&= \det(A_i^T A_i) \, ||\mathbf{a}_{i+1}^{\perp}||^2 \\
&\overset{(**)}{=} V_i(\mathbf{a}_1, \dots, \mathbf{a}_i)^2 \, ||\mathbf{a}_{i+1}^{\perp}||^2 \\
&\overset{(***)}{=} V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1})^2,
\end{aligned}$$

where (*) follows by Laplace expansion along the rightmost column, (**) follows from the induction hypothesis, and (***) follows from the definition of $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1})$.

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

From the previous slide:

$$\det(A_{i+1}^T A_{i+1}) \;=\; V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1})^2.$$

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

From the previous slide:

$$\det(A_{i+1}^T A_{i+1}) \;=\; V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1})^2.$$

Since $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) \geq 0$ (by Proposition 7.10.1), we may now take the square root of both sides to obtain

$$V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) \;=\; \sqrt{\det(A_{i+1}^T A_{i+1})}.$$

This completes the induction. □

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Corollary 7.10.3

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Then $V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) = |\det(\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix})|$.

Proof.

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Corollary 7.10.3

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Then $V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) = |\det(\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix})|$.

Proof. First of all, we note that $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix}$ is an $n \times n$ matrix (with entries in $\mathbb{R}$), and so it has a determinant.

Theorem 7.10.2

Let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \dots, \mathbf{a}_m) = \sqrt{\det(A^T A)}.$$

Corollary 7.10.3

Let $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$. Then $V_n(\mathbf{a}_1, \dots, \mathbf{a}_n) = |\det(\begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix})|$.

Proof. First of all, we note that $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ is an $n \times n$ matrix (with entries in $\mathbb{R}$), and so it has a determinant. We now compute:

$$\begin{aligned} V_n(\mathbf{a}_1, \dots, \mathbf{a}_n) &= \sqrt{\det(A^T A)} && \text{by Theorem 7.10.2} \\ &= \sqrt{\det(A^T)\det(A)} && \text{by Theorem 7.5.2} \\ &= \sqrt{\det(A)^2} && \text{by Theorem 7.1.3} \\ &= |\det(A)|. \end{aligned}$$

□

Theorem 7.10.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) \quad = \quad \sqrt{\det(A^T A)}.$$

Corollary 7.10.4

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $\sigma \in S_m$. Then $V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m) = V_m(\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(m)})$.

Proof.

Theorem 7.10.2

Let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. Then

$$V_m(\mathbf{a}_1, \dots, \mathbf{a}_m) \;=\; \sqrt{\det(A^T A)}.$$

Corollary 7.10.4

Let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ and $\sigma \in S_m$. Then $V_m(\mathbf{a}_1, \dots, \mathbf{a}_m) = V_m(\mathbf{a}_{\sigma(1)}, \dots, \mathbf{a}_{\sigma(m)})$.

Proof. Set $A := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$ and $A_\sigma := \begin{bmatrix} \mathbf{a}_{\sigma(1)} & \dots & \mathbf{a}_{\sigma(m)} \end{bmatrix}$, and consider P_σ, the matrix of the permutation σ. By Theorem 2.3.15(c), we have that $A_\sigma = AP_\sigma^T$, and by Proposition 7.1.1, we have that $\det(P_\sigma) = \text{sgn}(\sigma)$. But now (next slide):

Proof (continued).

$$
\begin{aligned}
V_m(\mathbf{a}_{\sigma(1)}, \dots, \mathbf{a}_{\sigma(m)}) &\overset{(*)}{=} \sqrt{\det(A_\sigma^T A_\sigma)} \\
&= \sqrt{\det\big((AP_\sigma^T)^T (AP_\sigma^T)\big)} \\
&= \sqrt{\det(P_\sigma A^T A P_\sigma^T)} \\
&\overset{(**)}{=} \sqrt{\det(P_\sigma)\det(A^T A)\det(P_\sigma^T)} \\
&\overset{(***)}{=} \sqrt{\det(P_\sigma)\det(A^T A)\,\det(P_\sigma)} \\
&= \sqrt{\mathrm{sgn}(\sigma)^2 \det(A^T A)} \\
&= \sqrt{\det(A^T A)} \\
&\overset{(*)}{=} V_m(\mathbf{a}_1, \dots, \mathbf{a}_m),
\end{aligned}
$$

where both instances of (*) follow from Theorem 7.10.2, (**) follows from Theorem 7.5.2, and (***) follows from Theorem 7.1.3. □

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n\times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \quad = \quad |\det(A)| \; V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

Proof.

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n\times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \quad = \quad |\det(A)|\ V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

Proof. Set $B := \begin{bmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{bmatrix}$ and $C := \begin{bmatrix} A\mathbf{v}_1 & \ldots & A\mathbf{v}_n \end{bmatrix} = AB$.

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \quad = \quad |\det(A)| \; V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

Proof. Set $B := \left[\begin{array}{ccc} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{array} \right]$ and $C := \left[\begin{array}{ccc} A\mathbf{v}_1 & \ldots & A\mathbf{v}_n \end{array} \right] = AB$. Note that A, B, and $C = AB$ all belong to $\mathbb{R}^{n \times n}$, and so all three matrices have determinants.

Corollary 7.10.5

Let $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \dots, A\mathbf{v}_n) \quad = \quad |\det(A)| \; V_n(\mathbf{v}_1, \dots, \mathbf{v}_n).$$

Proof. Set $B := \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$ and $C := \begin{bmatrix} A\mathbf{v}_1 & \dots & A\mathbf{v}_n \end{bmatrix} = AB$. Note that A, B, and $C = AB$ all belong to $\mathbb{R}^{n \times n}$, and so all three matrices have determinants. We now compute:

$$\begin{aligned}
V_n(A\mathbf{v}_1, \dots, A\mathbf{v}_n) &\overset{\text{Thm. 7.10.2}}{=} \sqrt{\det(C^T C)} \\
&= \sqrt{\det\Big((AB)^T (AB)\Big)} \\
&= \sqrt{\det(B^T A^T AB)} \\
&\overset{\text{Thm. 7.5.2}}{=} \sqrt{\det(B^T)\det(A^T)\det(A)\det(B)} \\
&\overset{\text{Thm. 7.1.3}}{=} \sqrt{\det(A)^2 \det(B^T)\det(B)} \\
&\overset{\text{Thm. 7.5.2}}{=} \sqrt{\det(A)^2 \, \det(B^T B)} \\
&= |\det(A)| \sqrt{\det(B^T B)} \\
&\overset{\text{Thm. 7.10.2}}{=} |\det(A)| \; V_n(\mathbf{v}_1, \dots, \mathbf{v}_n).
\end{aligned}$$

□

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \quad = \quad |\det(A)| \; V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \;=\; |\det(A)|\; V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

- **Remark:** For $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ $(m \neq n)$ and $A \in \mathbb{R}^{n \times n}$, the formula from Corollary 7.10.5 fails, i.e.

$$V_m(A\mathbf{v}_1, \ldots, A\mathbf{v}_m) \;\neq\; |\det(A)|\; V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m).$$

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \quad = \quad |\det(A)| \; V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

- **Remark:** For $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ $(m \neq n)$ and $A \in \mathbb{R}^{n \times n}$, the formula from Corollary 7.10.5 fails, i.e.

$$V_m(A\mathbf{v}_1, \ldots, A\mathbf{v}_m) \quad \not\asymp \quad |\det(A)| \; V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m).$$

  - For instance, for $m = 1$ and $n = 2$, we can take

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

  so that $A\mathbf{v}_1 = \mathbf{v}_1$.

Corollary 7.10.5

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

$$V_n(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) \quad = \quad |\det(A)| \; V_n(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

- **Remark:** For $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ $(m \neq n)$ and $A \in \mathbb{R}^{n \times n}$, the formula from Corollary 7.10.5 fails, i.e.

$$V_m(A\mathbf{v}_1, \ldots, A\mathbf{v}_m) \quad \not\asymp \quad |\det(A)| \; V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m).$$

  - For instance, for $m = 1$ and $n = 2$, we can take

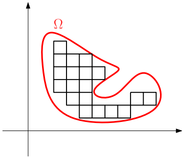
$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

    so that $A\mathbf{v}_1 = \mathbf{v}_1$.
  - Then
    - $V_1(A\mathbf{v}_1) = V_1(\mathbf{v}_1) = ||\mathbf{v}_1|| = 1$,
    - $\det(A) = 0$,

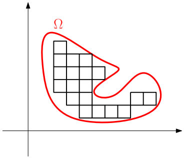
    and so $V_1(A\mathbf{v}_1) \neq |\det(A)| \; V_1(\mathbf{v}_1)$.

- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.

- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
    - We will not go into the technical details of how this can be done, but the idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes (which are simply n-parallelepipeds, and so we know how to compute their n-volume) will give us an ever better approximation of the n-volume of Ω that we wish to define.

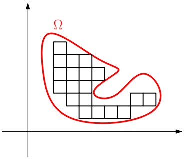


- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
    - We will not go into the technical details of how this can be done, but the idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes (which are simply n-parallelepipeds, and so we know how to compute their n-volume) will give us an ever better approximation of the n-volume of Ω that we wish to define.



    - To obtain the actual n-volume of Ω, we take the limit of these ever-finer approximations. If the limit exists, then Ω will have an n-volume (defined to be this limit). If the limit does not exist, then n-volume is undefined for Ω.

- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
    - We will not go into the technical details of how this can be done, but the idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes (which are simply n-parallelepipeds, and so we know how to compute their n-volume) will give us an ever better approximation of the n-volume of Ω that we wish to define.

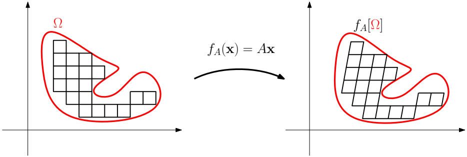


    - To obtain the actual n-volume of Ω, we take the limit of these ever-finer approximations. If the limit exists, then Ω will have an n-volume (defined to be this limit). If the limit does not exist, then n-volume is undefined for Ω.
    - It is actually pretty difficult to construct Ω for which volume is undefined! Any reasonably pretty object Ω will have a volume, although that volume may possibly be zero.

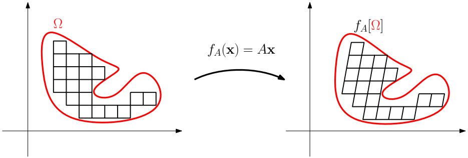
- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.

- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.
- We consider the linear function $f_A : \mathbb{R}^n \to \mathbb{R}^n$ whose standard matrix is A (i.e. for all $\mathbf{x} \in \mathbb{R}^n$, we have $f_A(\mathbf{x}) = A\mathbf{x}$).

- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.
- We consider the linear function $f_A : \mathbb{R}^n \to \mathbb{R}^n$ whose standard matrix is A (i.e. for all $\mathbf{x} \in \mathbb{R}^n$, we have $f_A(\mathbf{x}) = A\mathbf{x}$).
- Then each of the small n-hypercubes gets mapped onto a small n-parallelepiped; if the small n-hypercubes each had volume V, then by Corollary 7.10.5, the small n-parallelepipeds that these n-hypercubes get mapped onto via f_A will have volume $|\det(A)|\ V$.



- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.
- We consider the linear function $f_A : \mathbb{R}^n \to \mathbb{R}^n$ whose standard matrix is A (i.e. for all $\mathbf{x} \in \mathbb{R}^n$, we have $f_A(\mathbf{x}) = A\mathbf{x}$).
- Then each of the small n-hypercubes gets mapped onto a small n-parallelepiped; if the small n-hypercubes each had volume V, then by Corollary 7.10.5, the small n-parallelepipeds that these n-hypercubes get mapped onto via f_A will have volume $|\det(A)|\ V$.



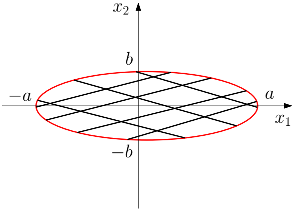
- So, we get the following formula for the n-volume of the image of Ω under f_A:

$$V_n(f_A[\Omega]) \quad = \quad |\det(A)|\ V_n(\Omega).$$

Example 7.10.6

Let a and b be positive real numbers. Compute the area (i.e. 2-volume) of the region bounded by the ellipse whose equation is

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1.$$



Solution. We need compute the area of the region

$$E \;:=\; \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \le 1 \right\}.$$

Solution. We need compute the area of the region

$$E \quad := \quad \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ \tfrac{x_1^2}{a^2} + \tfrac{x_2^2}{b^2} \leq 1 \right\}.$$

Consider the unit disk

$$D \quad := \quad \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ x_1^2 + x_2^2 \leq 1 \right\}$$

and the matrix

$$A \quad = \quad \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Solution. We need compute the area of the region

$$E \quad := \quad \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ \tfrac{x_1^2}{a^2} + \tfrac{x_2^2}{b^2} \leq 1 \right\}.$$

Consider the unit disk

$$D \quad := \quad \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ x_1^2 + x_2^2 \leq 1 \right\}$$

and the matrix

$$A \quad = \quad \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Let $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear function whose standard matrix is A, so that for all $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \in \mathbb{R}^2$, we have

$$f_A\Big(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Big) \quad = \quad \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad = \quad \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix}.$$

Solution. We need compute the area of the region

$$E \;:=\; \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\; \tfrac{x_1^2}{a^2} + \tfrac{x_2^2}{b^2} \le 1 \right\}.$$

Consider the unit disk

$$D \;:=\; \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\; x_1^2 + x_2^2 \le 1 \right\}$$

and the matrix

$$A \;=\; \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

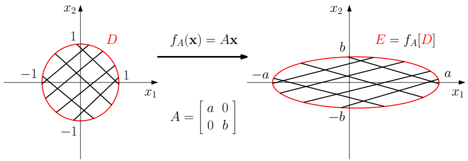
Let $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear function whose standard matrix is A, so that for all $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \in \mathbb{R}^2$, we have

$$f_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \;=\; \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;=\; \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix}.$$

WTA $f_A[D] = E$.

Solution (continued). We now see that

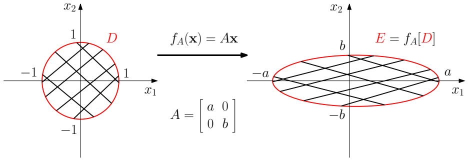
$$
\begin{aligned}
f_A[D] &= \left\{ f_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) \mid x_1, x_2 \in \mathbb{R},\ x_1^2 + x_2^2 \leq 1 \right\} \\
&= \left\{ \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ x_1^2 + x_2^2 \leq 1 \right\} \\
&= \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \mid y_1, y_2 \in \mathbb{R},\ (\tfrac{y_1}{a})^2 + (\tfrac{y_2}{b})^2 \leq 1 \right\} \\
&= \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \mid y_1, y_2 \in \mathbb{R},\ \tfrac{y_1^2}{a^2} + \tfrac{y_2^2}{b^2} \leq 1 \right\} \\
&= E.
\end{aligned}
$$



Example 7.10.6

Let a and b be positive real numbers. Compute the area (i.e. 2-volume) of the region bounded by the ellipse whose equation is

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1.$$

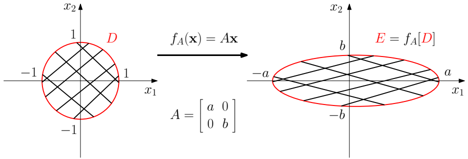


Solution (continued). Reminder: $f_A[D] = E$.

Example 7.10.6

Let a and b be positive real numbers. Compute the area (i.e. 2-volume) of the region bounded by the ellipse whose equation is

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1.$$



Solution (continued). Reminder: $f_A[D] = E$.

Therefore, the area of E is

$$\text{area}(E) = \underbrace{|\det(A)|}_{=ab} \underbrace{\text{area}(D)}_{=1^2\pi} = ab\pi.$$

□

2. Eigenvalues and eigenvectors of linear functions and square matrices

2. Eigenvalues and eigenvectors of linear functions and square matrices

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

2. Eigenvalues and eigenvectors of linear functions and square matrices

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

- So, the eigenvectors of f are those **non-zero** vectors in V that simply get scaled by f, and the eigenvalues are the scalars that the eigenvectors get scaled by.

2. Eigenvalues and eigenvectors of linear functions and square matrices

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

- So, the eigenvectors of f are those **non-zero** vectors in V that simply get scaled by f, and the eigenvalues are the scalars that the eigenvectors get scaled by.
- By definition, an eigenvector cannot be $\mathbf{0}$, but an eigenvalue may possibly be 0.

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

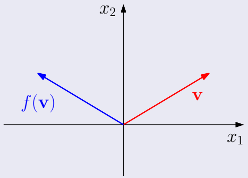
- **Remark:** Note that eigenvectors and eigenvalues are only defined for those linear functions whose domain is the same as the codomain.

Example 8.1.1

Consider the linear function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

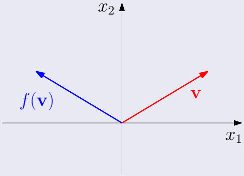
$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{R}$. So, f is the reflection about the x_2-axis (see the picture below), and its standard matrix is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

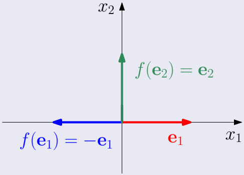


As usual, $\mathbf{e}_1$ and $\mathbf{e}_2$ are the standard basis vectors of $\mathbb{R}^2$. Then (next slide)

Example 8.1.1



- $\mathbf{e}_1$ is an eigenvector of f associated with the eigenvalue $\lambda_1 := -1$, since $f(\mathbf{e}_1) = -\mathbf{e}_1 = \lambda_1 \mathbf{e}_1$;
- $\mathbf{e}_2$ is an eigenvector of f associated with the eigenvalue $\lambda_2 := 1$, since $f(\mathbf{e}_2) = \mathbf{e}_2 = \lambda_2 \mathbf{e}_2$.

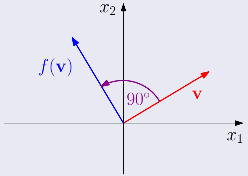


Example 8.1.2

Consider the linear function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{R}$. So, f is the counterclockwise rotation by 90° about the origin (see the picture below), and its standard matrix is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. This function has no eigenvectors (and consequently, it has no eigenvalues), since it does not simply scale any non-zero vector in $\mathbb{R}^2$.



Example 8.1.3

Consider the linear function $f : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. (This is the same formula as the one from Example 8.1.2, except that we are now working over $\mathbb{C}$, rather than over $\mathbb{R}$.) Then

Example 8.1.3

Consider the linear function $f : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. (This is the same formula as the one from Example 8.1.2, except that we are now working over $\mathbb{C}$, rather than over $\mathbb{R}$.) Then

- $\mathbf{v}_1 := \begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvector of f associated with the eigenvalue $\lambda_1 := i$, since $f(\mathbf{v}_1) = \begin{bmatrix} -1 \\ i \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix} = \lambda_1 \mathbf{v}_1$;

Example 8.1.3

Consider the linear function $f : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. (This is the same formula as the one from Example 8.1.2, except that we are now working over $\mathbb{C}$, rather than over $\mathbb{R}$.) Then

- $\mathbf{v}_1 := \begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvector of f associated with the eigenvalue $\lambda_1 := i$, since $f(\mathbf{v}_1) = \begin{bmatrix} -1 \\ i \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix} = \lambda_1 \mathbf{v}_1$;
- $\mathbf{v}_2 := \begin{bmatrix} -i \\ 1 \end{bmatrix}$ is an eigenvector of f associated with the eigenvalue $\lambda_2 := -i$, since $f(\mathbf{v}_2) = \begin{bmatrix} -1 \\ -i \end{bmatrix} = (-i)\begin{bmatrix} -i \\ 1 \end{bmatrix} = \lambda_2 \mathbf{v}_2$.

- Example 8.1.2: $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f\left(\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]\right) = \left[\begin{array}{rr} 0 & -1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{r} -x_2 \\ x_1 \end{array}\right] \qquad \forall x_1, x_2 \in \mathbb{R}$$

(counterclockwise rotation by 90° about the origin);

- Example 8.1.3: $f : \mathbb{C}^2 \to \mathbb{C}^2$,

$$f\left(\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]\right) = \left[\begin{array}{rr} 0 & -1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{r} -x_2 \\ x_1 \end{array}\right] \qquad \forall x_1, x_2 \in \mathbb{C}.$$

- Example 8.1.2: $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{R}$$

  (counterclockwise rotation by 90° about the origin);
- Example 8.1.3: $f : \mathbb{C}^2 \to \mathbb{C}^2$,

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{C}.$$

- **Remark:** It may be somewhat surprising that the linear function f from Example 8.1.2 has no eigenvectors and no eigenvalues, whereas the one from Example 8.1.3 has them.

- Example 8.1.2: $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{R}$$

  (counterclockwise rotation by 90° about the origin);
- Example 8.1.3: $f : \mathbb{C}^2 \to \mathbb{C}^2$,

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{C}.$$

- **Remark:** It may be somewhat surprising that the linear function f from Example 8.1.2 has no eigenvectors and no eigenvalues, whereas the one from Example 8.1.3 has them.
- As we shall see once we learn how to actually compute eigenvalues and eigenvectors (this will involve finding roots of polynomials), the essential difference is that $\mathbb{C}$ is an algebraically closed field, whereas $\mathbb{R}$ is not.

- Reminder:

Definition

An *algebraically closed field* is a field $\mathbb{F}$ that has the property that every non-constant polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

- Reminder:

Definition

An *algebraically closed field* is a field $\mathbb{F}$ that has the property that every non-constant polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

- If $\mathbb{F}$ is an algebraically closed field, and $p(x)$ is non-constant polynomial with coefficients in $\mathbb{F}$, then $p(x)$ can be factored into linear terms.

- Reminder:

Definition

An *algebraically closed field* is a field $\mathbb{F}$ that has the property that every non-constant polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

- If $\mathbb{F}$ is an algebraically closed field, and $p(x)$ is non-constant polynomial with coefficients in $\mathbb{F}$, then $p(x)$ can be factored into linear terms.
- $\mathbb{C}$ is algebraically closed.
- $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{Z}_p$ (where p is a prime number) are **not** algebraically closed.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda \mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda\mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda\mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.
- Note that, for an eigenvalue λ of f, the elements of the eigenspace $E_\lambda(f)$ are precisely the zero vector and the eigenvectors of f associated with λ.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda\mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.
- Note that, for an eigenvalue λ of f, the elements of the eigenspace $E_\lambda(f)$ are precisely the zero vector and the eigenvectors of f associated with λ.
  - By definition, $\mathbf{0}$ cannot be an eigenvector.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda \mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.
- Note that, for an eigenvalue λ of f, the elements of the eigenspace $E_\lambda(f)$ are precisely the zero vector and the eigenvectors of f associated with λ.
    - By definition, $\mathbf{0}$ cannot be an eigenvector.
- On the other hand, if λ is not an eigenvalue of f, then we simply have that $E_\lambda(f) = \{\mathbf{0}\}$, and we do not refer to $E_\lambda(f)$ as an eigenspace.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline). (a) For $\lambda \in \mathbb{F}$:

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline). (a) For $\lambda \in \mathbb{F}$:

- we check that $E_\lambda(f)$ contains $\mathbf{0}$ and is closed under vector addition and scalar multiplication, and we deduce (by Theorem 3.1.7) that $E_\lambda(f)$ is a subspace of V;

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

- (a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;
- (b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline). (a) For $\lambda \in \mathbb{F}$:

- we check that $E_\lambda(f)$ contains $\mathbf{0}$ and is closed under vector addition and scalar multiplication, and we deduce (by Theorem 3.1.7) that $E_\lambda(f)$ is a subspace of V;
- any non-zero vector in $E_\lambda(f)$ is an eigenvector of f associated with λ, and so $E_\lambda(f)$ is non-trivial iff λ is an eigenvalue of f.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline, continued). (b) Fix distinct $\lambda_1, \lambda_2 \in \mathbb{F}$.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

- (a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;
- (b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline, continued). (b) Fix distinct $\lambda_1, \lambda_2 \in \mathbb{F}$. Obviously, $\mathbf{0} \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline, continued). (b) Fix distinct $\lambda_1, \lambda_2 \in \mathbb{F}$. Obviously, $\mathbf{0} \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$. On the other hand, for $\mathbf{v} \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$:

$$\begin{array}{l} f(\mathbf{v}) = \lambda_1 \mathbf{v} \text{ (because } \mathbf{v} \in E_{\lambda_1}(f)) \text{ and} \\ f(\mathbf{v}) = \lambda_2 \mathbf{v} \text{ (because } \mathbf{v} \in E_{\lambda_2}(f)) \end{array}$$

$$\Longrightarrow \quad \lambda_1 \mathbf{v} = \lambda_2 \mathbf{v} \implies \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \mathbf{v} = \mathbf{0} \implies \mathbf{v} = \mathbf{0}.$$

So, $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$. $\square$

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

- **Terminology:** Suppose that V is a vector space over a field $\mathbb{F}$, and that λ is an eigenvalue of a linear function $f : V \to V$.
  - The *geometric multiplicity* of the eigenvalue λ is defined to be $\dim\big(E_\lambda(f)\big)$.
  - So, the geometric multiplicity of an eigenvalue is the dimension of the associated eigenspace.

Definition

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a **square** matrix. An *eigenvector* of A is a vector $\mathbf{v} \in \mathbb{F}^n \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of A associated with the eigenvector $\mathbf{v}$, s.t.

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

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of A associated with the eigenvalue λ.

- Eigenvectors are, by definition, non-zero, whereas eigenvalues may possibly be zero.

- For a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is some field), and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(A) \quad := \quad \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda \mathbf{v}\}.$$

If λ is an eigenvalue of A, then $E_\lambda(A)$ is called the *eigenspace* of A associated with the eigenvalue λ.

- For a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is some field), and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(A) \quad := \quad \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

  If λ is an eigenvalue of A, then $E_\lambda(A)$ is called the *eigenspace* of A associated with the eigenvalue λ.
    - Note that, for an eigenvalue λ of A, the elements of the eigenspace $E_\lambda(A)$ are precisely the zero vector and the eigenvectors of A associated with λ.

- For a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is some field), and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(A) \quad := \quad \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

  If λ is an eigenvalue of A, then $E_\lambda(A)$ is called the *eigenspace* of A associated with the eigenvalue λ.
    - Note that, for an eigenvalue λ of A, the elements of the eigenspace $E_\lambda(A)$ are precisely the zero vector and the eigenvectors of A associated with λ.
    - On the other hand, if λ is not an eigenvalue of A, then we simply have that $E_\lambda(A) = \{\mathbf{0}\}$, and we do not refer to $E_\lambda(A)$ as an eigenspace.

Proposition 8.1.5

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^n \to \mathbb{F}^n$ be a linear function, and let A be the standard matrix of f. Then f and A have exactly the same eigenvalues and the associated eigenectors. Moreover, for all eigenvalues λ of f and A, we have that $E_\lambda(f) = E_\lambda(A)$.

Proof. This follows immediately from the appropriate definitions. □

- Reminder:

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

- Reminder:

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

- For square matrices, we have the following analog of Proposition 8.1.4.

Proposition 8.1.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then all the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of A;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$.

Proof.

Proposition 8.1.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then all the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of A;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$.

Proof. Consider the function $f_A : \mathbb{F}^n \to \mathbb{F}^n$, given by $f_A(\mathbf{v}) = A\mathbf{v}$ for all vectors $\mathbf{v} \in \mathbb{F}^n$. Then f_A is linear (by Proposition 1.10.4), and moreover, A is the standard matrix of f_A.

So, by Proposition 8.1.5, we have that for all $\lambda \in \mathbb{F}$, $E_\lambda(A) = E_\lambda(f_A)$.

The result now follows immediately from Proposition 8.1.4. □

Proposition 8.1.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then all the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of A;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$.

- **Terminology:** Suppose that $\mathbb{F}$ is a field, and that λ is an eigenvalue of a square matrix $A \in \mathbb{F}^{n \times n}$.
  - The *geometric multiplicity* of the eigenvalue λ is defined to be $\dim\big(E_\lambda(A)\big)$.
  - So, the geometric multiplicity of an eigenvalue is the dimension of the associated eigenspace.

Proposition 8.1.7

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then for all $\lambda \in \mathbb{F}$, we have that

$$E_\lambda\Big({}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \Big) \;=\; \Big\{ \big[\ \mathbf{v}\ \big]_{\mathcal{B}} \mid \mathbf{v} \in E_\lambda(f) \Big\}.$$

Consequently, the linear function f and the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ have exactly the same eigenvalues, with exactly the same corresponding geometric multiplicities.

- Proof: Lecture Notes.
- Proposition 8.1.7 states that $E_\lambda\Big({}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \Big)$ is the image of $E_\lambda(f)$ under the coordinate transformation $\big[\ \cdot\ \big]_{\mathcal{B}}$.

- In view of Propositions 8.1.5 ("linear functions and their standard matrices have the same eigenvalues, eigenvectors, and eigenspaces") and 8.1.7 (previous slide), we see that the study of eigenvalues and eigenvectors of linear functions from a non-trivial, finite-dimensional vector space to itself is essentially equivalent to the study of eigenvalues and eigenvectors of square matrices.
  - The computational tools that we develop for finding eigenvectors and eigenvalues will primarily be for square matrices.
  - On the other hand, some of the theoretical results that we prove will be for linear functions instead, and we will obtain corresponding results for matrices as more or less immediate corollaries.

3. The characteristic polynomial and spectrum

Definition

Given a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the *characteristic polynomial* of A is defined to be

$$p_A(\lambda) \;\; := \;\; \det(\lambda I_n - A).$$

The *characteristic equation* of A is the equation

$$\det(\lambda I_n - A) \;\; = \;\; 0.$$

So, the roots of the characteristic polynomial of A are precisely the solutions of the characteristic equation of A.

Example 8.2.1

Compute the characteristic polynomial of the following matrix in $\mathbb{C}^{3\times 3}$:

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

Solution.

Example 8.2.1

Compute the characteristic polynomial of the following matrix in $\mathbb{C}^{3\times 3}$:

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

Solution. The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) = \begin{vmatrix} \lambda - 1 & 2 & -3 \\ 1 & \lambda & -2 \\ -2 & 1 & \lambda + 3 \end{vmatrix} \\ &= \lambda^3 + 2\lambda^2 - 9\lambda - 3. \end{aligned}$$

□

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.
    - In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$.

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.
  - In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$.
  - By Proposition 7.2.3, we have that $\det(A - \lambda I_n) = (-1)^n \det(\lambda I_n - A)$, and so the polynomials $\det(\lambda I_n - A)$ and $\det(A - \lambda I_n)$ have exactly the same roots, with the same corresponding multiplicities, which is what we will actually care about when it comes to the characteristic polynomial.

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.
    - In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$.
    - By Proposition 7.2.3, we have that $\det(A - \lambda I_n) = (-1)^n \det(\lambda I_n - A)$, and so the polynomials $\det(\lambda I_n - A)$ and $\det(A - \lambda I_n)$ have exactly the same roots, with the same corresponding multiplicities, which is what we will actually care about when it comes to the characteristic polynomial.
    - The main advantage of using $\det(\lambda I_n - A)$ rather than $\det(A - \lambda I_n)$ is that the former polynomial has leading coefficient 1, whereas the latter has leading coefficient $(-1)^n$, which is -1 if n is odd.

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \mathrm{Nul}(\lambda_0 I_n - A) \;=\; \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof.

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \mathrm{Nul}(\lambda_0 I_n - A) \;=\; \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof. Obviously, for all $\mathbf{v} \in \mathbb{F}^n$, we have that $(\lambda_0 I_n - A)\mathbf{v} = \mathbf{0}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0}$. So, $\mathrm{Nul}(\lambda_0 I_n - A) = \mathrm{Nul}(A - \lambda_0 I_n)$.

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \quad = \quad \text{Nul}(\lambda_0 I_n - A) \quad = \quad \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). Further, we compute:

$$\begin{aligned} E_{\lambda_0}(A) &= \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda_0\mathbf{v}\} \\ &= \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda_0 I_n\mathbf{v}\} \\ &= \{\mathbf{v} \in \mathbb{F}^n \mid (\lambda_0 I_n - A)\mathbf{v} = \mathbf{0}\} \\ &= \text{Nul}(\lambda_0 I_n - A). \end{aligned}$$

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). We have now shown that

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). We have now shown that

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

It remains to show that (1), (2), and (3) are equivalent. The fact that (2) and (3) are equivalent follows immediately from the appropriate definitions. It remains to prove that (1) and (3) are equivalent.

Theorem 8.2.2

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). Reminder: $E_{\lambda_0}(A) = \text{Nul}(\lambda_0 I_n - A)$.

$$\underbrace{\lambda_0 \text{ is an eigenvalue of } A}_{(1)} \overset{\text{Prop. 8.1.6}}{\iff} E_{\lambda_0}(A) \neq \{\mathbf{0}\}$$

$$\iff \underbrace{\text{Nul}(\lambda_0 I_n - A)}_{=E_{\lambda_0}(A)} \neq \{\mathbf{0}\}$$

$$\overset{\text{IMT}}{\iff} \text{the matrix } \lambda_0 I_n - A \text{ is not invertible}$$

$$\overset{\text{IMT}}{\iff} \underbrace{\det(\lambda_0 I_n - A) = 0}_{(3)}$$

□

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \mathrm{Nul}(\lambda_0 I_n - A) \;=\; \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.
- Since $\deg(p_A(\lambda)) = n$, the sum of algebraic multiplicities of the eigenvalues of the matrix $A \in \mathbb{F}^{n \times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of the eigenvalues of A is exactly n.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.
- Since $\deg(p_A(\lambda)) = n$, the sum of algebraic multiplicities of the eigenvalues of the matrix $A \in \mathbb{F}^{n \times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of the eigenvalues of A is exactly n.
  - Indeed, if $\mathbb{F}$ is algebraically closed, then the characteristic polynomial $p_A(\lambda)$ can be written as a product of linear factors, and there are n of those factors.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.
- Since $\deg(p_A(\lambda)) = n$, the sum of algebraic multiplicities of the eigenvalues of the matrix $A \in \mathbb{F}^{n \times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of the eigenvalues of A is exactly n.
  - Indeed, if $\mathbb{F}$ is algebraically closed, then the characteristic polynomial $p_A(\lambda)$ can be written as a product of linear factors, and there are n of those factors.
  - If $\mathbb{F}$ is not algebraically closed, we might or might not be able to factor $p_A(\lambda)$ in this way, which is why the sum of algebraic multiplicities of the eigenvalues of A is at most n (possibly strictly smaller than n).

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

- Proof: Later!
- Schematically, Theorem 8.2.3 states that for an eigenvalue λ of A:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

- Proof: Later!
- Schematically, Theorem 8.2.3 states that for an eigenvalue λ of A:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

- For now, we have only stated Theorem 8.2.3. We will not use this theorem before proving it.

- The *spectrum* of a square matrix $A \in \mathbb{F}^{n \times n}$ is the multiset of all eigenvalues of A, with algebraic multiplicities taken into account.
    - This means that the number of times that an eigenvalue appears in the spectrum is equal to the algebraic multiplicity of that eigenvalue. The order in which we list the eigenvalues in the spectrum does not matter, but repetitions do matter.

- The *spectrum* of a square matrix $A \in \mathbb{F}^{n \times n}$ is the multiset of all eigenvalues of A, with algebraic multiplicities taken into account.
  - This means that the number of times that an eigenvalue appears in the spectrum is equal to the algebraic multiplicity of that eigenvalue. The order in which we list the eigenvalues in the spectrum does not matter, but repetitions do matter.
- For example, if a matrix $A \in \mathbb{C}^{5 \times 5}$ has eigenvalues 1 (with algeraic multiplicity 1), $1 + i$ (with algebraic multiplicity 2), and $1 - i$ (with algebraic multiplicity 2), then the spectrum of A is $\{1, 1+i, 1+i, 1-i, 1-i\}$.

- The *spectrum* of a square matrix $A \in \mathbb{F}^{n \times n}$ is the multiset of all eigenvalues of A, with algebraic multiplicities taken into account.
  - This means that the number of times that an eigenvalue appears in the spectrum is equal to the algebraic multiplicity of that eigenvalue. The order in which we list the eigenvalues in the spectrum does not matter, but repetitions do matter.
- For example, if a matrix $A \in \mathbb{C}^{5 \times 5}$ has eigenvalues 1 (with algeraic multiplicity 1), $1 + i$ (with algebraic multiplicity 2), and $1 - i$ (with algebraic multiplicity 2), then the spectrum of A is $\{1, 1+i, 1+i, 1-i, 1-i\}$.
- In general, the spectrum of a matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field) has at most n elements; if the field $\mathbb{F}$ is algebraically closed, then the spectrum of A has exactly n elements.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

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

Solution. (a) The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda - 4 & 0 & 2 \\ -2 & \lambda - 5 & -4 \\ 0 & 0 & \lambda - 5 \end{vmatrix} \\ &= (\lambda - 4)(\lambda - 5)^2 \qquad \text{via Laplace expansion along 2nd column} \\ &= \lambda^3 - 14\lambda^2 + 65\lambda - 100. \end{aligned}$$

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

Solution. (a) The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda - 4 & 0 & 2 \\ -2 & \lambda - 5 & -4 \\ 0 & 0 & \lambda - 5 \end{vmatrix} \\ &= (\lambda - 4)(\lambda - 5)^2 \qquad \text{via Laplace expansion along 2nd column} \\ &= \lambda^3 - 14\lambda^2 + 65\lambda - 100. \end{aligned}$$

- **Remark:** We did not really need to expand in the last line.
  - We only really care about the roots of the characteristic polynomial, and it is more convenient to have a form that is already factored.
  - So, $p_A(\lambda) = (\lambda - 4)(\lambda - 5)^2$ is a "better" answer than $p_A(\lambda) = \lambda^3 - 14\lambda^2 + 65\lambda - 100$, although they are both correct.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

Solution (continued). Reminder: (a) $p_A(\lambda) = (\lambda - 4)(\lambda - 5)^2$.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

a. Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.
b. Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.
c. For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

Solution (continued). Reminder: (a) $p_A(\lambda) = (\lambda - 4)(\lambda - 5)^2$.

(b) From part (a), we see that A has two eigenvalues, namely, the eigenvalue $\lambda_1 = 4$ (with algebraic multiplicity 1), and the eigenvalue $\lambda_2 = 5$ (with algebraic multiplicity 2). So, the spectrum of A is $\{4, 5, 5\}$.

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

Solution (continued). Reminder: the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 5$.

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

Solution (continued). Reminder: the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 5$.

(c) For each $i \in \{1, 2\}$, we have that

$$E_{\lambda_i}(A) \;\; = \;\; \mathrm{Nul}(\lambda_i I_3 - A),$$

which is precisely the set of all solutions of the characteristic equation

$$(\lambda_i I_3 - A)\mathbf{x} \;\; = \;\; \mathbf{0}.$$

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

Solution (continued). Reminder: the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 5$.

(c) For each $i \in \{1, 2\}$, we have that

$$E_{\lambda_i}(A) \quad = \quad \mathrm{Nul}(\lambda_i I_3 - A),$$

which is precisely the set of all solutions of the characteristic equation

$$(\lambda_i I_3 - A)\mathbf{x} \quad = \quad \mathbf{0}.$$

Let us now compute a basis of each of the two eigenspaces.

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\mathrm{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \frac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \tfrac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T \right\}$ is a basis of the eigespace $E_{\lambda_1}(A) = \text{Nul}\big(A - \lambda_1 I_n\big)$,

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \tfrac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T \right\}$ is a basis of the eigespace $E_{\lambda_1}(A) = \text{Nul}(A - \lambda_1 I_n)$, and we see that the eigenvalue $\lambda_1 = 4$ has geometric multiplicity 1.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

Solution (continued). (c) Similarly, for $\lambda_2 = 5$, we get that $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace $E_{\lambda_2}(A) = \text{Nul}(A - \lambda_2 I_n)$, and we see that the eigenvalue $\lambda_2 = 5$ has geometric multiplicity 2 (details: Lecture Notes). $\square$

- Reminder:

Proposition 7.3.1

Let $\mathbb{F}$ be a field, and let $A = \left[\ a_{i,j}\ \right]_{n \times n}$ be a triangular matrix in $\mathbb{F}^{n \times n}$. Then

$$\det(A) \quad = \quad \prod_{i=1}^{n} a_{i,i} \quad = \quad a_{1,1}a_{2,2}\dots a_{n,n},$$

that is, $\det(A)$ is equal to the product of entries on the main diagonal of A.

$$\begin{bmatrix} * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * \\ 0 & 0 & * & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & * & * \\ 0 & 0 & 0 & \dots & 0 & * \end{bmatrix}$$

upper triangular matrix

$$\begin{bmatrix} * & 0 & 0 & \dots & 0 & 0 \\ * & * & 0 & \dots & 0 & 0 \\ * & * & * & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & * & \dots & * & 0 \\ * & * & * & \dots & * & * \end{bmatrix}$$

lower triangular matrix

Proposition 8.2.7

Let $\mathbb{F}$ be a field, and let $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times n}$ be a triangular matrix in $\mathbb{F}^{n\times n}$. Then the characteristic polynomial of A is

$$p_A(\lambda) \;=\; \prod_{i=1}^{n}(\lambda - a_{i,i}) \;=\; (\lambda - a_{1,1})(\lambda - a_{2,2})\dots(\lambda - a_{n,n}),$$

the eigenvalues of A are precisely the entries of A on its main diagonal, and moreover, the algebraic multiplicity of each eigenvalue is precisely the number of times that it appears on the main diagonal of A.[a] Consequently, the spectrum of A is $\{a_{1,1}, a_{2,2}, \dots, a_{n,n}\}$, i.e. the multiset formed precisely by the main diagonal entries of A, with each number appearing in the spectrum of A the same number of times as on the main diagonal of A.

---

[a] However, the geometric multiplicity may possibly be smaller.

- For example, for the matrix

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

  in $\mathbb{C}^{5\times 5}$, we have the following:
  - the characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= (\lambda-1)(\lambda-2)(\lambda-1)(\lambda-3)(\lambda-3) \\ &= (\lambda-1)^2(\lambda-2)(\lambda-3)^2; \end{aligned}$$

  - the spectrum of A is $\{1, 1, 2, 3, 3\}$.