Linear Algebra 2

Lecture #20

Applications of determinants: volume and polynomials

Irena Penev

April 10, 2024

- This lecture has three parts:

- This lecture has three parts:
    1. Determinants and volume

- This lecture has three parts:
  1. Determinants and volume
  2. Algebraically closed fields

- This lecture has three parts:
  1. Determinants and volume
  2. Algebraically closed fields
  3. Common roots of polynomials via determinants

1. Determinants and volume

1. Determinants and volume

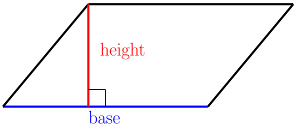
- In our study of determinants and volume, we assume throughout that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

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 $||\cdot||$.

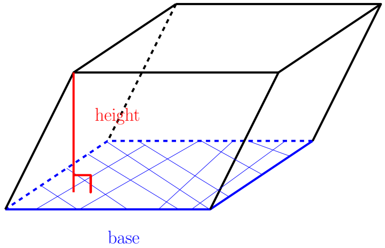
- For a parallelogram, we have the familiar formula

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



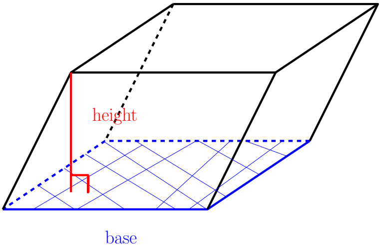
- We have a similar formula for the volume of a parallelepiped:

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



- 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

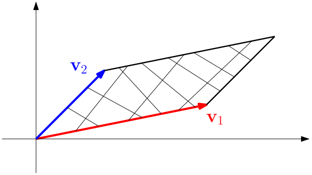
$$\left\{ c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \;\middle|\; c_1, \ldots, c_m \in \mathbb{R},\ 0 \leq c_1, \ldots, c_m \leq 1 \right\}.$$

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

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

- 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.

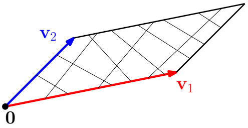


Definition

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

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

- 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

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

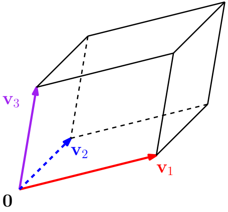
  which is 1-dimensional (a line segment) if $\mathbf{v}_1 \neq \mathbf{0}$, and is 0-dimensional (containing only the zero vector) if $\mathbf{v}_1 = \mathbf{0}$.
- We can think of these as "degenerate parallelograms."

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 $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.
  - 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

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

- 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."

- 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

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

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

Definition

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

$$\Big\{c_1\mathbf{v}_1 + \dots + c_m\mathbf{v}_m \ \Big| \ c_1, \dots, c_m \in \mathbb{R},\ 0 \leq c_1, \dots, c_m \leq 1\Big\}.$$

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

---

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

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

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

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

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

- 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

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

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

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

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

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

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

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

Proof. $\forall i \in \{1, \dots, m\}$: $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$. We will prove inductively that $\forall i \in \{1, \dots, m\}$: $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. 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. Fix $i \in \{1, \dots, m-1\}$, and assume inductively that $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. 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: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
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: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
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: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$.

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

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

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

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

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

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

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

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.

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

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

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

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

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

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

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.

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

Then (next slide):

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$;
$\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \cdots - c_i\mathbf{a}_i$.

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

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$;
$\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \cdots - c_i\mathbf{a}_i$.

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

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

- $R_{i+1} \to R_{i+1} - c_1 R_1$;

  $\vdots$
- $R_{i+1} \to R_{i+1} - c_i R_i$.

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

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$;
$\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \cdots - c_i\mathbf{a}_i$.

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

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

- $R_{i+1} \to R_{i+1} - c_1 R_1$;

  $\vdots$
- $R_{i+1} \to R_{i+1} - c_i R_i$.

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

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$;
$\mathbf{a}_{i+1}^{\perp} = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \cdots - c_i\mathbf{a}_i$.

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

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

- $R_{i+1} \to R_{i+1} - c_1 R_1$;

  $\vdots$

- $R_{i+1} \to R_{i+1} - c_i R_i$.

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

Proof (continued). Reminder: $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$;
WTS $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$;
$\mathbf{a}_{i+1}^\perp = \mathbf{a}_{i+1} - c_1\mathbf{a}_1 - \cdots - c_i\mathbf{a}_i$.

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

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

- $R_{i+1} \to R_{i+1} - c_1 R_1$;

  $\vdots$
- $R_{i+1} \to R_{i+1} - c_i R_i$.

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

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

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

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

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

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

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

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

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 $(\mathbf{a}_{i+1}^\perp)^T \mathbf{a}_{i+1}^\perp = \mathbf{a}_{i+1}^\perp \cdot \mathbf{a}_{i+1}^\perp = ||\mathbf{a}_{i+1}^\perp||$.

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

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

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 $(\mathbf{a}_{i+1}^\perp)^T \mathbf{a}_{i+1}^\perp = \mathbf{a}_{i+1}^\perp \cdot \mathbf{a}_{i+1}^\perp = ||\mathbf{a}_{i+1}^\perp||$.

We now compute (next slide):

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

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

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

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

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

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:

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

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

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

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

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

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

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

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

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

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.

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).

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

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

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

- **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.

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

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

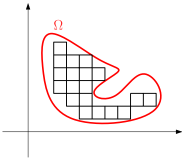
$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

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

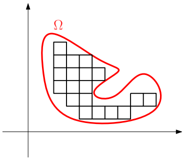
    and so $V_1(A\mathbf{v}_1) \neq |\det(A)| \; V_1(\mathbf{v}_1)$.

- 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.

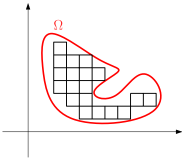


- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
  - We will not go into the technical details of how this can be done, but the idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes (which 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.

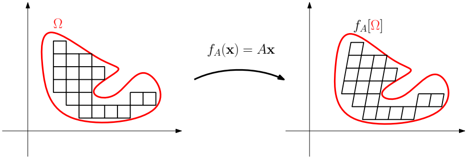


  - To obtain the actual n-volume of Ω, we take the limit of these ever-finer approximations. If the limit exists, then Ω will have an n-volume (defined to be this limit). If the limit does not exist, then n-volume is undefined for Ω.
  - It is actually pretty difficult to construct Ω for which volume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

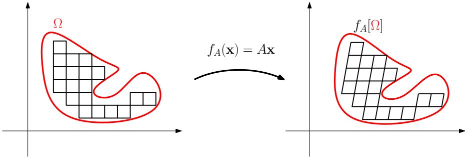
- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.

- 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 $|\det(A)|\ V$.



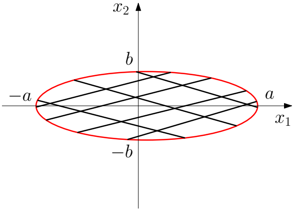
- So, we get the following formula for the n-volume of the image of Ω under f_A:

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

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

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

Solution. We need compute the area of the region

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

Consider the unit disk

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

and the matrix

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

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

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

Solution. We need compute the area of the region

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

Consider the unit disk

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

and the matrix

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

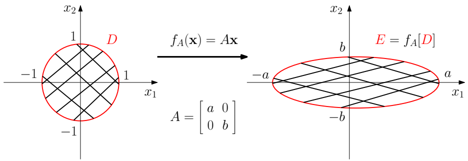
Let $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear function whose standard matrix is A, so that for all $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \in \mathbb{R}^2$, we have

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

WTA $f_A[D] = E$.

Solution (continued). We now see that

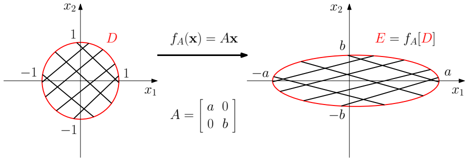
$$
\begin{aligned}
f_A[D] &= \left\{ f_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) \mid x_1, x_2 \in \mathbb{R},\ x_1^2 + x_2^2 \leq 1 \right\} \\
&= \left\{ \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R},\ x_1^2 + x_2^2 \leq 1 \right\} \\
&= \left\{ \begin{bmatrix}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



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

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

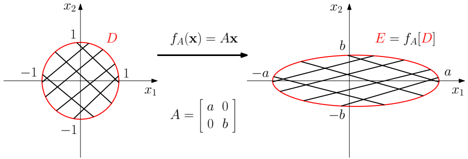


Solution (continued). Reminder: $f_A[D] = E$.

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

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

$\square$

2. Algebraically closed fields (subsec. 2.4.5 of the Lecture Notes)

2. Algebraically closed fields (subsec. 2.4.5 of the Lecture Notes)

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 $\mathbb{F}$.

2. Algebraically closed fields (subsec. 2.4.5 of the Lecture Notes)

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 $\mathbb{F}$.

The Fundamental Theorem of Algebra

Any non-constant polynomial with complex coefficients has a complex root.

- By the Fundamental Theorem of Algebra, the field $\mathbb{C}$ is algebraically closed.

2. Algebraically closed fields (subsec. 2.4.5 of the Lecture Notes)

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 $\mathbb{F}$.

The Fundamental Theorem of Algebra

Any non-constant polynomial with complex coefficients has a complex root.

- By the Fundamental Theorem of Algebra, the field $\mathbb{C}$ is algebraically closed.
- On the other hand, $\mathbb{R}$ is not algebraically closed, and similarly, neither is $\mathbb{Q}$.
    - For example, the polynomial $x^2 + 1$ has no roots in $\mathbb{R}$ (and in particular, it has no roots in $\mathbb{Q}$).
    - It does, however, have two complex roots, namely, i and $-i$.

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 $\mathbb{F}$.

- No finite field is algebraically closed.

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 $\mathbb{F}$.

- No finite field is algebraically closed.
  - To see this, consider any finite field $\mathbb{F} = \{f_1, \ldots, f_t\}$ ($t \geq 2$), and consider the polynomial

$$p(x) = (x - f_1) \ldots (x - f_t) + 1,$$

which is a polynomial of degree t with coefficients in $\mathbb{F}$.

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 $\mathbb{F}$.

- No finite field is algebraically closed.
  - To see this, consider any finite field $\mathbb{F} = \{f_1, \ldots, f_t\}$ $(t \geq 2)$, and consider the polynomial

$$p(x) \quad = \quad (x - f_1) \ldots (x - f_t) + 1,$$

    which is a polynomial of degree t with coefficients in $\mathbb{F}$.
  - Then for each $i \in \{1, \ldots, t\}$, we have that $p(f_i) = 1$, and consequently, f_i is not a root of $p(x)$.

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 $\mathbb{F}$.

- No finite field is algebraically closed.
  - To see this, consider any finite field $\mathbb{F} = \{f_1, \dots, f_t\}$ ($t \geq 2$), and consider the polynomial

    $$p(x) \quad = \quad (x - f_1) \dots (x - f_t) + 1,$$

    which is a polynomial of degree t with coefficients in $\mathbb{F}$.
  - Then for each $i \in \{1, \dots, t\}$, we have that $p(f_i) = 1$, and consequently, f_i is not a root of $p(x)$.
  - Since $\mathbb{F} = \{f_1, \dots, f_t\}$, we see that $p(x)$ has no roots in $\mathbb{F}$.

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 $\mathbb{F}$.

- No finite field is algebraically closed.
  - To see this, consider any finite field $\mathbb{F} = \{f_1, \dots, f_t\}$ ($t \geq 2$), and consider the polynomial

    $$p(x) \quad = \quad (x - f_1) \dots (x - f_t) + 1,$$

    which is a polynomial of degree t with coefficients in $\mathbb{F}$.
  - Then for each $i \in \{1, \dots, t\}$, we have that $p(f_i) = 1$, and consequently, f_i is not a root of $p(x)$.
  - Since $\mathbb{F} = \{f_1, \dots, f_t\}$, we see that $p(x)$ has no roots in $\mathbb{F}$.
- Thus, of the fields that we have seen so far, namely, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_p$ (where p is a prime number), only the field $\mathbb{C}$ is algebraically closed.

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 $\mathbb{F}$.

- No finite field is algebraically closed.
  - To see this, consider any finite field $\mathbb{F} = \{f_1, \dots, f_t\}$ $(t \geq 2)$, and consider the polynomial

    $$p(x) \quad = \quad (x - f_1) \dots (x - f_t) + 1,$$

    which is a polynomial of degree t with coefficients in $\mathbb{F}$.
  - Then for each $i \in \{1, \dots, t\}$, we have that $p(f_i) = 1$, and consequently, f_i is not a root of $p(x)$.
  - Since $\mathbb{F} = \{f_1, \dots, f_t\}$, we see that $p(x)$ has no roots in $\mathbb{F}$.
- Thus, of the fields that we have seen so far, namely, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_p$ (where p is a prime number), only the field $\mathbb{C}$ is algebraically closed.
- Other algebraically closed fields do exist, but we will not study them in this course.

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 $\mathbb{F}$.

- It can be shown (though we will not give a formal proof) that any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ can be factored into linear terms in a unique way.

- More precisely, if $p(x)$ is a polynomial of degree $n \geq 1$, and with coefficients in an algebraically closed field $\mathbb{F}$, then there exist numbers $a, \alpha_1, \ldots, \alpha_\ell$ in $\mathbb{F}$ s.t. $a \neq 0$ and s.t. $\alpha_1, \ldots, \alpha_\ell$ are pairwise distinct, and positive integers $n_1, \ldots, n_\ell$ satisfying $n_1 + \cdots + n_\ell = n$, s.t.

$$p(x) = a(x - \alpha_1)^{n_1} \ldots (x - \alpha_\ell)^{n_\ell}.$$

  Moreover, $a, \alpha_1, \ldots, \alpha_\ell, n_1, \ldots, n_\ell$ are uniquely determined by the polynomial $p(x)$, up to a permutation of the α_i's and the corresponding n_i's.

- More precisely, if $p(x)$ is a polynomial of degree $n \geq 1$, and with coefficients in an algebraically closed field $\mathbb{F}$, then there exist numbers $a, \alpha_1, \dots, \alpha_\ell$ in $\mathbb{F}$ s.t. $a \neq 0$ and s.t. $\alpha_1, \dots, \alpha_\ell$ are pairwise distinct, and positive integers $n_1, \dots, n_\ell$ satisfying $n_1 + \dots + n_\ell = n$, s.t.

$$p(x) = a(x - \alpha_1)^{n_1} \dots (x - \alpha_\ell)^{n_\ell}.$$

  Moreover, $a, \alpha_1, \dots, \alpha_\ell, n_1, \dots, n_\ell$ are uniquely determined by the polynomial $p(x)$, up to a permutation of the α_i's and the corresponding n_i's.
- Here, a is the leading coefficient of $p(x)$, i.e. the coefficient in front of x^n. Numbers $\alpha_1, \dots, \alpha_\ell$ are the roots of $p(x)$ with *multiplicities* $n_1, \dots, n_\ell$, respectively.

- More precisely, if $p(x)$ is a polynomial of degree $n \geq 1$, and with coefficients in an algebraically closed field $\mathbb{F}$, then there exist numbers $a, \alpha_1, \dots, \alpha_\ell$ in $\mathbb{F}$ s.t. $a \neq 0$ and s.t. $\alpha_1, \dots, \alpha_\ell$ are pairwise distinct, and positive integers $n_1, \dots, n_\ell$ satisfying $n_1 + \dots + n_\ell = n$, s.t.

$$p(x) = a(x - \alpha_1)^{n_1} \dots (x - \alpha_\ell)^{n_\ell}.$$

  Moreover, $a, \alpha_1, \dots, \alpha_\ell, n_1, \dots, n_\ell$ are uniquely determined by the polynomial $p(x)$, up to a permutation of the α_i's and the corresponding n_i's.
- Here, a is the leading coefficient of $p(x)$, i.e. the coefficient in front of x^n. Numbers $\alpha_1, \dots, \alpha_\ell$ are the roots of $p(x)$ with *multiplicities* $n_1, \dots, n_\ell$, respectively.
- If we think of each α_i as being a root "n_i times" (due to its multiplicity), then we see that the n-th degree polynomial $p(x)$ has exactly n roots in $\mathbb{F}$.

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 $\mathbb{F}$.

- The discussion from the previous slide is often summarized as follows:

  Every n-th degree polynomial (with $n \geq 1$) with coefficients in an algebraically closed field has exactly n roots in that field, when multiplicities are taken into account.

3. Common roots of polynomials via determinants

3. Common roots of polynomials via determinants

- Any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ has a root in $\mathbb{F}$. However, there is no general formula for computing such a root.

3. Common roots of polynomials via determinants

- Any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ has a root in $\mathbb{F}$. However, there is no general formula for computing such a root.
- So, it may be surprising that, given arbitrary polynomials $p(x)$ and $q(x)$ with coefficients in an algebraically closed field $\mathbb{F}$, we can use determinants to determine whether $p(x)$ and $q(x)$ have a common root, i.e. whether there exists a number $x_0 \in \mathbb{F}$ for which we have $p(x_0) = 0$ and $q(x_0) = 0$ (next slide).

3. Common roots of polynomials via determinants

- Any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ has a root in $\mathbb{F}$. However, there is no general formula for computing such a root.
- So, it may be surprising that, given arbitrary polynomials $p(x)$ and $q(x)$ with coefficients in an algebraically closed field $\mathbb{F}$, we can use determinants to determine whether $p(x)$ and $q(x)$ have a common root, i.e. whether there exists a number $x_0 \in \mathbb{F}$ for which we have $p(x_0) = 0$ and $q(x_0) = 0$ (next slide).
- However, the determinant in question will only tell us whether such a common root exists; it provides no information on how one might actually compute such a root.

Theorem 7.11.1

Let $\mathbb{F}$ be an **algebraically closed field**. Let m and n be positive integers, and let $p(x) = \sum_{i=0}^{m} a_i x^i$ ($a_m \neq 0$) and $q(x) = \sum_{i=0}^{n} b_i x^i$ ($b_n \neq 0$) be polynomials with coefficients in $\mathbb{F}$. Let P be the $n \times (n+m)$ matrix whose j-th row (for $j \in \{1, \ldots, n\}$) is

$$\left[\quad \underbrace{0 \quad \ldots \quad 0}_{j-1} \quad a_m \quad a_{m-1} \quad \ldots \quad a_0 \quad \underbrace{0 \quad \ldots \quad 0}_{n-j} \quad \right],$$

and let Q be the $m \times (n+m)$ matrix whose j-th row (for $j \in \{1, \ldots, m\}$) is

$$\left[\quad \underbrace{0 \quad \ldots \quad 0}_{j-1} \quad b_n \quad b_{n-1} \quad \ldots \quad b_0 \quad \underbrace{0 \quad \ldots \quad 0}_{m-j} \quad \right].$$

Then $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$ iff

$$\det\left(\left[\begin{array}{c} P \\ \hline Q \end{array} \right] \right) = 0.$$

- First a more detailed explanation of how out matrix is formed, then an example, then a proof.

- For example, if $m = 3$ and $n = 5$, so that
  - $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$,
  - $q(x) = b_5x^5 + b_4x^4 + b_3x^3 + b_2x^2 + b_1x + b_0$,

  then we have

$$\left[\begin{array}{c} P \\ \hline Q \end{array}\right] = \left[\begin{array}{cccccccc} a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 \\ \hline b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \end{array}\right]_{8\times 8}.$$

Example 7.11.2

Determine whether the polynomials $p(x) = 5x^3 - 2x^2 + x - 4$ and $q(x) = 7x^2 - 6x - 1$ have a common complex root.

Proof.

Example 7.11.2

Determine whether the polynomials $p(x) = 5x^3 - 2x^2 + x - 4$ and $q(x) = 7x^2 - 6x - 1$ have a common complex root.

Proof. In this case, it is easy to see that $p(1) = 0$ and $q(1) = 0$, and so 1 is a common root of $p(x)$ and $q(x)$. However, let us use Theorem 7.11.1, in order to illustrate how this theorem can be applied.

Using the notation of Theorem 7.11.1, we have that $m = 3$, $n = 2$, and the matrices P and Q are given by

- $P = \begin{bmatrix} 5 & -2 & 1 & -4 & 0 \\ 0 & 5 & -2 & 1 & -4 \end{bmatrix}$;
- $Q = \begin{bmatrix} 7 & -6 & -1 & 0 & 0 \\ 0 & 7 & -6 & -1 & 0 \\ 0 & 0 & 7 & -6 & -1 \end{bmatrix}$.

Example 7.11.2

Determine whether the polynomials $p(x) = 5x^3 - 2x^2 + x - 4$ and $q(x) = 7x^2 - 6x - 1$ have a common complex root.

Proof (continued). We now have that

$$\det\left(\begin{bmatrix} P \\ \hline Q \end{bmatrix}\right) = \left|\begin{array}{ccccc} 5 & -2 & 1 & -4 & 0 \\ 0 & 5 & -2 & 1 & -4 \\ \hline 7 & -6 & -1 & 0 & 0 \\ 0 & 7 & -6 & -1 & 0 \\ 0 & 0 & 7 & -6 & -1 \end{array}\right| = 0.$$

Theorem 7.11.2 now guarantees that $p(x)$ and $q(x)$ have a common complex root. □

Theorem 7.11.1

Let $\mathbb{F}$ be an **algebraically closed field**. Let m and n be positive integers, and let $p(x) = \sum_{i=0}^{m} a_i x^i$ ($a_m \neq 0$) and $q(x) = \sum_{i=0}^{n} b_i x^i$ ($b_n \neq 0$) be polynomials with coefficients in $\mathbb{F}$. Let P be the $n \times (n+m)$ matrix whose j-th row (for $j \in \{1, \dots, n\}$) is

$$\left[\quad \underbrace{0 \quad \dots \quad 0}_{j-1} \quad a_m \quad a_{m-1} \quad \dots \quad a_0 \quad \underbrace{0 \quad \dots \quad 0}_{n-j} \quad \right],$$

and let Q be the $m \times (n+m)$ matrix whose j-th row (for $j \in \{1, \dots, m\}$) is

$$\left[\quad \underbrace{0 \quad \dots \quad 0}_{j-1} \quad b_n \quad b_{n-1} \quad \dots \quad b_0 \quad \underbrace{0 \quad \dots \quad 0}_{m-j} \quad \right].$$

Then $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$ iff

$$\det\left(\left[\begin{array}{c} P \\ \hline Q \end{array} \right] \right) = 0.$$

- Let's prove the theorem!

Proof.

Claim. *Polynomials $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$ iff there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ that satisfy the following:*

- $deg(r(x)) \leq n - 1$;
- $deg(s(x)) \leq m - 1$;
- $r(x)p(x) + s(x)q(x) = 0$.

Proof of the Claim.

Proof.

Claim. *Polynomials $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$ iff there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ that satisfy the following:*

- *$deg(r(x)) \leq n - 1$;*
- *$deg(s(x)) \leq m - 1$;*
- *$r(x)p(x) + s(x)q(x) = 0$.*

Proof of the Claim. Suppose first that $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$, say α.

Proof.

Claim. *Polynomials $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$ iff there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ that satisfy the following:*

- $deg(r(x)) \leq n-1$;
- $deg(s(x)) \leq m-1$;
- $r(x)p(x) + s(x)q(x) = 0$.

Proof of the Claim. Suppose first that $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$, say α. Then we set

$$r(x) := \frac{q(x)}{x-\alpha} \qquad \text{and} \qquad s(x) := -\frac{p(x)}{x-\alpha},$$

and we observe that $\deg(r(x)) = \deg(q(x)) - 1 = n-1$, $\deg(s(x)) = \deg(p(x)) - 1 = m-1$, and

$$r(x)p(x) + s(x)q(x) \;=\; \frac{q(x)p(x)}{x-\alpha} - \frac{p(x)q(x)}{x-\alpha} \;=\; 0.$$

Proof of the Claim (continued). Suppose conversely there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ s.t.

- $\deg\big(r(x)\big) \leq n - 1$;
- $\deg\big(s(x)\big) \leq m - 1$;
- $r(x)p(x) + s(x)q(x) = 0$.

WTS $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$.

Proof of the Claim (continued). Suppose conversely there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ s.t.

- $\deg(r(x)) \leq n-1$;
- $\deg(s(x)) \leq m-1$;
- $r(x)p(x) + s(x)q(x) = 0$.

WTS $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$.

Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with coefficients in $\mathbb{F}$, and they have exactly the same roots with the same corresponding multiplicities.

Proof of the Claim (continued). Suppose conversely there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ s.t.

- $\deg(r(x)) \leq n - 1$;
- $\deg(s(x)) \leq m - 1$;
- $r(x)p(x) + s(x)q(x) = 0$.

WTS $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$.

Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with coefficients in $\mathbb{F}$, and they have exactly the same roots with the same corresponding multiplicities.

Since $\deg(p(x)) = m$, we know that $p(x)$ has exactly m roots in $\mathbb{F}$ (when multiplicities are taken into account).

- Here, we are using the fact that $\mathbb{F}$ is algebraically closed.

Proof of the Claim (continued). Suppose conversely there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ s.t.

- $\deg(r(x)) \leq n - 1$;
- $\deg(s(x)) \leq m - 1$;
- $r(x)p(x) + s(x)q(x) = 0$.

WTS $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$.

Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with coefficients in $\mathbb{F}$, and they have exactly the same roots with the same corresponding multiplicities.

Since $\deg(p(x)) = m$, we know that $p(x)$ has exactly m roots in $\mathbb{F}$ (when multiplicities are taken into account).

- Here, we are using the fact that $\mathbb{F}$ is algebraically closed.

But $\deg(s(x)) \leq m - 1$, and so at least one of the roots of $p(x)$ either fails to be a root of $s(x)$, or is a root of $s(x)$ but has smaller multiplicity in $s(x)$ than in $p(x)$.

Proof of the Claim (continued). Suppose conversely there exist non-zero polynomials $r(x)$ and $s(x)$ with coefficients in $\mathbb{F}$ s.t.

- $\deg(r(x)) \leq n - 1$;
- $\deg(s(x)) \leq m - 1$;
- $r(x)p(x) + s(x)q(x) = 0$.

WTS $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$.

Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with coefficients in $\mathbb{F}$, and they have exactly the same roots with the same corresponding multiplicities.

Since $\deg(p(x)) = m$, we know that $p(x)$ has exactly m roots in $\mathbb{F}$ (when multiplicities are taken into account).

- Here, we are using the fact that $\mathbb{F}$ is algebraically closed.

But $\deg(s(x)) \leq m - 1$, and so at least one of the roots of $p(x)$ either fails to be a root of $s(x)$, or is a root of $s(x)$ but has smaller multiplicity in $s(x)$ than in $p(x)$. This root of $p(x)$ must therefore be a root of $q(x)$. ♦

Proof (continued). *We have now proven the Claim below:*

Claim. *Polynomials* $p(x)$ *and* $q(x)$ *have a common root in* $\mathbb{F}$ *iff there exist non-zero polynomials* $r(x)$ *and* $s(x)$ *with coefficients in* $\mathbb{F}$ *that satisfy the following:*

- $deg(r(x)) \leq n-1$;
- $deg(s(x)) \leq m-1$;
- $r(x)p(x) + s(x)q(x) = 0$.

Proof (continued). In view of the Claim, it now suffices to determine if there exist non-zero polynomials $r(x) = \sum_{i=0}^{n-1} c_i x^i$ and $s(x) = \sum_{i=0}^{m-1} d_i x^i$ s.t. $r(x)p(x) + s(x)q(x) = 0$.

Proof (continued). In view of the Claim, it now suffices to determine if there exist non-zero polynomials $r(x) = \sum_{i=0}^{n-1} c_i x^i$ and $s(x) = \sum_{i=0}^{m-1} d_i x^i$ s.t. $r(x)p(x) + s(x)q(x) = 0$.

So, we need to determine if there exist $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1} \in \mathbb{F}$ s.t. at least one of $c_0, \dots, c_{n-1}$ is non-zero and at least one of $d_0, \dots, d_{m-1}$ is non-zero, and s.t.

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)} \underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)} \underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} \quad = \quad 0.$$

Proof (continued). In view of the Claim, it now suffices to determine if there exist non-zero polynomials $r(x) = \sum_{i=0}^{n-1} c_i x^i$ and $s(x) = \sum_{i=0}^{m-1} d_i x^i$ s.t. $r(x)p(x) + s(x)q(x) = 0$.

So, we need to determine if there exist $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1} \in \mathbb{F}$ s.t. at least one of $c_0, \dots, c_{n-1}$ is non-zero and at least one of $d_0, \dots, d_{m-1}$ is non-zero, and s.t.

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} \quad = \quad 0.$$

But obviously, if $c_0, \dots, c_{n-1}$ are all zero, then $d_0, \dots, d_{m-1}$ are all zero, and vice versa.

Proof (continued). In view of the Claim, it now suffices to determine if there exist non-zero polynomials $r(x) = \sum_{i=0}^{n-1} c_i x^i$ and $s(x) = \sum_{i=0}^{m-1} d_i x^i$ s.t. $r(x)p(x) + s(x)q(x) = 0$.

So, we need to determine if there exist $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1} \in \mathbb{F}$ s.t. at least one of $c_0, \dots, c_{n-1}$ is non-zero and at least one of $d_0, \dots, d_{m-1}$ is non-zero, and s.t.

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)} \underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)} \underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} = 0.$$

But obviously, if $c_0, \dots, c_{n-1}$ are all zero, then $d_0, \dots, d_{m-1}$ are all zero, and vice versa. So, we in fact need to determine if the above equality holds for some numbers $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1} \in \mathbb{F}$, at least one of which is non-zero.

Proof (continued). Reminder:

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} \quad = \quad 0.$$

We now write the polynomial on the left-hand-side in the standard form, and we set all the coefficients that we obtain equal to zero.

- We can do this since our polynomial is identically zero, i.e. it is zero as a polynomial. This means precisely that all its coefficients are zero.

Proof (continued). Reminder:

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} \quad = \quad 0.$$

This yields a system of $n+m$ linear equations in the variables $c_{n-1}, \ldots, c_0, d_{m-1}, \ldots, d_0$ (we treat $a_m, \ldots, a_0, b_n, \ldots, b_0$ as constants).

Proof (continued). Reminder:

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} \quad = \quad 0.$$

This yields a system of $n+m$ linear equations in the variables $c_{n-1}, \dots, c_0, d_{m-1}, \dots, d_0$ (we treat $a_m, \dots, a_0, b_n, \dots, b_0$ as constants).

In each equation, we arrange the variables $c_{n-1}, \dots, c_0, d_{m-1}, \dots, d_0$ in this order from left to right. We arrange the equations for the coefficients in front of $x^{n+m-1}, \dots, x^1, x^0$ from top to bottom.

Proof (continued). Reminder:

$$\underbrace{\Big(\sum_{i=0}^{n-1} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{m} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{m-1} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{n} b_i x^i\Big)}_{=q(x)} \quad = \quad 0.$$

This yields a system of $n+m$ linear equations in the variables $c_{n-1}, \ldots, c_0, d_{m-1}, \ldots, d_0$ (we treat $a_m, \ldots, a_0, b_n, \ldots, b_0$ as constants).

In each equation, we arrange the variables $c_{n-1}, \ldots, c_0, d_{m-1}, \ldots, d_0$ in this order from left to right. We arrange the equations for the coefficients in front of $x^{n+m-1}, \ldots, x^1, x^0$ from top to bottom. We then rewrite this linear system as a matrix-vector equation

$$A \left[\begin{array}{ccc:ccc} c_{n-1} & \ldots & c_0 & d_{m-1} & \ldots & d_0 \end{array}\right]^T \quad = \quad \mathbf{0},$$

and we observe that the coefficient matrix A satisfies $A^T = \left[\begin{array}{c} P \\ \hdashline Q \end{array}\right]$.

- Intermission: Let's look at an example with $m = 3$ and $n = 5$.

Intermission: Example with $m = 3$ and $n = 5$. Then

- $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$,
- $q(x) = b_5x^5 + b_4x^4 + b_3x^3 + b_2x^2 + b_1x + b_0$,
- $r(x) = c_4x^4 + c_3x^3 + c_2x^2 + c_1x + c_0$,
- $s(t) = d_2x^2 + d_1x + d_0$,

then our equation becomes

$$\underbrace{\Big(\sum_{i=0}^{4} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{3} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{2} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{5} b_i x^i\Big)}_{=q(x)} \quad = \quad 0$$

which yields the system of linear equations on the next slide (we consider the coefficients in front of $x^7, x^6, x^5, x^4, x^3, x^2, x^1, x^0$ from top to bottom, and we arrange the variables $c_4, c_3, c_2, c_1, c_0, d_2, d_1, d_0$ from left to right).

Intermission (continued): Example with $m = 3$ and $n = 5$.
Reminder: our equation was

$$\underbrace{\left(\sum_{i=0}^{4} c_i x^i\right)}_{=r(x)} \underbrace{\left(\sum_{i=0}^{3} a_i x^i\right)}_{=p(x)} + \underbrace{\left(\sum_{i=0}^{2} d_i x^i\right)}_{=s(x)} \underbrace{\left(\sum_{i=0}^{5} b_i x^i\right)}_{=q(x)} \quad = \quad 0$$

	c_4		c_3		c_2		c_1		c_0		d_2		d_1		d_0		
x^7	$a_3 c_4$									$+$	$b_5 d_2$					$=$	0
x^6	$a_2 c_4$	$+$	$a_3 c_3$							$+$	$b_4 d_2$	$+$	$b_5 d_1$			$=$	0
x^5	$a_1 c_4$	$+$	$a_2 c_3$	$+$	$a_3 c_2$					$+$	$b_3 d_2$	$+$	$b_4 d_1$	$+$	$b_5 d_0$	$=$	0
x^4	$a_0 c_4$	$+$	$a_1 c_3$	$+$	$a_2 c_2$	$+$	$a_3 c_1$			$+$	$b_2 d_2$	$+$	$b_3 d_1$	$+$	$b_4 d_0$	$=$	0
x^3			$a_0 c_3$	$+$	$a_1 c_2$	$+$	$a_2 c_1$	$+$	$a_3 c_0$	$+$	$b_1 d_2$	$+$	$b_2 d_1$	$+$	$b_3 d_0$	$=$	0
x^2					$a_0 c_2$	$+$	$a_1 c_1$	$+$	$a_2 c_0$	$+$	$b_0 d_2$	$+$	$b_1 d_1$	$+$	$b_2 d_0$	$=$	0
x^1							$a_0 c_1$	$+$	$a_1 c_0$			$+$	$b_0 d_1$	$+$	$b_1 d_0$	$=$	0
x^0									$a_0 c_0$					$+$	$b_0 d_0$	$=$	0

Intermission (continued): Example with $m = 3$ and $n = 5$.
Reminder: our equation was

$$\underbrace{\Big(\sum_{i=0}^{4} c_i x^i\Big)}_{=r(x)}\underbrace{\Big(\sum_{i=0}^{3} a_i x^i\Big)}_{=p(x)} + \underbrace{\Big(\sum_{i=0}^{2} d_i x^i\Big)}_{=s(x)}\underbrace{\Big(\sum_{i=0}^{5} b_i x^i\Big)}_{=q(x)} = 0$$

	c_4		c_3		c_2		c_1		c_0		d_2		d_1		d_0		
x^7	$a_3 c_4$									$+$	$b_5 d_2$					$=$	0
x^6	$a_2 c_4$	$+$	$a_3 c_3$							$+$	$b_4 d_2$	$+$	$b_5 d_1$			$=$	0
x^5	$a_1 c_4$	$+$	$a_2 c_3$	$+$	$a_3 c_2$					$+$	$b_3 d_2$	$+$	$b_4 d_1$	$+$	$b_5 d_0$	$=$	0
x^4	$a_0 c_4$	$+$	$a_1 c_3$	$+$	$a_2 c_2$	$+$	$a_3 c_1$			$+$	$b_2 d_2$	$+$	$b_3 d_1$	$+$	$b_4 d_0$	$=$	0
x^3			$a_0 c_3$	$+$	$a_1 c_2$	$+$	$a_2 c_1$	$+$	$a_3 c_0$	$+$	$b_1 d_2$	$+$	$b_2 d_1$	$+$	$b_3 d_0$	$=$	0
x^2					$a_0 c_2$	$+$	$a_1 c_1$	$+$	$a_2 c_0$	$+$	$b_0 d_2$	$+$	$b_1 d_1$	$+$	$b_2 d_0$	$=$	0
x^1							$a_0 c_1$	$+$	$a_1 c_0$			$+$	$b_0 d_1$	$+$	$b_1 d_0$	$=$	0
x^0									$a_0 c_0$					$+$	$b_0 d_0$	$=$	0

This linear system, in turn, translates into the following matrix-vector equation (next slide):

Intermission (continued): Example with $m = 3$ and $n = 5$.

$$\left[\begin{array}{ccccc:ccc} a_3 & 0 & 0 & 0 & 0 & b_5 & 0 & 0 \\ a_2 & a_3 & 0 & 0 & 0 & b_4 & b_5 & 0 \\ a_1 & a_2 & a_3 & 0 & 0 & b_3 & b_4 & b_5 \\ a_0 & a_1 & a_2 & a_3 & 0 & b_2 & b_3 & b_4 \\ 0 & a_0 & a_1 & a_2 & a_3 & b_1 & b_2 & b_3 \\ 0 & 0 & a_0 & a_1 & a_2 & b_0 & b_1 & b_2 \\ 0 & 0 & 0 & a_0 & a_1 & 0 & b_0 & b_1 \\ 0 & 0 & 0 & 0 & a_0 & 0 & 0 & b_0 \end{array}\right] \left[\begin{array}{c} c_4 \\ c_3 \\ c_2 \\ c_1 \\ c_0 \\ \hdashline d_2 \\ d_1 \\ d_0 \end{array}\right] = \mathbf{0}.$$

Intermission (continued): Example with $m = 3$ and $n = 5$.

$$\left[\begin{array}{ccccc|ccc} a_3 & 0 & 0 & 0 & 0 & b_5 & 0 & 0 \\ a_2 & a_3 & 0 & 0 & 0 & b_4 & b_5 & 0 \\ a_1 & a_2 & a_3 & 0 & 0 & b_3 & b_4 & b_5 \\ a_0 & a_1 & a_2 & a_3 & 0 & b_2 & b_3 & b_4 \\ 0 & a_0 & a_1 & a_2 & a_3 & b_1 & b_2 & b_3 \\ 0 & 0 & a_0 & a_1 & a_2 & b_0 & b_1 & b_2 \\ 0 & 0 & 0 & a_0 & a_1 & 0 & b_0 & b_1 \\ 0 & 0 & 0 & 0 & a_0 & 0 & 0 & b_0 \end{array}\right] \left[\begin{array}{c} c_4 \\ c_3 \\ c_2 \\ c_1 \\ c_0 \\ \hline d_2 \\ d_1 \\ d_0 \end{array}\right] = \mathbf{0}.$$

The transpose of the coefficient matrix that we obtained is precisely the matrix

$$\left[\begin{array}{c} P \\ \hline Q \end{array}\right] = \left[\begin{array}{cccccccc} a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 \\ \hline b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \end{array}\right]_{8\times 8}.$$

Proof (continued). We now have the following sequence of equivalent statements:

$$p(x) \text{ and } q(x) \text{ have a common root in } \mathbb{F} \iff A \left[\, c_{n-1} \quad \dots \quad c_0 \;\vdots\; d_{m-1} \quad \dots \quad d_0 \,\right]^T = \mathbf{0} \text{ has a non-zero solution}$$

$$\stackrel{(*)}{\iff} A \text{ is non-invertible}$$

$$\stackrel{(*)}{\iff} A^T = \begin{bmatrix} \text{-} \; P \; \text{-} \\ \hline \text{-} \; Q \; \text{-} \end{bmatrix} \text{ is non-invertible}$$

$$\stackrel{(*)}{\iff} \det\left(\begin{bmatrix} \text{-} \; P \; \text{-} \\ \hline \text{-} \; Q \; \text{-} \end{bmatrix} \right) = 0,$$

where all three instances of (*) follow from the Invertible Matrix Theorem. □

Theorem 7.11.1

Let $\mathbb{F}$ be an **algebraically closed field**. Let m and n be positive integers, and let $p(x) = \sum_{i=0}^{m} a_i x^i$ ($a_m \neq 0$) and $q(x) = \sum_{i=0}^{n} b_i x^i$ ($b_n \neq 0$) be polynomials with coefficients in $\mathbb{F}$. Let P be the $n \times (n+m)$ matrix whose j-th row (for $j \in \{1, \dots, n\}$) is

$$\left[\begin{array}{ccccccc} \underbrace{0 \quad \dots \quad 0}_{j-1} & a_m & a_{m-1} & \dots & a_0 & \underbrace{0 \quad \dots \quad 0}_{n-j} \end{array}\right],$$

and let Q be the $m \times (n+m)$ matrix whose j-th row (for $j \in \{1, \dots, m\}$) is

$$\left[\begin{array}{ccccccc} \underbrace{0 \quad \dots \quad 0}_{j-1} & b_n & b_{n-1} & \dots & b_0 & \underbrace{0 \quad \dots \quad 0}_{m-j} \end{array}\right].$$

Then $p(x)$ and $q(x)$ have a common root in $\mathbb{F}$ iff

$$\det\left(\left[\begin{array}{c} P \\ \hline Q \end{array}\right]\right) = 0.$$