Linear Algebra 2: Lecture 17

Irena Penev

Summer 2023

1 The Vandermonde matrix

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

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

called the *Vandermonde matrix*.

Proposition 1.1. *For all positive integers n and real numbers $a_0, a_1, \dots, a_n$, we have that*

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

and consequently, the matrix $V(a_0, a_1, \dots, a_n)$ is invertible if and only if $a_0, a_1, \dots, a_n$ are pairwise distinct.

Proof. By Theorem 5.1 from Lecture Notes 15, we know that a square matrix is invertible if and only if 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.

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

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

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 $\det\Big(V(a_0, a_1, \dots, a_n)\Big) = \prod_{i=1}^{n}\prod_{j=0}^{i-1}(a_i - a_j)$. We must show that the formula is correct for $n + 1$. Fix $a_0, a_1, \dots, a_n, a_{n+1} \in \mathbb{R}$; we will show that

$\det\Big(V(a_0, a_1, \dots, a_n, a_{n+1})\Big) = \prod_{i=1}^{n+1} \prod_{j=0}^{i-1}(a_i - a_j)$. If some two of the numbers $a_0, a_1, \dots, a_n, a_{n+1}$ 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 (by Proposition 2.4 of Lecture Notes 15) 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, we may assume that $a_0, a_1, \dots, a_n, a_{n+1}$ are pairwise distinct.

Set

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

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

$$k \quad := \quad \begin{vmatrix} 1 & 1 & \dots & 1 \\ a_0 & a_1 & \dots & a_n \\ a_0^2 & a_1^2 & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_0^n & a_1^n & \dots & a_n^n \end{vmatrix} \quad \stackrel{(*)}{=} \quad \prod_{i=1}^{n} \prod_{j=0}^{i-1}(a_i - a_j),$$

where (*) follows from the induction hypothesis. Moreover, for each $i \in \{0, 1, \dots, n\}$, $f(a_i)$ is the determinant of a square matrix that has two identical columns and is therefore (by Proposition 2.4 of Lecture Notes 15) equal to zero. Thus, $a_0, a_1, \dots, a_n$ are all roots of the $(n+1)$-th degree polynomial $f(t)$. So, $f(t)$ can be factored as

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

Consequently,

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

This completes the induction. □

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 $\mathbb{F}$. For instance, the polynomial

$$x^2 + 1$$

has real (in fact, rational) coefficients, but it has no real roots. It does, of course, have two complex roots, namely i and $-i$. The Fundamental Theorem of Algebra also fails for $\mathbb{Z}_p$ (for any prime p). In the case of $\mathbb{Z}_2$, it is easy to see that the polynomial

$$x^2 + x + 1$$

has no roots in $\mathbb{Z}_2$ (indeed, neither 0 nor 1 is a root). On the other hand, for prime numbers $p \geq 3$, the polynomial

$$x^{p-1} + 1$$

has no roots in $\mathbb{Z}_p$. Obviously, 0 is not a root of this polynomial. On the other hand, for all $a \in \mathbb{Z}_p \setminus \{0\}$, Fermat's Little Theorem guarantees that $a^{p-1} = 1$, and so $a^{p-1} + 1 = 2 \neq 0$ (since $p \geq 3$), and it follows that a is not a root of the polynomial $x^{p-1} + 1$ either.

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 $a, \alpha_1, \dots, \alpha_\ell$ such that $a \neq 0$ and such that $\alpha_1, \dots, \alpha_\ell$ are pairwise distinct, and positive integers $n_1, \dots, n_\ell$ satisfying $n_1 + \dots + n_\ell = n$, such that

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

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. 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.)

In view of the above, it may be surprising that given arbitrary polynomials $p(x)$ and $q(x)$ with complex coefficients, we can use determinants to determine whether $p(x)$ and $q(x)$ have a common root, i.e. whether there exists a complex number x_0 for which we have $p(x_0) = 0$ and $q(x_0) = 0$ (see Theorem 2.1 below). 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 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*

$$\begin{bmatrix} \underbrace{0 \ \ldots \ 0}_{j-1} & a_m & a_{m-1} & \ldots & a_0 & \underbrace{0 \ \ldots \ 0}_{n-j} \end{bmatrix},$$

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

$$\begin{bmatrix} \underbrace{0 \ \ldots \ 0}_{j-1} & b_n & b_{n-1} & \ldots & b_0 & \underbrace{0 \ \ldots \ 0}_{m-j} \end{bmatrix}.$$

Then $p(x)$ and $q(x)$ have a common complex root if and only if

$$det\left(\begin{bmatrix} P \\ \hdashline Q \end{bmatrix} \right) = 0.$$

Remark#1: For example, if $m = 3$ and $n = 5$, so that

- $p(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0$,
- $q(x) = b_5 x^5 + b_4 x^4 + b_3 x^3 + b_2 x^2 + b_1 x + b_0$,

then we have

$$\begin{bmatrix} P \\ \hdashline Q \end{bmatrix} = \begin{bmatrix} a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & 0 & 0 & 0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 \\ \hdashline b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix}_{8 \times 8}.$$

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 $\mathbb{F}$.

Proof. We begin by proving a claim.

> **Claim.** Polynomials $p(x)$ and $q(x)$ have a common complex root if and only if 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

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

Suppose conversely there exist non-zero polynomials $r(x)$ and $s(x)$ with complex coefficients such that $\deg(r(x)) \leq n-1$, $\deg(s(x)) \leq m-1$, and $r(x)p(x) + s(x)q(x) = 0$. 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 $q(x)$.[1] ♦

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 $s(x) = \sum_{i=0}^{m-1} d_i x^i$ such that $r(x)p(x) + s(x)q(x) = 0$. So, we need to determine if there exist complex numbers $c_0, \dots, c_{n-1}, d_0, \dots, d_{m-1}$ such that 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 such that

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

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. 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. 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.[2] 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

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

[1] We are using the fact that $r(x)p(x)$ and $s(x)q(x)$ have the same roots with the same corresponding multiplicities.

[2] 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.

and we observe that the coefficient matrix A satisfies $A^T = \begin{bmatrix} P \\ \hline Q \end{bmatrix}$.[3]

---

[3]For example, if $m = 3$ and $n = 5$, so that

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

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 below (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).

	c_4		c_3		c_2		c_1		c_0		d_2		d_1		d_0		
x^7	a_3c_4									$+$	b_5d_2					$=$	0
x^6	a_2c_4	$+$	a_3c_3							$+$	b_4d_2	$+$	b_5d_1			$=$	0
x^5	a_1c_4	$+$	a_2c_3	$+$	a_3c_2					$+$	b_3d_2	$+$	b_4d_1	$+$	b_5d_0	$=$	0
x^4	a_0c_4	$+$	a_1c_3	$+$	a_2c_2	$+$	a_3c_1			$+$	b_2d_2	$+$	b_3d_1	$+$	b_4d_0	$=$	0
x^3			a_0c_3	$+$	a_1c_2	$+$	a_2c_1	$+$	a_3c_0	$+$	b_1d_2	$+$	b_2d_1	$+$	b_3d_0	$=$	0
x^2					a_0c_2	$+$	a_1c_1	$+$	a_2c_0	$+$	b_0d_2	$+$	b_1d_1	$+$	b_2d_0	$=$	0
x^1							a_0c_1	$+$	a_1c_0			$+$	b_0d_1	$+$	b_1d_0	$=$	0
x^0									a_0c_0					$+$	b_0d_0	$=$	0

This system, in turn, translates into the following matrix-vector equation:

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

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

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

from Remark#1 following the statement of the theorem.

We now have the following equivalences:

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

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

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

$$\overset{(***)}{\iff} \det\left(\begin{bmatrix} P \\ \hdashline Q \end{bmatrix}\right) = 0,$$

where (*) follows from Corollary 5.1 of Lecture Notes 4, (**) follows from Corollary 4.2 of Lecture Notes 7, and (***) follows from Theorem 5.1 of Lecture Notes 15. This completes the argument. □

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

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

We now have that

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

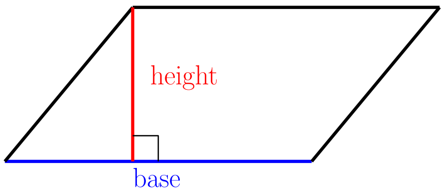
Theorem 2.1 now guarantees that $p(x)$ and $q(x)$ have a common complex root. □

3 Determinants and volume

Throughout this section, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

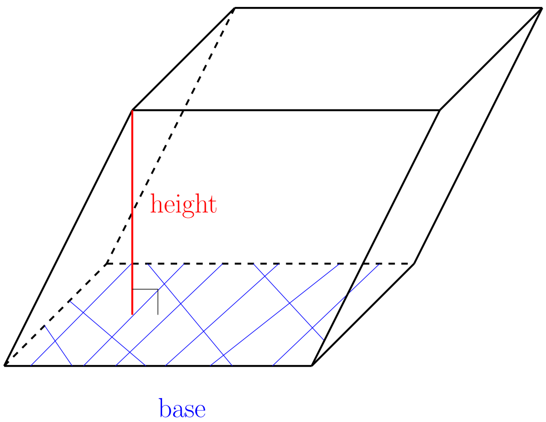
For a parallelogram, we have the familiar formula

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



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

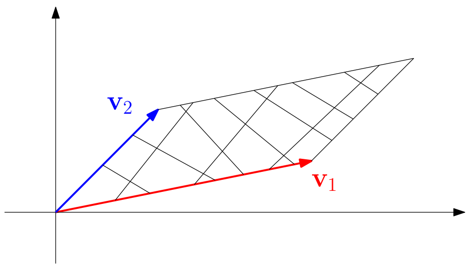
$$\begin{pmatrix} \text{volume of} \\ \text{parallelepiped} \end{pmatrix} = (\text{area of base}) \times (\text{height}).$$



We would now like to generalize this to arbitrary dimensions. 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

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

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 (see the picture below).

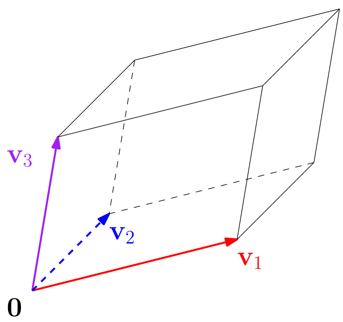


For vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, neither of which is a scalar multiple of each other, the 2-parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2$ is still a parallelogram, but this parallelogram lies in the plane (2-dimensional subspace) $\text{Span}(\mathbf{v}_1, \mathbf{v}_2)$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

$$\begin{aligned} & \Big\{ c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \mid c_1, c_2 \in \mathbb{R},\ 0 \le c_1, c_2 \le 1 \Big\} \\ = & \Big\{ c_1\mathbf{v}_1 + c_2\alpha\mathbf{v}_1 \mid c_1, c_2 \in \mathbb{R},\ 0 \le c_1, c_2 \le 1 \Big\} \\ = & \Big\{ (c_1 + c_2\alpha)\mathbf{v}_1 \mid c_1, c_2 \in \mathbb{R},\ 0 \le c_1, c_2 \le 1 \Big\} \\ = & \Big\{ c(1+\alpha)\mathbf{v}_1 \mid c \in \mathbb{R},\ 0 \le c \le 1 \Big\}, \end{aligned}$$

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

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).



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 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 $\mathbf{v}_{m+1}^{\perp} = \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)^{\perp}}(\mathbf{v}_{m+1})$.[4]

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

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

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). The former readily follows from the definition of volume (see Proposition 3.1 below). However, the latter is not entirely obvious. We prove this in Corollary 3.4, which in fact follows from the formula for volume as the determinant of a certain square matrix (see Theorem 3.2).

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 if and only if* $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ *is a linearly dependent set.*

Proof. 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 if and only if at least one of the vectors $\mathbf{v}_1, \mathbf{v}_2^{\perp}, \ldots, \mathbf{v}_m^{\perp}$ is $\mathbf{0}$. We will show that the latter happens if and only if the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly dependent.

[4]Equivalently: $\mathbf{v}_{m+1}^{\perp} = \mathbf{v}_{m+1} - \text{proj}_{\text{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_m)}(\mathbf{v}_{m+1})$. (The equivalence follows from Corollary 1.3 of Lecture Notes 13.)

Suppose first that at least one of $\mathbf{v}_1, \mathbf{v}_2^{\perp}, \dots, \mathbf{v}_m^{\perp}$ is $\mathbf{0}$. If $\mathbf{v}_1 = \mathbf{0}$, then obviously, $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is linearly dependent. Suppose now that $\mathbf{v}_i^{\perp} = \mathbf{0}$ for some index $i \in \{2, \dots, m\}$. Then $\text{proj}_{\text{Span}(\mathbf{v}_1,\dots,\mathbf{v}_{i-1})}(\mathbf{v}_i) = \mathbf{v}_i - \mathbf{v}_i^{\perp} = \mathbf{v}_i$, and so $\mathbf{v}_i \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$, i.e. $\mathbf{v}_i$ is a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}$. So, $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is linearly dependent.

Suppose now that $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is linearly dependent. Then there exists some $i \in \{1, \dots, m\}$ such that $\mathbf{v}_i$ is a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}$.[5] If $i = 1$, then $\mathbf{v}_1 = \mathbf{0}$. On the other hand, if $i \geq 2$, then $\mathbf{v}_i \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$, and so $\mathbf{v}_i^{\perp} = \mathbf{0}$.[6] $\square$

Theorem 3.2. *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. For each $i \in \{1, \dots, m\}$, set $A_i := \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_i \end{bmatrix}$. We will prove inductively that for all $i \in \{1, \dots, m\}$, we have that $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. Obviously, this is enough, since $A_m = A$.

For $i = 1$, we observe that

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

and consequently,

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

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 $V_i(\mathbf{a}_1, \dots, \mathbf{a}_i) = \sqrt{\det(A_i^T A_i)}$. We must show that $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) = \sqrt{\det(A_{i+1}^T A_{i+1})}$. 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 $c_1, \dots, c_i \in \mathbb{R}$ such that $\mathbf{a}_{i+1}^{||} = c_1\mathbf{a}_1 + \dots + c_i\mathbf{a}_i$, and consequently,

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

[5] This was Problem 2 of HW#7 from Linear Algebra 1 (winter 2022).

[6] Indeed, if $\mathbf{v}_i \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$, then $\text{proj}_{\text{Span}(\mathbf{v}_1,\dots,\mathbf{v}_{i-1})}(\mathbf{v}_i) = \mathbf{v}_i$, and consequently, $\mathbf{v}_i^{\perp} = \mathbf{v}_i - \text{proj}_{\text{Span}(\mathbf{v}_1,\dots,\mathbf{v}_{i-1})}(\mathbf{v}_i) = \mathbf{0}$.

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.

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

Then

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

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

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

  $\vdots$

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

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, $B_{i+1} = A_{i+1} E_1^T \dots E_i^T$. By Theorem 4.2(c) of Lecture Notes 15, we see that $\det(E_1) = \dots = \det(E_i) = 1$.[7] We now compute:

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

where (*) follows from Theorem 1.3 of Lecture Notes 16, and (**) follows from Theorem 2.2 of Lecture Notes 15. But note that $B_{i+1} = \begin{bmatrix} A_i & \vdots & \mathbf{a}_{i+1}^{\perp} \end{bmatrix}$, and so

[7] Indeed, for each $j \in \{1, \dots, i\}$, the matrix E_i is obtained by performing the row operation $R_{i+1} \to R_{i+1} - c_j R_j$ on the identity matrix I_{i+1}, and so by Theorem 4.2(c) of Lecture Notes 15, we have that $\det(E_j) = \det(I_{i+1}) = 1$.

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

where in (*), we used the fact that $\mathbf{a}_{i+1}^\perp$ is orthogonal to the columns of A, and so $A^T \mathbf{a}_{i+1}^\perp = \mathbf{0}$.[8] We now compute:

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

where (*) follows by Laplace expansion along the last column, (**) follows from the induction hypothesis, and (***) follows from the definition of $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1})$. Since $V_{i+1}(\mathbf{a}_1, \dots, \mathbf{a}_i, \mathbf{a}_{i+1}) \geq 0$ (by Proposition 3.1), we may now take the square root of both sides to obtain

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

This completes the induction. □

[8] Indeed,

$$A^T \mathbf{a}_{i+1}^\perp = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_i^T \end{bmatrix} \mathbf{a}_{i+1}^\perp = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_{i+1} \\ \vdots \\ \mathbf{a}_1 \cdot \mathbf{a}_{i+1} \end{bmatrix}.$$

Since $\mathbf{a}_{i+1}^\perp \in \mathrm{Span}(\mathbf{a}_1, \dots, \mathbf{a}_i)^\perp$, we see that $\mathbf{a}_1 \cdot \mathbf{a}_{i+1}^\perp = \cdots = \mathbf{a}_i \cdot \mathbf{a}_{i+1}^\perp = 0$, and it follows that $A^T \mathbf{a}_{i+1}^\perp = \mathbf{0}$.

The following corollary of Theorem 3.2 gives a geometric interpretation of the determinant.

Corollary 3.3. *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:

$$\begin{aligned} V_n(\mathbf{a}_1, \dots, \mathbf{a}_n) &= \sqrt{\det(A^T A)} && \text{by Theorem 3.2} \\ &= \sqrt{\det(A^T)\det(A)} && \text{by Theorem 1.3 of Lecture Notes 16} \\ &= \sqrt{\det(A)^2} && \text{by Theorem 2.2 of Lecture Notes 15} \\ &= |\det(A)|. \end{aligned}$$

This completes the argument. □

Our next corollary states that the m-volume of an m-parallelepiped remains unchanged if we merely change the order of the vectors that determine our m-parallelepiped.

Corollary 3.4. *Let* $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ *and* $\sigma \in S_m$. *Then* $V_m(\mathbf{a}_1, \dots, \mathbf{a}_m) = V_m(\mathbf{a}_{\sigma(1)}, \dots, \mathbf{a}_{\sigma(m)})$.

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

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

for all $i, j \in \{1, \dots, m\}$.[9] Note that $\det(E_\sigma) = \text{sgn}(\sigma)$.[10] Moreover, $A_\sigma^T = E_\sigma A^T$,[11] and so $A_\sigma = A E_\sigma^T$. But now

[9]Note that E_σ has exactly one 1 in every row and in every column, and zeroes everywhere else. Matrices with this property are called *permutation matrices.*

[10]This follows immediately from the definition of the determinant (details?).

[11]Details?

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

where both instances of (*) follow from Theorem 3.2, (**) follows from Theorem 1.3 of Lecture Notes 16, and (***) follows from Theorem 2.2 of Lecture Notes 15. □

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*

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

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, C, AB all belong to $\mathbb{R}^{n\times n}$, and so all four matrices have determinants. We now compute:

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

where both instances of (*) follow from Theorem 3.2, both instances of (**) follow from Theorem 1.3 of Lecture Notes 16, and (***) follows from Theorem 2.2 of Lecture Notes 15. □

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.

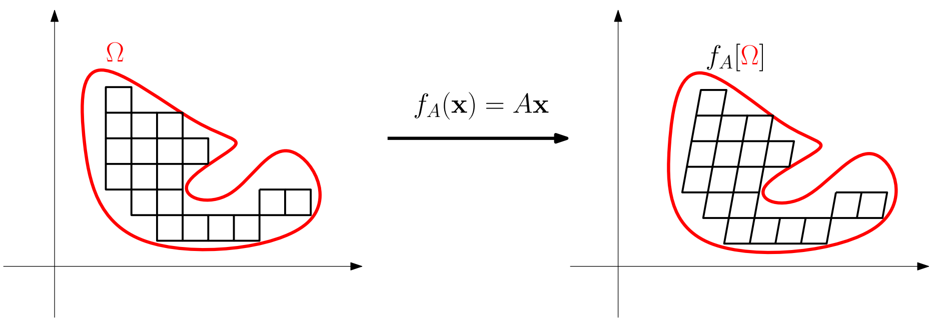
$$V_m(A\mathbf{v}_1, \ldots, A\mathbf{v}_m) \;\;\not\asymp\;\; |\det(A)|\, V_m(\mathbf{v}_1, \ldots, \mathbf{v}_m).$$

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. 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 $A \in \mathbb{R}^{n \times n}$. 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:

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

For the case $n = 2$, see the picture below.



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. We need compute the area of the region

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

Consider the unit disk

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

and the matrix

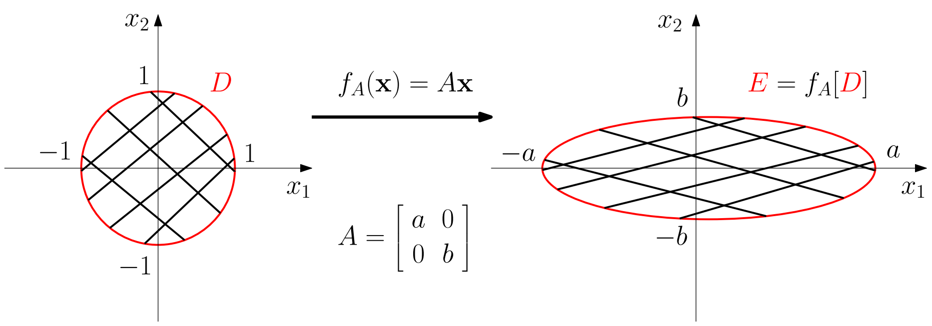
$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Let $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear 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

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



Therefore, the area of E is

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

□