Lecture Notes: Matrix Algebra
Part D, appendix: Determinants and Volumes

Peter J. Hammond

revised 2023 September 15th;
typeset from `matrixAlgDapp23.tex`

Outline

Determinants and "Volume" in Two or Three Dimensions

Eric Renault has asked me to explain
the change of variables formula
for the kind of multiple integral that occurs in econometrics.

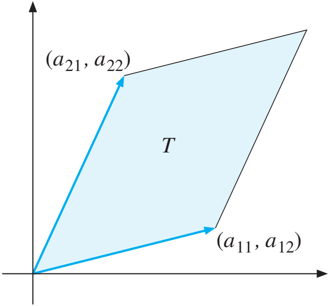
It helps to understand first
the relationship between volume and determinant.

To start the discussion, see the two subsections,
both entitled "A Geometric Interpretation", which can be found:

- in Section 13.1 on pp. 497–498
  and in Section 13.2 on pp. 502–503 of EMEA6;
- with one figure less,
  in Sections 16.1 and 16.2 of previous editions of EMEA.

The next few slides are adapted from Section 13.1 of EMEA6.

Determinants and "Volume" in Two Dimensions Illustrated

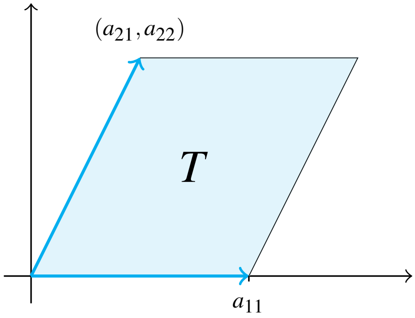


Suppose we represent the two rows of the matrix $\mathbf{A} = (a_{ij})_{2\times 2}$ as the two 2-vectors shown in the figure.

Then $|\mathbf{A}|$ equals the shaded area T of the parallelogram.

But if we swap the two rows,
the determinant becomes minus this shaded area.

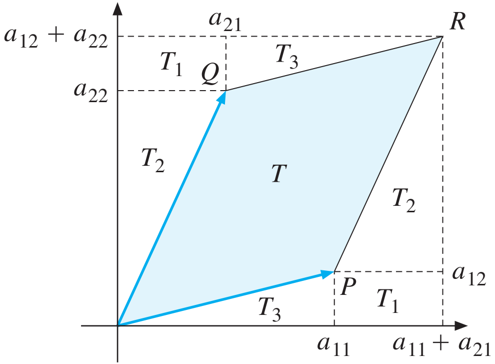
Determinants and Area in Two Dimensions: Special Case



This figure shows the area corresponding to the 2×2 determinant in the special case when $a_{12} = 0$.

Here the area of T is simply the product of its base and height, which is just $a_{11}a_{22} = |\mathbf{A}|$ because $a_{12}a_{21} = 0$.

Determinants and Area in Two Dimensions: General Case



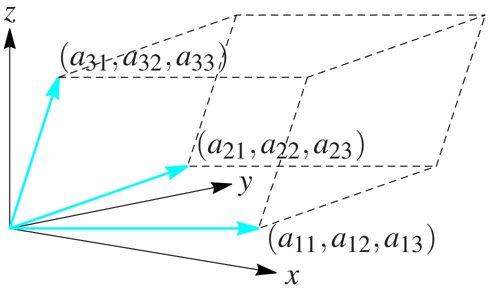
The area of the large rectangle is $(a_{11} + a_{21})(a_{12} + a_{22})$.

It is the sum of the relevant area T
and the two triangular or rectangular areas marked.

This sum of areas equals $2T_1 + 2T_2 + 2T_3 + T$
where $T_1 = a_{12}a_{21}$, $T_2 = \frac{1}{2}a_{21}a_{22}$, and $T_3 = \frac{1}{2}a_{11}a_{12}$.

By elementary algebra, it follows that $T = a_{11}a_{22} - a_{21}a_{12} = |\mathbf{A}|$.

Determinants and Volume in Three Dimensions



The rows of a 3×3 determinant correspond
to three different 3-vectors represented in the diagram.

Instead of a cuboid whose six faces are all rectangles,
one has a parallelepiped whose six faces are all parallelograms.

The volume of this parallelepiped must equal
the absolute value of the determinant $|\mathbf{A}|$.

Higher dimensions require new definitions.

Outline

Summing Sets in n Dimensions: Definition

Definition

Let $\{S_k \mid k = 1, 2, \ldots, r\}$ be any collection of r subsets of $\mathbb{R}^n$.
Define the vector sum $S := \sum_{k=1}^r S_k$ of these sets as the set

$$S := \{\mathbf{x} \in \mathbb{R}^n \mid \forall k \in \mathbb{N}_r; \exists \mathbf{x}^k \in S_k : \mathbf{x} = \sum_{k=1}^r \mathbf{x}^k\}$$

of all possible selections $\mathbb{N}_r \ni k \mapsto \mathbf{x}^k \in \mathbb{R}^n$
of r n-vectors from the family of sets,
or equivalently, from the correspondence $\mathbb{N}_r \ni k \mapsto\!\!\!\to \mathbf{S}_k \subset \mathbb{R}^n$.

Exercise

Once you have learned about convex sets,
prove that the vector sum of convex subsets of $\mathbb{R}^n$ is convex.

Two Examples of Summing Sets in Two Dimensions

Example

In two dimensions, the rows of the general 2×2 matrix $\mathbf{A}$
are vectors described by the two line intervals in $\mathbb{R}^2$ given by

$$I_1 := [(0,0),(a_{11},a_{12})] \quad \text{and} \quad I_2 := [(0,0),(a_{21},a_{22})]$$

The respective sets of endpoints for these two intervals are

$$Z_1 := \{(0,0),(a_{11},a_{12})\} \quad \text{and} \quad Z_2 := \{(0,0),(a_{21},a_{22})]\}$$

The vector sum of these two sets is the set

$$Z_1 + Z_2 = \{(0,0),(a_{11},a_{12}),(a_{21},a_{22}),(a_{11}+a_{21},a_{12}+a_{22})\}$$

consisting of the four vertices of a parallelogram in $\mathbb{R}^2$.

The vector sum $I_1 + I_2$ of the convex intervals, however,
is a convex set equal to the whole parallelogram.

An Exercise in Summing Sets in Three Dimensions

Exercise

Extend the previous example from two dimensions to three.
In particular:

1. *Show that the rows of the general* 3×3 *matrix* $\mathbf{A}$ *determine three line intervals* I_k $(k = 1, 2, 3)$ *that each join the origin* $\mathbf{0}$ *of* $\mathbb{R}^3$ *to a point* $\mathbf{a}_k \in \mathbb{R}^3$ *representing one row of the matrix.*
2. *If* Z_k *denotes, for each* $k = 1, 2, 3$, *the pair set* $\{\mathbf{0}, \mathbf{a}_k\}$ *of endpoints, show that the vector sum* $Z_1 + Z_2 + Z_3$ *is a set consisting of eight vertices of a parallelipiped.*
3. *Show that the vector sum* $I_1 + I_2 + I_3$ *of the three line intervals is the parallelipiped that is the convex hull of the eight points in* $Z_1 + Z_2 + Z_3$.

An Iterative Definition of Parallelipiped Volumes

Definition

Consider the $n \times n$ matrix $\mathbf{A}$
whose rows are the n row n-vectors $(\mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_n)$.

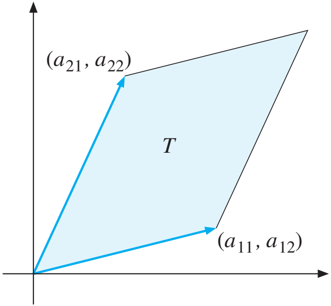
First, define $\text{vol}_1(\mathbf{A}) := \|\mathbf{a}_1\|$ and then, for each $k = 2, 3, \ldots n$, define each successive volume by $\text{vol}_k(\mathbf{A}) = d_k(\mathbf{A}) \, \text{vol}_{k-1}(\mathbf{A})$ where

$$d_k(\mathbf{A}) := \min_{\lambda_1, \lambda_2, \ldots, \lambda_{k-1}} \|\mathbf{a}_k - \sum\nolimits_{r=1}^{k-1} \lambda_r \mathbf{a}_r\|$$

denotes the (shortest) distance of $\mathbf{a}_k$ from the linear subspace spanned by the first $k-1$ rows of $\mathbf{A}$. □

The idea is that the volume $\text{vol}_k(\mathbf{A})$
of the parallelipiped $\sum_{r=1}^{k} [\mathbf{0}, \mathbf{a}_r]$ spanned by the first k rows of $\mathbf{A}$
is equal to the product of:
(i) the volume $\text{vol}_{k-1}(\mathbf{A})$ of the "base" parallelipiped;
(ii) the "height" $d_k(\mathbf{A})$ of the vector $\mathbf{a}_k$
above the linear space spanned by the first $k-1$ rows of $\mathbf{A}$.

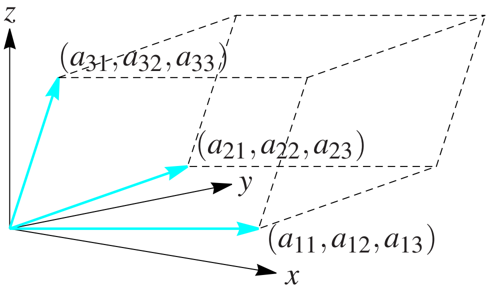
The Two-Dimensional Case



When $\mathbf{A}$ is 2×2, the area T
is given by $\text{vol}_2(\mathbf{A}) = d_2(\mathbf{A})\,\text{vol}_1(\mathbf{A})$ where:

1. $\text{vol}_1(\mathbf{A})$ is the length $\|\mathbf{a}_1\|$ of $\mathbf{A}$'s first row $\mathbf{a}_1 = (a_{11}, a_{12})$;
2. $d_2(\mathbf{A})$ equals the (minimum) distance $\min_{\lambda_1} \|\mathbf{a}_2 - \lambda_1 \mathbf{a}_1\|$
   between the point $\mathbf{a}_2 = (a_{21}, a_{22})$
   that corresponds to the second row
   and the closest point to $\mathbf{a}_2$ on the line through $\mathbf{a}_1$.

The Three-Dimensional Case



When $\mathbf{A}$ is 3×3, its volume
is given by $\text{vol}_3(\mathbf{A}) = d_3(\mathbf{A})\,\text{vol}_2(\mathbf{A})$ where:

1. $\text{vol}_2(\mathbf{A})$ is the area of the parallelogram spanned by the first two rows of $\mathbf{A}$;
2. $d_3(\mathbf{A})$ equals the distance $\min_{\lambda_1,\lambda_2} \|\mathbf{a}_3 - \lambda_1 \mathbf{a}_1 - \lambda_2 \mathbf{a}_2\|$ between the third row vector $\mathbf{a}_3 = (a_{31}, a_{32}, a_{33})$ and the closest point to $\mathbf{a}_3$ on the plane spanned by the first two rows of $\mathbf{A}$.

Outline

Determinants and Volumes of Diagonal Matrices

Recall the notation $\mathbf{e}_i$ for the ith row vector of the canonical basis, whose jth component, for $j = 1, 2, \ldots, n$, equals the Kronecker delta δ_{ij}.

Then $\lambda_i \mathbf{e}_i$ is the ith row of the $n \times n$ diagonal matrix $\Lambda = \mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Exercise

Prove by induction that, for $k = 1, 2, \ldots, n$, one has $d_k(\Lambda) = |\lambda_k|$ and

$$\mathrm{vol}_k(\Lambda) = \prod_{i=1}^{k} |\lambda_i| = |\det \Lambda|$$

Remark

The answer to this exercise accords with the geometric intuition that the volume of an n-dimensional cuboid, whose faces are all $n-1$-dimensional cuboids, equals the product of the lengths $\|\lambda_i \mathbf{e}_i\| = |\lambda_i| \|\mathbf{e}_i\| = |\lambda_i|$ of all its n row vectors.

Diagonal Matrices with Permuted Columns

Suppose that the $n \times n$ diagonal matrix $\Lambda = \mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is post-multiplied by the sign corrected determinant preserving permutation matrix $\hat{\mathbf{P}} = (-1)^{\operatorname{sgn} \pi}(\delta_{i\pi(j)})_{n \times n}$
for some permutation $\pi \in \Pi_n$ which is applied to the columns of $\mathbf{I}_n$.

Then each (i, j) element of the matrix $\Lambda\hat{\mathbf{P}}$ is given by

$$(\Lambda\hat{\mathbf{P}})_{ij} = (-1)^{\operatorname{sgn} \pi} \sum\nolimits_{h=1}^{n} \lambda_i \delta_{ih} \delta_{h\pi(j)} = (-1)^{\operatorname{sgn} \pi} \lambda_i \delta_{i\pi(j)}$$

This is a matrix whose only non-zero element in the ith row occurs when $\pi(j) = 1$, so the ith row is $(-1)^{\operatorname{sgn} \pi} \lambda_i \mathbf{e}_{\pi^{-1}(i)}$.

Arguing as in the case of a diagonal matrix,
it follows that for each $k = 1, 2, \ldots, n$,
one has $d_k(\Lambda) = |\lambda_k|$ and $\operatorname{vol}_k(\Lambda) = \prod_{i=1}^{k} |\lambda_i|$.

This proves that $\operatorname{vol}_n(\Lambda\hat{\mathbf{P}}) = \operatorname{vol}_n(\Lambda) = \prod_{i=1}^{k} |\lambda_i|$.

Outline

General Invariance to Row Permutations

Lemma

Given any $n \times n$ matrix $\mathbf{A}$ and any $n \times n$ permutation matrix $\mathbf{P}$, one has $\mathrm{vol}_n(\mathbf{PA}) = \mathrm{vol}_n(\mathbf{A})$.

Proof.

Pre-multiplying the matrix $\mathbf{A}$ by the permutation matrix $\mathbf{P}$ permutes the rows of $\mathbf{A}$.

This pre-multiplication merely reorders the terms in the vector sum $\sum_{r=1}^{n} [\mathbf{0}, \mathbf{a}_r] \subset \mathbb{R}^n$ of line intervals.

But this vector sum corresponds to the paralleliped that is spanned by all the n rows of $\mathbf{A}$.

Geometrically, therefore, since the paralleliped defined by the sum $\sum_{r=1}^{n} [\mathbf{0}, \mathbf{a}_r]$ must be unchanged, so is its volume. □

Outline

The Effect of an Elementary Row Operation, I

Recall the notation $\mathbf{E}_{r+\alpha s}$ for the $n \times n$ matrix that represents the elementary and determinant preserving row operation of adding α times row s to row r.

We start with the special case when $r = n$, so a multiple of row s is added to the last row n.

Lemma

Given any $n \times n$ matrix $\mathbf{A}$ and any row $s < n$, for all $k = 1, 2, \ldots, n-1$ one has $\text{vol}_k(\mathbf{E}_{n+\alpha s}\mathbf{A}) = \text{vol}_k(\mathbf{A})$.

Proof.

Note that the operation $\mathbf{E}_{n+\alpha s}$ affects only the last row n of any $n \times n$ matrix.

From the definition of $\text{vol}_k(\mathbf{A})$ for any $n \times n$ matrix $\mathbf{A}$, it follows that, for $k = 1, 2, \ldots, n-1$, one has $\text{vol}_k(\mathbf{E}_{n+\alpha s}\mathbf{A}) = \text{vol}_k(\mathbf{A})$. □

The Effect of an Elementary Row Operation, II

Lemma

Given any $n \times n$ matrix $\mathbf{A}$ and any row $s < n$, one has $d_n(\mathbf{E}_{n+\alpha s}\mathbf{A}) = d_n(\mathbf{A})$.

Proof.

Because $(\mathbf{a}_n + \alpha \mathbf{a}_s) - \sum_{i=1}^{n-1} \lambda_i \mathbf{a}_i = \mathbf{a}_n - \sum_{i=1}^{n-1}(\lambda_i - \delta_{is}\alpha)\mathbf{a}_i$,
one has $d_n(\mathbf{E}_{n+\alpha s}\mathbf{A}) = \min_{(\lambda_1,\lambda_2,\ldots,\lambda_{n-1})} \|\mathbf{a}_n - \sum_{i=1}^{n-1}(\lambda_i - \delta_{is}\alpha)\mathbf{a}_i\|$.

Suppose that $\min_{(\lambda_1,\lambda_2,\ldots,\lambda_{n-1})} \|\mathbf{a}_n - \sum_{i=1}^{n-1} \lambda_i \mathbf{a}_i\|$
is achieved at $(\lambda_1^*, \ldots \lambda_{n-1}^*)$, so $d_n(\mathbf{A}) = \|\mathbf{a}_n - \sum_{i=1}^{n-1} \lambda_i^* \mathbf{a}_i\|$.

Then $\min_{(\lambda_1,\lambda_2,\ldots,\lambda_{n-1})} \|\mathbf{a}_n - \sum_{i=1}^{n-1}(\lambda_i - \delta_{is}\alpha)\mathbf{a}_i\|$ is achieved
when for each $i = 1, 2, \ldots, n-1$
one has $\lambda_i - \delta_{is}\alpha = \lambda_i^*$ and so $\lambda_i = \lambda_i^* + \delta_{is}\alpha$.

This proves that $d_n(\mathbf{E}_{n+\alpha s}\mathbf{A}) = \|\mathbf{a}_n - \sum_{i=1}^{n-1} \lambda_i^* \mathbf{a}_i\| = d_n(\mathbf{A})$. □

The Effect of an Elementary Row Operation, III

Proposition

Given any $n \times n$ matrix $\mathbf{A}$ and any row $s < n$, one has $\text{vol}_n(\mathbf{E}_{n+\alpha s}\mathbf{A}) = \text{vol}_n(\mathbf{A})$.

Proof.

The definitions of $\text{vol}_n(\mathbf{A})$ and of $\text{vol}_n(\mathbf{E}_{n+\alpha s}\mathbf{A})$ imply that

$$\begin{aligned} \text{vol}_n(\mathbf{E}_{n+\alpha s}\mathbf{A}) &= d_n(\mathbf{E}_{n+\alpha s}\mathbf{A})\,\text{vol}_{n-1}(\mathbf{E}_{n+\alpha s}\mathbf{A}) \\ \text{vol}_n(\mathbf{A}) &= d_n(\mathbf{A})\,\text{vol}_{n-1}(\mathbf{A}) \end{aligned}$$

But the two previous lemmas imply that $\text{vol}_{n-1}(\mathbf{E}_{n+\alpha s}\mathbf{A}) = \text{vol}_{n-1}(\mathbf{A})$ and $d_n(\mathbf{E}_{n+\alpha s}\mathbf{A}) = d_n(\mathbf{A})$.

Thus, the right-hand sides of the two displayed equations are equal, implying that the left-hand sides must also be equal. □

The Effect of an Elementary Row Operation, IV

Proposition

Given any $n \times n$ matrix $\mathbf{A}$ and any elementary row operation $\mathbf{E}_{r+\alpha s}$, one has $\text{vol}_n(\mathbf{E}_{r+\alpha s}\mathbf{A}) = \text{vol}_n(\mathbf{A})$.

Proof.

Note that for all $r \in \mathbb{N}_{n-1}$ and all $s \neq r$,
one has $\mathbf{E}_{r+\alpha s} = \hat{\mathbf{T}}_{r\leftrightarrow n}\mathbf{E}_{n+\alpha s}\hat{\mathbf{T}}_{r\leftrightarrow n}$
where $\hat{\mathbf{T}}_{r\leftrightarrow n} = -\mathbf{T}_{r\leftrightarrow n}$ is the sign-adjusted version
of the row operation $\mathbf{T}_{r\leftrightarrow n}$ that transposes rows r and n.

Because transpositions are permutations that preserve the volume, it follows from the previous lemma that

$$\begin{aligned}\text{vol}_n(\mathbf{A}) = \text{vol}_n(\hat{\mathbf{T}}_{r\leftrightarrow n}\mathbf{A}) &= \text{vol}_n(\mathbf{E}_{n+\alpha s}\hat{\mathbf{T}}_{r\leftrightarrow n}\mathbf{A}) \\ &= \text{vol}_n(\hat{\mathbf{T}}_{r\leftrightarrow n}\mathbf{E}_{n+\alpha s}\hat{\mathbf{T}}_{r\leftrightarrow n}\mathbf{A}) = \text{vol}_n(\mathbf{E}_{r+\alpha s}\mathbf{A}) \quad \square\end{aligned}$$

Outline

The Case of Linearly Dependent Rows

Proposition

The $n \times n$ matrix $\mathbf{A}$ has linearly dependent rows
if and only if $\text{vol}_n(\mathbf{A}) = 0$.

Proof.

If the rows of $\mathbf{A}$ are linearly dependent,
then after permuting its rows $\mathbf{a}_1, \ldots, \mathbf{a}_n$ if necessary,
there must exist $n-1$ real constants $\lambda_1, \ldots, \lambda_{n-1}$
such that $\mathbf{a}_n = \sum_{i=1}^{n-1} \lambda_i \mathbf{a}_i$.
But then $d_n(\mathbf{A}) = \|\mathbf{a}_n - \sum_{i=1}^{n-1} \lambda_i \mathbf{a}_i\| = 0$ and so $\text{vol}_n(\mathbf{A}) = 0$.

Conversely, if $\text{vol}_n(\mathbf{A}) = 0$, then there exists $k \in \mathbb{N}_n$
and $k-1$ real constants $\lambda_1, \ldots, \lambda_{k-1}$
such that $d_k(\mathbf{A}) = \|\mathbf{a}_k - \sum_{i=1}^{k-1} \lambda_i \mathbf{a}_i\| = 0$.

So $\mathbf{a}_k = \sum_{i=1}^{k-1} \lambda_i \mathbf{a}_i$,
implying that the rows of $\mathbf{A}$ are linearly dependent. □

Determinants and Volumes: Main Theorem

Theorem

Given any $n \times n$ matrix $\mathbf{A}$,
the volume $\text{vol}_n(\mathbf{A})$ of the parallelipiped spanned by its rows equals $|\det \mathbf{A}|$, the absolute value of the determinant.

Determinants and Volumes: Proof of Main Theorem

Proof.

In the case when $\mathbf{A}$ is singular, so its rows are linearly dependent, the previous Proposition implies that $\text{vol}_n(\mathbf{A}) = \det \mathbf{A} = 0$.

Otherwise there exists a maximal diagonalization $\mathbf{R}\mathbf{A}\hat{\mathbf{P}} = \mathbf{D}$, where the diagonal matrix $\mathbf{D}$ is non-singular.

Furthermore $\mathbf{R}$ is invertible, and $\mathbf{R}^{-1}$ is the product of matrices which are either sign corrected permutations,
or elementary row operations of the form $\mathbf{E}_{r+\alpha s}$.

Finally $\hat{\mathbf{P}}$ is an invertible permutation, so $\mathbf{A} = \mathbf{R}^{-1}\mathbf{D}\hat{\mathbf{P}}^{-1}$, where both $\mathbf{R}^{-1}$ and $\hat{\mathbf{P}}^{-1}$ are determinant preserving.

From previous results, in the non-singular case it follows that

$$\text{vol}_n(\mathbf{A}) = \text{vol}_n(\mathbf{D}\hat{\mathbf{P}}^{-1}) = \text{vol}_n(\mathbf{D}) = |\det \mathbf{D}| = |\det \mathbf{A}| > 0 \quad \square$$