Linear Algebra 2

Lecture #17
Applications of determinants: polynomials and volume

Irena Penev

March 30, 2023

- This lecture has three parts:
  1. the Vandermonde matrix;
  2. common roots of polynomials via determinants;
  3. determinants and volume

1. The Vandermonde matrix

1. The Vandermonde matrix

Definition

For a positive integer n and real numbers $a_0, a_1, \ldots, a_n$, we define the matrix

$$V(a_0, a_1, \ldots, a_n) := \begin{bmatrix} 1 & 1 & \ldots & 1 \\ a_0 & a_1 & \ldots & a_n \\ a_0^2 & a_1^2 & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_0^n & a_1^n & \ldots & a_n^n \end{bmatrix}_{(n+1)\times(n+1)},$$

called the *Vandermonde matrix*.

Proposition 1.1

For all positive integers n and real numbers $a_0, a_1, \ldots, a_n$, we have that

$$\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i>j}(a_i - a_j) = \prod_{i=1}^{n}\prod_{j=0}^{i-1}(a_i - a_j).$$

and consequently, the matrix $V(a_0, a_1, \ldots, a_n)$ is invertible iff $a_0, a_1, \ldots, a_n$ are pairwise distinct.

Proof.

Proposition 1.1

For all positive integers n and real numbers $a_0, a_1, \ldots, a_n$, we have that

$$\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i>j}(a_i - a_j) = \prod_{i=1}^{n}\prod_{j=0}^{i-1}(a_i - a_j).$$

and consequently, the matrix $V(a_0, a_1, \ldots, a_n)$ is invertible iff $a_0, a_1, \ldots, a_n$ are pairwise distinct.

Proof. By Theorem 5.1 from Lecture Notes 15, we know that a square matrix is invertible iff its determinant is non-zero. So, the second statement of the proposition follows immediately from the first, i.e. from the formula for the determinant of the Vandermonde matrix.

Proof (continued). WTS $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

Proof (continued). WTS $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

We prove the formula for the determinant of the Vandermonde by induction on n.

Proof (continued). WTS $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

We prove the formula for the determinant of the Vandermonde by induction on n. For $n = 1$, we note that $\forall a_0, a_1 \in \mathbb{R}$, we have that

$$\det\Big(V(a_0, a_1)\Big) = \begin{vmatrix} 1 & 1 \\ a_0 & a_1 \end{vmatrix} = (a_1 - a_0).$$

Proof (continued). WTS $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

We prove the formula for the determinant of the Vandermonde by induction on n. For $n = 1$, we note that $\forall a_0, a_1 \in \mathbb{R}$, we have that

$$\det\Big(V(a_0, a_1)\Big) = \begin{vmatrix} 1 & 1 \\ a_0 & a_1 \end{vmatrix} = (a_1 - a_0).$$

Now, fix a positive integer n, and assume inductively that our formula is correct for n, i.e. that for all real numbers $a_0, a_1, \ldots, a_n$, we have that $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$. WTS the formula is correct for $n + 1$.

Proof (continued). WTS $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

We prove the formula for the determinant of the Vandermonde by induction on n. For $n = 1$, we note that $\forall a_0, a_1 \in \mathbb{R}$, we have that

$$\det\Big(V(a_0, a_1)\Big) = \begin{vmatrix} 1 & 1 \\ a_0 & a_1 \end{vmatrix} = (a_1 - a_0).$$

Now, fix a positive integer n, and assume inductively that our formula is correct for n, i.e. that for all real numbers $a_0, a_1, \ldots, a_n$, we have that $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1}(a_i - a_j)$. WTS the formula is correct for $n + 1$. Fix $a_0, a_1, \ldots, a_n, a_{n+1} \in \mathbb{R}$; WTS $\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1}(a_i - a_j)$.

Proof (continued). WTS $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

We prove the formula for the determinant of the Vandermonde by induction on n. For $n = 1$, we note that $\forall a_0, a_1 \in \mathbb{R}$, we have that

$$\det\Big(V(a_0, a_1)\Big) = \begin{vmatrix} 1 & 1 \\ a_0 & a_1 \end{vmatrix} = (a_1 - a_0).$$

Now, fix a positive integer n, and assume inductively that our formula is correct for n, i.e. that for all real numbers $a_0, a_1, \dots, a_n$, we have that $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$. WTS the formula is correct for $n + 1$. Fix $a_0, a_1, \dots, a_n, a_{n+1} \in \mathbb{R}$; WTS $\det\Big(V(a_0, a_1, \dots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

If some two of $a_0, \dots, a_{n+1}$ are the same, then this is obvious. Indeed, in this case, the matrix $V(a_0, a_1, \dots, a_n, a_{n+1})$ has two identical columns and therefore has determinant zero, and clearly, $\prod_{i=1}^{n+1} \prod_{j=0}^{i} (a_i - a_j) = 0$ (because one of the factors is zero).

Proof (continued). WTS $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n}\prod_{j=0}^{i-1}(a_i - a_j)$.

We prove the formula for the determinant of the Vandermonde by induction on n. For $n = 1$, we note that $\forall a_0, a_1 \in \mathbb{R}$, we have that

$$\det\Big(V(a_0, a_1)\Big) = \begin{vmatrix} 1 & 1 \\ a_0 & a_1 \end{vmatrix} = (a_1 - a_0).$$

Now, fix a positive integer n, and assume inductively that our formula is correct for n, i.e. that for all real numbers $a_0, a_1, \dots, a_n$, we have that $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n}\prod_{j=0}^{i-1}(a_i - a_j)$. WTS the formula is correct for $n + 1$. Fix $a_0, a_1, \dots, a_n, a_{n+1} \in \mathbb{R}$; WTS $\det\Big(V(a_0, a_1, \dots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1}\prod_{j=0}^{i-1}(a_i - a_j)$.

If some two of $a_0, \dots, a_{n+1}$ are the same, then this is obvious. Indeed, in this case, the matrix $V(a_0, a_1, \dots, a_n, a_{n+1})$ has two identical columns and therefore has determinant zero, and clearly, $\prod_{i=1}^{n+1}\prod_{j=0}^{i}(a_i - a_j) = 0$ (because one of the factors is zero). So, from now on, WMA $a_0, \dots, a_{n+1}$ are pairwise distinct.

Proof (continued). Reminder: $a_0, \ldots, a_{n+1}$ are pairwise distinct;

induction hypothesis: $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;

WTS $\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

Proof (continued). Reminder: $a_0, \ldots, a_{n+1}$ are pairwise distinct;

induction hypothesis: $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;

WTS $\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

$$f(t) \quad := \quad \begin{vmatrix} 1 & 1 & \ldots & 1 & 1 \\ a_0 & a_1 & \ldots & a_n & t \\ a_0^2 & a_1^2 & \ldots & a_n^2 & t^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_0^n & a_1^n & \ldots & a_n^n & t^n \\ a_0^{n+1} & a_1^{n+1} & \ldots & a_n^{n+1} & t^{n+1} \end{vmatrix},$$

so that $f(a_{n+1}) = V(a_0, a_1, \ldots, a_n, a_{n+1})$.

Proof (continued). Reminder: $a_0, \ldots, a_{n+1}$ are pairwise distinct; induction hypothesis: $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;

WTS $\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

$$f(t) \quad := \quad \begin{vmatrix} 1 & 1 & \ldots & 1 & 1 \\ a_0 & a_1 & \ldots & a_n & t \\ a_0^2 & a_1^2 & \ldots & a_n^2 & t^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_0^n & a_1^n & \ldots & a_n^n & t^n \\ a_0^{n+1} & a_1^{n+1} & \ldots & a_n^{n+1} & t^{n+1} \end{vmatrix},$$

so that $f(a_{n+1}) = V(a_0, a_1, \ldots, a_n, a_{n+1})$. By performing Laplace expansion along the last column, we see that $f(t)$ is a polynomial of degree $n+1$, and that its leading coefficient (i.e. coefficient in front of t^{n+1}) is $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) \overset{\text{ind. hyp.}}{=} \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$.

Proof (continued). Reminder: $a_0, \dots, a_{n+1}$ are pairwise distinct; induction hypothesis: $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;

WTS $\det\Big(V(a_0, a_1, \dots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

$$f(t) \quad := \quad \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ a_0 & a_1 & \dots & a_n & t \\ a_0^2 & a_1^2 & \dots & a_n^2 & t^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_0^n & a_1^n & \dots & a_n^n & t^n \\ a_0^{n+1} & a_1^{n+1} & \dots & a_n^{n+1} & t^{n+1} \end{vmatrix},$$

so that $f(a_{n+1}) = V(a_0, a_1, \dots, a_n, a_{n+1})$. By performing Laplace expansion along the last column, we see that $f(t)$ is a polynomial of degree $n+1$, and that its leading coefficient (i.e. coefficient in front of t^{n+1}) is $\det\Big(V(a_0, a_1, \dots, a_n)\Big) \overset{\text{ind. hyp.}}{=} \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$. Also, $a_0, \dots, a_n$ are the roots of $f(t)$.

Proof (continued). Reminder: $a_0, \dots, a_{n+1}$ are pairwise distinct; induction hypothesis: $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;

WTS $\det\Big(V(a_0, a_1, \dots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

$$f(t) \quad := \quad \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ a_0 & a_1 & \dots & a_n & t \\ a_0^2 & a_1^2 & \dots & a_n^2 & t^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_0^n & a_1^n & \dots & a_n^n & t^n \\ a_0^{n+1} & a_1^{n+1} & \dots & a_n^{n+1} & t^{n+1} \end{vmatrix},$$

so that $f(a_{n+1}) = V(a_0, a_1, \dots, a_n, a_{n+1})$. By performing Laplace expansion along the last column, we see that $f(t)$ is a polynomial of degree $n+1$, and that its leading coefficient (i.e. coefficient in front of t^{n+1}) is $\det\Big(V(a_0, a_1, \dots, a_n)\Big) \overset{\text{ind. hyp.}}{=} \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$. Also, $a_0, \dots, a_n$ are the roots of $f(t)$. So, $f(t)$ can be factored as (next slide):

Proof (continued). Reminder: $a_0, \ldots, a_{n+1}$ are pairwise distinct; induction hypothesis: $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;
WTS $\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

$$f(t) \;=\; k \prod_{j=0}^{n} (t - a_j) \;=\; \Big(\prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j) \Big) \Big(\prod_{j=0}^{n} (t - a_j) \Big).$$

Proof (continued). Reminder: $a_0, \ldots, a_{n+1}$ are pairwise distinct; induction hypothesis: $\det\Big(V(a_0, a_1, \ldots, a_n)\Big) = \prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j)$;
WTS $\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j)$.

$$f(t) \;=\; k \prod_{j=0}^{n} (t - a_j) \;=\; \Big(\prod_{i=1}^{n} \prod_{j=0}^{i-1} (a_i - a_j) \Big) \Big(\prod_{j=0}^{n} (t - a_j) \Big).$$

Consequently,

$$\det\Big(V(a_0, a_1, \ldots, a_n, a_{n+1})\Big) \;=\; f(a_{n+1}) \;=\; \prod_{i=1}^{n+1} \prod_{j=0}^{i-1} (a_i - a_j).$$

This completes the induction. Q.E.D.

$$V(a_0, a_1, \ldots, a_n) \quad := \quad \begin{bmatrix} 1 & 1 & \ldots & 1 \\ a_0 & a_1 & \ldots & a_n \\ a_0^2 & a_1^2 & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_0^n & a_1^n & \ldots & a_n^n \end{bmatrix}_{(n+1)\times(n+1)},$$

Proposition 1.1

For all positive integers n and real numbers $a_0, a_1, \ldots, a_n$, we have that

$$\det\Big(V(a_0, a_1, \ldots, a_n)\Big) \quad = \quad \prod_{i>j}(a_i - a_j) \quad = \quad \prod_{i=1}^{n}\prod_{j=0}^{i-1}(a_i - a_j).$$

and consequently, the matrix $V(a_0, a_1, \ldots, a_n)$ is invertible iff $a_0, a_1, \ldots, a_n$ are pairwise distinct.

2. Common roots of polynomials via determinants

2. Common roots of polynomials via determinants

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

2. Common roots of polynomials via determinants

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- **Remark:** The Fundamental Theorem of Algebra applies to the field $\mathbb{C}$, but it does **not** work for general fields $\mathbb{F}$.

2. Common roots of polynomials via determinants

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- **Remark:** The Fundamental Theorem of Algebra applies to the field $\mathbb{C}$, but it does **not** work for general fields $\mathbb{F}$.
    - It fails for the fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{Z}_p$ (for any prime p).

2. Common roots of polynomials via determinants

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- **Remark:** The Fundamental Theorem of Algebra applies to the field $\mathbb{C}$, but it does **not** work for general fields $\mathbb{F}$.
  - It fails for the fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{Z}_p$ (for any prime p).
  - Examples: Lecture Notes.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- The Fundamental Theorem of Algebra readily implies that any polynomial $p(x)$ with complex coefficients and of degree $n \geq 1$ can be factored into n linear factors.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- The Fundamental Theorem of Algebra readily implies that any polynomial $p(x)$ with complex coefficients and of degree $n \geq 1$ can be factored into n linear factors.
- More precisely, for such a polynomial $p(x)$, there exist complex numbers $a, \alpha_1, \dots, \alpha_\ell$ 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 $n_1 + \cdots + n_\ell = n$, s.t.

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

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- The Fundamental Theorem of Algebra readily implies that any polynomial $p(x)$ with complex coefficients and of degree $n \geq 1$ can be factored into n linear factors.
- More precisely, for such a polynomial $p(x)$, there exist complex numbers $a, \alpha_1, \dots, \alpha_\ell$ 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.

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

- Here, a is the leading coefficient of $p(x)$, i.e. the coefficient in front of x^n. Complex numbers $\alpha_1, \dots, \alpha_\ell$ are the roots of $p(x)$ with *multiplicities* $n_1, \dots, n_\ell$, respectively.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- If $n \leq 4$, then there are formulas that allow us to compute the roots of $p(x)$.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- If $n \leq 4$, then there are formulas that allow us to compute the roots of $p(x)$.
- However, if $n \geq 5$, then no such formula exists: we know that $p(x)$ has n complex roots (when multiplicities are taken into account), but in general, there is no formula for computing these roots.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- If $n \leq 4$, then there are formulas that allow us to compute the roots of $p(x)$.
- However, if $n \geq 5$, then no such formula exists: we know that $p(x)$ has n complex roots (when multiplicities are taken into account), but in general, there is no formula for computing these roots.
- In fact, not only is no such formula known, but using Galois theory, one can show that no such formula can exist for polynomials of degree at least five.

The Fundamental Theorem of Algebra

Any non-constant polynomial $p(x)$ with complex coefficients has a complex root.

- If $n \leq 4$, then there are formulas that allow us to compute the roots of $p(x)$.
- However, if $n \geq 5$, then no such formula exists: we know that $p(x)$ has n complex roots (when multiplicities are taken into account), but in general, there is no formula for computing these roots.
- In fact, not only is no such formula known, but using Galois theory, one can show that no such formula can exist for polynomials of degree at least five.
    - Of course, we may be able to use various tricks to compute the roots of some special high-degree polynomials. However, none of these tricks will work in the general case.

Theorem 2.1

Let m and n be positive integers, and let $p(x) = \sum_{i=0}^{m} a_i x^i$ (with $a_m \neq 0$) and $q(x) = \sum_{i=0}^{n} b_i x^i$ ($b_n \neq 0$) be polynomials with complex coefficients. 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 complex root iff

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

- **Remark#1:** 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

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

- **Remark#2:** Theorem 2.1 works for polynomials with **complex** coefficients, but not for polynomials with coefficients in an arbitrary field $\mathbb{F}$. This is essentially because the proof of Theorem 2.1 relies on the Fundamental Theorem of Algebra, which works for $\mathbb{C}$, but not for general fields $\mathbb{F}$.

- **Remark#2:** Theorem 2.1 works for polynomials with **complex** coefficients, but not for polynomials with coefficients in an arbitrary field $\mathbb{F}$. This is essentially because the proof of Theorem 2.1 relies on the Fundamental Theorem of Algebra, which works for $\mathbb{C}$, but not for general fields $\mathbb{F}$.
- **Remark#3:** Theorem 2.1 gives a recipe for determining whether two polynomials with complex coefficients have a complex root, but it does not give a recipe for actually computing such a common root (if it exists).

Theorem 2.1

Let m and n be positive integers, and let $p(x) = \sum_{i=0}^{m} a_i x^i$ (with $a_m \neq 0$) and $q(x) = \sum_{i=0}^{n} b_i x^i$ ($b_n \neq 0$) be polynomials with complex coefficients. 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 complex root iff

$$\det\left(\left[\frac{P}{Q} \right] \right) = 0.$$

- First an example, then a proof.

- First an example, then a proof.

Example 2.2

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

Solution.

- First an example, then a proof.

Example 2.2

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

Solution. 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 2.1, in order to illustrate how this theorem can be applied.

Using the notation of Theorem 2.1, 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 2.2

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

Solution (continued). We now have that

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

Theorem 2.1 now guarantees that $p(x)$ and $q(x)$ have a common complex root.

Theorem 2.1

Let m and n be positive integers, and let $p(x) = \sum_{i=0}^{m} a_i x^i$ (with $a_m \neq 0$) and $q(x) = \sum_{i=0}^{n} b_i x^i$ ($b_n \neq 0$) be polynomials with complex coefficients. 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 complex root iff

$$\det\left(\left[\frac{P}{Q} \right] \right) = 0.$$

Proof.

Claim. *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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 complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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, say α.

Proof.

Claim. *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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, say α. Then we set $r(x) := \frac{q(x)}{x-\alpha}$ and $s(x) := -\frac{p(x)}{x-\alpha}$, and we observe that $\deg(r(x)) = \deg(q(x)) - 1 = n - 1$, $\deg(s(x)) = \deg(p(x)) - 1 = m - 1$, and

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

Proof (continued).

Claim. *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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 (continued).

Proof (continued).

> ***Claim.*** *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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 (continued). Suppose conversely there exist $r(x), s(x)$ with the desired properties.

Proof (continued).

Claim. *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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 (continued). Suppose conversely there exist $r(x), s(x)$ with the desired properties. Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with complex coefficients, and they have exactly the same roots with the same corresponding multiplicities.

Proof (continued).

Claim. *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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 (continued). Suppose conversely there exist $r(x), s(x)$ with the desired properties. Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with complex coefficients, and they have exactly the same roots with the same corresponding multiplicities.

Since $\deg(p(x)) = m$, the Fundamental Theorem of Algebra implies that $p(x)$ has exactly m complex roots (when multiplicities are taken into account).

Proof (continued).

Claim. *Polynomials $p(x)$ and $q(x)$ have a common complex root iff there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients 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 (continued). Suppose conversely there exist $r(x), s(x)$ with the desired properties. Then $r(x)p(x)$ and $s(x)q(x)$ are non-constant polynomials with complex coefficients, and they have exactly the same roots with the same corresponding multiplicities.

Since $\deg(p(x)) = m$, the Fundamental Theorem of Algebra implies that $p(x)$ has exactly m complex roots (when multiplicities are taken into account). 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 $q(x)$. This proves the Claim.

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 complex numbers $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1}$ 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 complex numbers $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1}$ 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 complex numbers $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1}$, at least one of which is non-zero.

Proof (continued). Reminder:

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

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.

Proof (continued). Reminder:

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

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. 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).

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. 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:

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

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. 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. We then rewrite this linear system as a matrix-vector equation

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

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

Proof (continued). We now have the following equivalences:

$$
\begin{array}{lll}
\begin{array}{l} p(x) \text{ and } q(x) \\ \text{have a common} \\ \text{complex root} \end{array} & \iff & \begin{array}{l} A \left[\begin{array}{ccc:ccc} c_{n-1} & \dots & c_0 & d_{m-1} & \dots & d_0 \end{array} \right]^T = \mathbf{0} \\ \text{has a non-zero solution} \end{array} \\
& \iff & A \text{ is non-invertible} \\
& \iff & A^T = \left[\begin{array}{c} P \\ \hline Q \end{array} \right] \text{ is non-invertible} \\
& \iff & \det\left(\left[\begin{array}{c} P \\ \hline Q \end{array} \right] \right) = 0,
\end{array}
$$

Q.E.D.

3. Determinants and volume

3. Determinants and volume

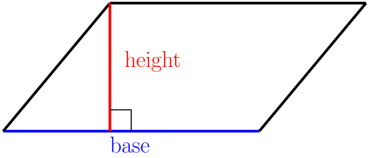
- From now on, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

3. Determinants and volume

- From now on, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

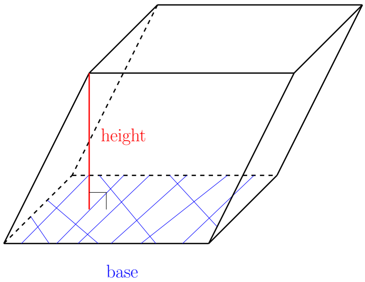
- For a parallelogram, we have the familiar formula

$$\left(\begin{array}{c}\text{area of}\\\text{parallelogram}\end{array}\right) = (\text{length of base}) \times (\text{height}).$$



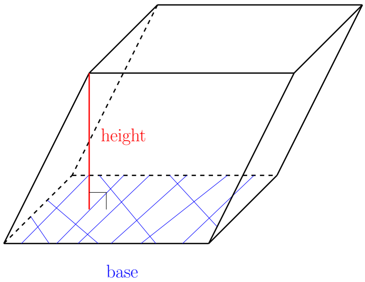
- We have a similar formula for the volume of a parallelepiped:

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



- 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.

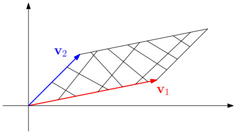
- Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the *m-parallelepiped* determined by vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

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

- 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

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

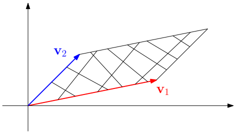
- For instance, given two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$, neither of which is a scalar multiple of the other, the 2-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2$ is a just a parallelogram.



- 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

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

- 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 a just a parallelogram.

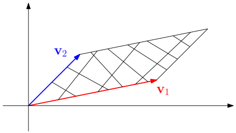


- For vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, neither of which is a scalar multiple of each other, the 2-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2$ is still a parallelogram, but this parallelogram lies in the plane (2-dimensional subspace) $\text{Span}(\mathbf{v}_1, \mathbf{v}_2)$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

- 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

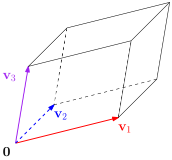
$$\left\{c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m \in \mathbb{R}^n \;\middle|\; c_1, \ldots, c_m \in \mathbb{R}, 0 \leq c_1, \ldots, c_m \leq 1\right\}.$$

- 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 a just a parallelogram.

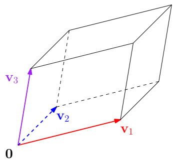


- For vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, neither of which is a scalar multiple of each other, the 2-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2$ is still a parallelogram, but this parallelogram lies in the plane (2-dimensional subspace) Span$(\mathbf{v}_1, \mathbf{v}_2)$ of $\mathbb{R}^n$.
- If one of $\mathbf{v}_1, \mathbf{v}_2$ is a scalar multiple of the other, then we get a "degenerate parallelogram" (line segment or just $\{\mathbf{0}\}$).

- 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.

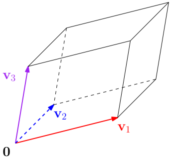


- Similarly, for three linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^n$, the 3-parallelepiped defined by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is just the usual parallelepiped whose edges are determined by these three vectors.



- 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."

- 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.



- 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.

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

- 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.

- 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.
  - The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be

$$V_1(\mathbf{v}_1) \quad := \quad ||\mathbf{v}_1||.$$

- 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.
  - The *1-volume* of the 1-parallelepiped determined by the vector $\mathbf{v}_1 \in \mathbb{R}^n$ is defined to be

$$V_1(\mathbf{v}_1) \quad := \quad ||\mathbf{v}_1||.$$

  - 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 $\mathbf{v}_{m+1}^{\perp} = \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$, or equivalently, $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)}(\mathbf{v}_{m+1})$.

- Reminder:
  - $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$;
  - $V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) := V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m)\ ||\mathbf{v}_{m+1}^{\perp}||$.

- Reminder:
  - $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$;
  - $V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) := V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \, ||\mathbf{v}_{m+1}^{\perp}||$.

- 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."

- Reminder:
  - $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$;
  - $V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) := V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m)\, ||\mathbf{v}_{m+1}^{\perp}||$.

- 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

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

- Reminder:
  - $V_1(\mathbf{v}_1) := ||\mathbf{v}_1||$;
  - $V_{m+1}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}) := V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m)\, ||\mathbf{v}_{m+1}^{\perp}||$.

- 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

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

- 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.

- Obviously, we would like volume to be non-negative and invariant under vector reordering (i.e. the m-volume of an m-parallelepiped should not change if we merely reorder the vectors determining this m-parallelepiped).

- Obviously, we would like volume to be non-negative and invariant under vector reordering (i.e. the m-volume of an m-parallelepiped should not change if we merely reorder the vectors determining this m-parallelepiped).

Proposition 3.1

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.

Corollary 3.4

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

- Obviously, we would like volume to be non-negative and invariant under vector reordering (i.e. the m-volume of an m-parallelepiped should not change if we merely reorder the vectors determining this m-parallelepiped).

Proposition 3.1

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

Corollary 3.4

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

- The proof of Corollary 3.4 uses Theorem 3.2, which gives a formula for the volume in terms of the determinant of a certain square matrix (later!).

Proposition 3.1

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 (outline).

Proposition 3.1

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 (outline). For each $i \in \{1, \ldots, m-1\}$, set $\mathbf{v}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_i)^{\perp}}(\mathbf{v}_{i+1})$.

Proposition 3.1

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 (outline). For each $i \in \{1, \ldots, m-1\}$, set
$\mathbf{v}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_i)^{\perp}}(\mathbf{v}_{i+1})$.
It then follows from our recursive definition of $V(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ that

$$V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \quad = \quad ||\mathbf{v}_1|| \, ||\mathbf{v}_2^{\perp}|| \, \ldots \, ||\mathbf{v}_m^{\perp}||.$$

Proposition 3.1

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 (outline). For each $i \in \{1, \ldots, m-1\}$, set
$\mathbf{v}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_i)^{\perp}}(\mathbf{v}_{i+1})$.
It then follows from our recursive definition of $V(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ that

$$V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \quad = \quad ||\mathbf{v}_1|| \, ||\mathbf{v}_2^{\perp}|| \, \ldots \, ||\mathbf{v}_m^{\perp}||.$$

Since the length of any vector is non-negative, we see that $V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \geq 0$.

Proposition 3.1

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 (outline). For each $i \in \{1, \ldots, m-1\}$, set
$\mathbf{v}_{i+1}^{\perp} := \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_i)^{\perp}}(\mathbf{v}_{i+1})$.
It then follows from our recursive definition of $V(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ that

$$V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \quad = \quad ||\mathbf{v}_1|| \, ||\mathbf{v}_2^{\perp}|| \, \ldots \, ||\mathbf{v}_m^{\perp}||.$$

Since the length of any vector is non-negative, we see that $V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \geq 0$. Moreover, equality holds iff at least one of the vectors $\mathbf{v}_1, \mathbf{v}_2^{\perp}, \ldots, \mathbf{v}_m^{\perp}$ is $\mathbf{0}$.

Proposition 3.1

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 (outline). For each $i \in \{1, \ldots, m-1\}$, set
$\mathbf{v}_{i+1}^{\perp} := \mathrm{proj}_{\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_i)^{\perp}}(\mathbf{v}_{i+1})$.
It then follows from our recursive definition of $V(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ that

$$V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \quad = \quad ||\mathbf{v}_1|| \, ||\mathbf{v}_2^{\perp}|| \; \ldots \; ||\mathbf{v}_m^{\perp}||.$$

Since the length of any vector is non-negative, we see that $V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m) \geq 0$. Moreover, equality holds iff at least one of the vectors $\mathbf{v}_1, \mathbf{v}_2^{\perp}, \ldots, \mathbf{v}_m^{\perp}$ is $\mathbf{0}$. But the latter happens iff the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly dependent (details: Lecture Notes).

Theorem 3.2

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.

Theorem 3.2

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

Theorem 3.2

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, and set $A := \left[\begin{array}{ccc} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{array} \right]$. Then

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

Proof. For each $i \in \{1, \ldots, m\}$, set $A_i := \left[\begin{array}{ccc} \mathbf{a}_1 & \ldots & \mathbf{a}_i \end{array} \right]$. We will prove inductively that for all $i \in \{1, \ldots, m\}$, we have that $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$.

Theorem 3.2

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

Theorem 3.2

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

For $i = 1$, we observe that

$$A_1^T A_1 \quad = \quad \begin{bmatrix} \mathbf{a}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{a}_1 \end{bmatrix} \quad = \quad \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 \end{bmatrix},$$

and consequently,

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

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

Proof (continued). 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 $V_i(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$.

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

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

Proof (continued). 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 $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$. Set

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

By Corollary 1.3 of Lecture Notes 13, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$.

Proof (continued). 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 $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$. Set

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

By Corollary 1.3 of Lecture Notes 13, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$. Since $\mathbf{a}_{i+1}^{||} \in \text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)$, there exist scalars $c_1, \dots, c_i \in \mathbb{R}$ 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). 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 $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$. Set

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

By Corollary 1.3 of Lecture Notes 13, we have that $\mathbf{a}_{i+1} = \mathbf{a}_{i+1}^{||} + \mathbf{a}_{i+1}^{\perp}$. Since $\mathbf{a}_{i+1}^{||} \in \text{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_i)$, there exist scalars $c_1, \ldots, c_i \in \mathbb{R}$ 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 last 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). Then

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

Proof (continued). Then

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

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

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

  $\vdots$

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

Proof (continued). Then

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

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 $E_1, \dots, E_i$ 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, $B_{i+1} = A_{i+1} E_1^T \dots E_i^T$.

Proof (continued). Then

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

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 4.2(c) of Lecture Notes 15, we see that $\det(E_1) = \cdots = \det(E_i) = 1$.

Proof (continued). We now compute:

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

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

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

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

Proof (continued). We now compute:

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

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

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

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

Since $V_{i+1}(\mathbf{a}_1, \ldots, \mathbf{a}_i, \mathbf{a}_{i+1}) \geq 0$ (by Proposition 3.1), we may now take the square root of both sides to obtain

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

This completes the induction. Q.E.D.

Theorem 3.2

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

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

Theorem 3.2

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 3.3

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Then

$$V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) = |\det(\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix})|.$$

Corollary 3.3

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Then

$$V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) \;=\; \left|\det\left(\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix}\right)\right|.$$

Proof.

Corollary 3.3

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Then

$$V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) \quad = \quad |\det(\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix})|.$$

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.

Corollary 3.3

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$. Then

$$V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) \;=\; \left|\det\left(\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix}\right)\right|.$$

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. We now compute:

$$\begin{aligned}
V_n(\mathbf{a}_1, \ldots, \mathbf{a}_n) &= \sqrt{\det(A^T A)} && \text{by Theorem 3.2} \\
&= \sqrt{\det(A^T)\det(A)} \\
&= \sqrt{\det(A)^2} \\
&= |\det(A)|.
\end{aligned}$$

Q.E.D.

Corollary 3.4

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.

Corollary 3.4

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. Set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$ and $A_\sigma := \begin{bmatrix} \mathbf{a}_{\sigma(1)} & \ldots & \mathbf{a}_{\sigma(m)} \end{bmatrix}$. Further, let $E_\sigma = [e_{i,j}]_{m\times m}$ be the matrix given by

$$e_{i,j} = \begin{cases} 1 & \text{if} \quad j = \sigma(i) \\ 0 & \text{if} \quad j \neq \sigma(i) \end{cases}$$

for all $i, j \in \{1, \ldots, m\}$.

Corollary 3.4

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. Set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$ and $A_\sigma := \begin{bmatrix} \mathbf{a}_{\sigma(1)} & \ldots & \mathbf{a}_{\sigma(m)} \end{bmatrix}$. Further, let $E_\sigma = [e_{i,j}]_{m \times m}$ be the matrix given by

$$e_{i,j} = \begin{cases} 1 & \text{if} \quad j = \sigma(i) \\ 0 & \text{if} \quad j \neq \sigma(i) \end{cases}$$

for all $i, j \in \{1, \ldots, m\}$. Note that $\det(E_\sigma) = \operatorname{sgn}(\sigma)$.

Corollary 3.4

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. Set $A := \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$ and $A_\sigma := \begin{bmatrix} \mathbf{a}_{\sigma(1)} & \ldots & \mathbf{a}_{\sigma(m)} \end{bmatrix}$. Further, let $E_\sigma = [e_{i,j}]_{m \times m}$ be the matrix given by

$$e_{i,j} = \begin{cases} 1 & \text{if} \quad j = \sigma(i) \\ 0 & \text{if} \quad j \neq \sigma(i) \end{cases}$$

for all $i, j \in \{1, \ldots, m\}$. Note that $\det(E_\sigma) = \text{sgn}(\sigma)$. Moreover, $A_\sigma^T = E_\sigma A^T$, and so $A_\sigma = A E_\sigma^T$.

Corollary 3.4

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 (continued). But now

$$\begin{aligned}
V_m(\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(m)}) &\overset{\text{Theorem 3.2}}{=} \sqrt{\det(A_\sigma^T A_\sigma)} \\
&= \sqrt{\det(E_\sigma A^T A E_\sigma^T)} \\
&= \sqrt{\det(E_\sigma)\det(A^T A)\det(E_\sigma^T)} \\
&= \sqrt{\det(E_\sigma)\det(A^T A)\det(E_\sigma)} \\
&= \sqrt{\operatorname{sgn}(\sigma)^2 \det(A^T A)} \\
&= \sqrt{\det(A^T A)} \\
&\overset{\text{Theorem 3.2}}{=} V_m(\mathbf{a}_1, \ldots, \mathbf{a}_m)
\end{aligned}$$

Q.E.D.

Corollary 3.5

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 3.5

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 3.5

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 $B := [\ \mathbf{v}_1 \quad \ldots \quad \mathbf{v}_n\ ]$ and $C := [\ A\mathbf{v}_1 \quad \ldots \quad A\mathbf{v}_n\ ] = AB$. Note that A, B, C, AB all belong to $\mathbb{R}^{n \times n}$, and so all four matrices have determinants. We now compute:

$$\begin{aligned}
V(A\mathbf{v}_1, \ldots, A\mathbf{v}_n) &\overset{\text{Theorem 3.2}}{=} \sqrt{\det(C^T C)} \\
&= \sqrt{\det\Big((AB)^T (AB)\Big)} \\
&= \sqrt{\det(B^T A^T AB)} \\
&= \sqrt{\det(B^T)\det(A^T)\det(A)\det(B)} \\
&= \sqrt{\det(A)^2 \det(B^T)\det(B)} \\
&= \sqrt{\det(A)^2\ \det(B^T B)} \\
&= |\det(A)| \sqrt{\det(B^T B)} \\
&\overset{\text{Theorem 3.2}}{=} |\det(A)|\ V_m(\mathbf{v}_1, \ldots, \mathbf{v}_n)
\end{aligned}$$

Q.E.D.

Corollary 3.5

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 3.5

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 3.5 fails, i.e.

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

  - For instance, for $m = 1$ and $n = 2$, we can take $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, so that $A\mathbf{v}_1 = \mathbf{v}_1$.

Corollary 3.5

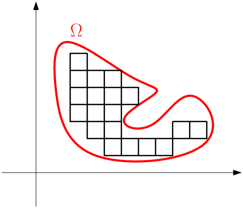
Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$. Then

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

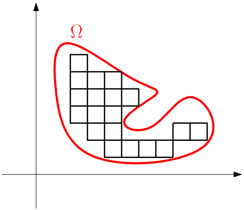
- **Remark:** For $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ $(m \neq n)$ and $A \in \mathbb{R}^{n \times n}$, the formula from Corollary 3.5 fails, i.e.

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

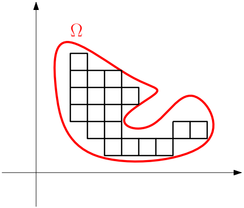
  - For instance, for $m = 1$ and $n = 2$, we can take $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, so that $A\mathbf{v}_1 = \mathbf{v}_1$.
  - Then $V_1(\mathbf{v}_1) = V_1(A\mathbf{v}_1) = ||\mathbf{v}_1|| = 1$, but $\det(A) = 0$, and so $V_1(A\mathbf{v}_1) \neq |\det(A)| \; V_1(\mathbf{v}_1)$.



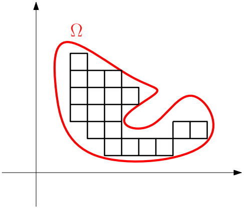
- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.



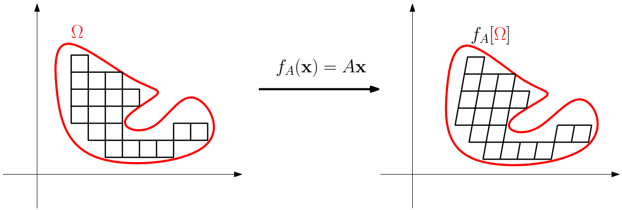
- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
    - The idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes will give us an ever better approximation of the n-volume of Ω that we wish to define.



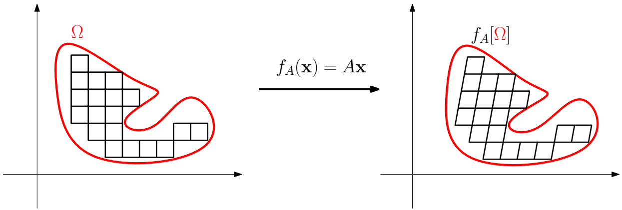
- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
  - The idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes will give us an ever better approximation of the n-volume of Ω that we wish to 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.



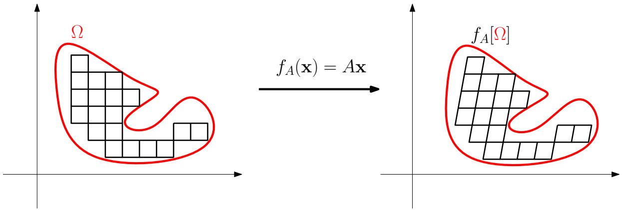
- Suppose that Ω is any object in $\mathbb{R}^n$ for which n-volume $V_n(\Omega)$ can be defined.
  - The idea is that we approximate Ω with ever smaller n-dimensional hypercubes; the sum of n-volumes of those n-hypercubes will give us an ever better approximation of the n-volume of Ω that we wish to 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.
  - Details: omitted!



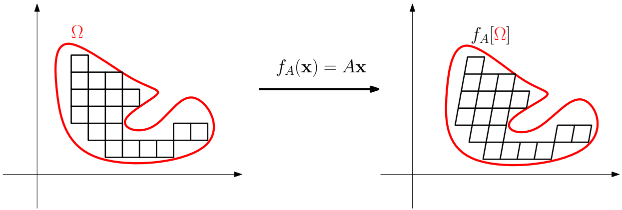
- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.



- 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 3.5, the small n-parallelepipeds that these n-hypercubes get mapped onto via f_A will have volume $|\det(A)|\, V$.



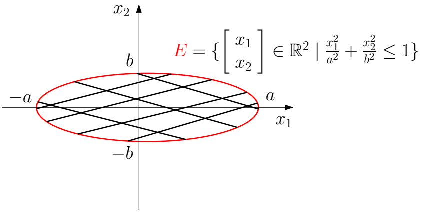
- Now, suppose we are given a matrix $A \in \mathbb{R}^{n \times n}$.
- We consider the linear function $f_A : \mathbb{R}^n \to \mathbb{R}^n$ whose standard matrix is A (i.e. for all $\mathbf{x} \in \mathbb{R}^n$, we have $f_A(\mathbf{x}) = A\mathbf{x}$).
- Then each of the small n-hypercubes gets mapped onto a small n-parallelepiped; if the small n-hypercubes each had volume V, then by Corollary 3.5, 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:

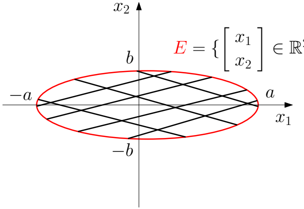
$$V_n(f_A[\Omega]) \quad = \quad |\det(A)|\ V_n(\Omega).$$

Example 3.6

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

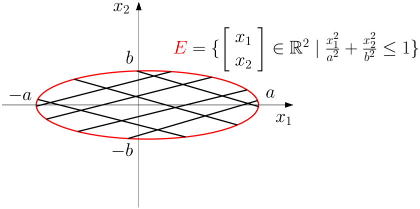
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1.$$





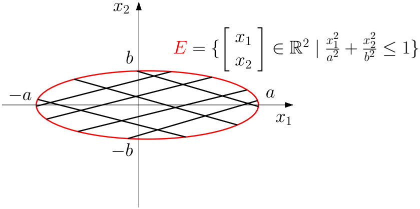
$$E = \{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \mid \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1 \}$$

Solution.



Solution. We need compute the area of the region

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



Solution. We need compute the area of the region

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

Consider the unit disk

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

and the matrix

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

Solution (continued). Let $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation whose standard matrix is A, so that for all $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, we have

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

Solution (continued). Let $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation whose standard matrix is A, so that for all $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, we have

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

We now see that

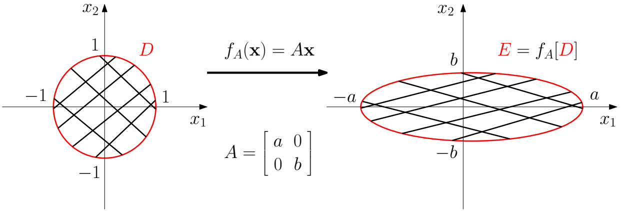
$$\begin{aligned} f_A[D] &= \left\{ f_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1 \right\} \\ &= \left\{ \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1 \right\} \\ &= \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}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

Example 3.6

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: $E = f_A[D]$.

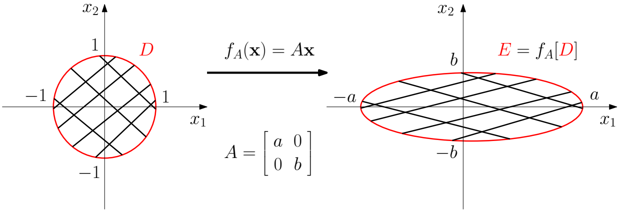


Example 3.6

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

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

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



Therefore, the area of E is

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