

Lecture Notes: Matrix Algebra

Part C: Determinants and Pivoting

Peter J. Hammond

revised 2025 September 17th
typeset from `matrixAlgC25.tex`

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

Determinants of Order 2: Definition

Consider again the pair of linear equations

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2\end{aligned}$$

with its associated coefficient matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Let us define the number $D := a_{11}a_{22} - a_{21}a_{12}$.

We saw earlier that, provided $D \neq 0$, the two simultaneous equations have a unique solution given by

$$x_1 = \frac{1}{D}(b_1a_{22} - b_2a_{12}), \quad x_2 = \frac{1}{D}(b_2a_{11} - b_1a_{21})$$

This number D is called the **determinant** of the matrix \mathbf{A} .

It is denoted by either $\det(\mathbf{A})$, or more concisely, by $|\mathbf{A}|$.

Determinants of Order 2: Simple Rule

Thus, for any 2×2 matrix \mathbf{A} , its determinant D is

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

For this special case of **order 2** determinants, a simple rule is:

1. multiply the diagonal elements together;
2. multiply the off-diagonal elements together;
3. subtract the product of the off-diagonal elements from the product of the diagonal elements.

Exercise

Show that the determinant satisfies

$$|\mathbf{A}| = a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{21}a_{12} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

Transposing the Rows or Columns

Example

Consider the two 2×2 matrices $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\mathbf{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Note that \mathbf{T} is orthogonal.

Also, one has $\mathbf{AT} = \begin{pmatrix} b & a \\ d & c \end{pmatrix}$ and $\mathbf{TA} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$.

Here \mathbf{T} is a **transposition matrix** which interchanges:

- (i) the **columns** of \mathbf{A} in \mathbf{AT} ;
- (ii) the **rows** of \mathbf{A} in \mathbf{TA} .

Evidently $|\mathbf{T}| = -1$ and $|\mathbf{TA}| = |\mathbf{AT}| = (bc - ad) = -|\mathbf{A}|$.

So interchanging **either** the two rows **or** the two columns of \mathbf{A} (but not both) changes the sign of $|\mathbf{A}|$.

Sign Adjusted Transpositions

Example

Next, consider the following three 2×2 matrices:

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\mathbf{T}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Note that, like \mathbf{T} , the matrix $\hat{\mathbf{T}}$ is orthogonal.

Here one has $\mathbf{A}\hat{\mathbf{T}} = \begin{pmatrix} b & -a \\ d & -c \end{pmatrix}$ and $\hat{\mathbf{T}}\mathbf{A} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}$.

Evidently $|\hat{\mathbf{T}}| = 1$ and $|\hat{\mathbf{T}}\mathbf{A}| = |\mathbf{A}\hat{\mathbf{T}}| = (ad - bc) = |\mathbf{A}|$.

The same is true of its transpose (and inverse) $\hat{\mathbf{T}}^T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

This key property makes both $\hat{\mathbf{T}}$ and $\hat{\mathbf{T}}^T$ **sign adjusted** versions of the transposition matrix \mathbf{T} .

Cramer's Rule in the 2×2 Case

Using determinant notation, the solution to the equations

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2\end{aligned}$$

can be written in the alternative form

$$x_1 = \frac{1}{D} \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \quad x_2 = \frac{1}{D} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

This accords with **Cramer's rule**,

which says that the solution to $\mathbf{Ax} = \mathbf{b}$ is the vector $\mathbf{x} = (x_i)_{i=1}^n$ each of whose components x_i is the fraction with:

1. denominator equal to the determinant D of the coefficient matrix \mathbf{A} (**provided**, of course, that $D \neq 0$);
2. numerator equal to the determinant of the matrix $[\mathbf{A}_{-i}/\mathbf{b}]$ formed from \mathbf{A} by excluding its i th column, then replacing it with the \mathbf{b} vector of right-hand side elements, while keeping all the columns in their original order.

Determinants of Order 3: Definition

Determinants of order 3 can be calculated from those of order 2 according to the formula

$$\begin{aligned} |\mathbf{A}| &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= \sum_{j=1}^3 (-1)^{1+j} a_{1j} |\mathbf{C}_{1j}| \end{aligned}$$

where, for $j = 1, 2, 3$, the 2×2 matrix \mathbf{C}_{1j} is the $(1, j)$ -**cofactor** obtained by removing both row 1 and column j from the matrix \mathbf{A} .

The result is the following sum

$$\begin{aligned} |\mathbf{A}| &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ &\quad - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned}$$

of $3! = 6$ terms, each the product of 3 elements chosen so that each row and each column is represented just once.

Determinants of Order 3: Cofactor Expansion

The determinant expansion

$$\begin{aligned} |\mathbf{A}| = & a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ & - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned}$$

is very symmetric, suggesting (correctly)
that the cofactor expansion **along the first row** (a_{11}, a_{12}, a_{13})

$$|\mathbf{A}| = \sum_{j=1}^3 (-1)^{1+j} a_{1j} |\mathbf{C}_{1j}|$$

gives the same answer as the other cofactor expansions

$$|\mathbf{A}| = \sum_{j=1}^3 (-1)^{r+j} a_{rj} |\mathbf{C}_{rj}| = \sum_{i=1}^3 (-1)^{i+s} a_{is} |\mathbf{C}_{is}|$$

along, respectively:

- ▶ **the r th row** (a_{r1}, a_{r2}, a_{r3})
- ▶ **the s th column** (a_{1s}, a_{2s}, a_{3s})

Determinants of Order 3: Alternative Expressions

One way of condensing the notation

$$|\mathbf{A}| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

is to reduce it to $|\mathbf{A}| = \sum_{\pi \in \Pi_3} \text{sgn}(\pi) \prod_{i=1}^3 a_{i\pi(i)}$
for the sign function $\Pi_3 \ni \pi \mapsto \text{sgn}(\pi) \in \{-1, +1\}$.

The six values of $\text{sgn}(\pi)$ can be read off as

$$\begin{aligned} \text{sgn}(\pi^{123}) &= +1; & \text{sgn}(\pi^{132}) &= -1; & \text{sgn}(\pi^{231}) &= +1; \\ \text{sgn}(\pi^{213}) &= -1; & \text{sgn}(\pi^{312}) &= +1; & \text{sgn}(\pi^{321}) &= -1. \end{aligned}$$

Exercise

Verify these values for each of the six $\pi \in \Pi_3$ by:

1. calculating the number of inversions directly;
2. expressing each π as the product of transpositions, and then counting these.

Sarrus's Rule: Diagram

An alternative way to evaluate determinants **only** of order 3 is to add two new columns that repeat the first and second columns:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Then add lines/arrows going up to the right or down to the right, as shown below

$$\begin{array}{cccccc} a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\ & \searrow & & \nearrow & & \nearrow & & \searrow & \\ a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\ & \nearrow & & \searrow & & \searrow & & \nearrow & \\ a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \end{array}$$

Note that some pairs of arrows in the middle cross each other.

Sarrus's Rule Defined

Now:

1. multiply along the three lines falling to the right, then sum these three products, to obtain

$$a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32}$$

2. multiply along the three lines rising to the right, then sum these three products, giving the sum a minus sign, to obtain

$$-a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31}$$

The sum of all six terms exactly equals the earlier formula for $|\mathbf{A}|$.

Note that this method, known as **Sarrus's rule**, **does not generalize** to determinants of order higher than 3.

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

The Determinant Function

For each $n \in \mathbb{N}$, let $\mathcal{M}_{n \times n}$ denote the domain of $n \times n$ matrices. It is evidently a copy of the space $\mathbb{R}^{n \times n} = \mathbb{R}^{n^2}$.

Definition

For all $n \in \mathbb{N}$, the **determinant function**

$$\mathcal{M}_{n \times n} \ni \mathbf{A} \mapsto \det \mathbf{A} = |\mathbf{A}| := \sum_{\pi \in \Pi_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{i\pi(i)} \in \mathbb{R}$$

specifies the determinant $|\mathbf{A}|$ of each $n \times n$ matrix \mathbf{A} as a function of its n row vectors $(\mathbf{a}_i^\top)_{i=1}^n = \left((a_{ij})_{j=1}^n \right)_{i=1}^n$. □

Here the multiplier $\operatorname{sgn}(\pi)$ attached to each product of n terms can be regarded as the **sign adjustment** associated with the permutation $\pi \in \Pi_n$.

A Four-Part Exercise

Exercise

Use the formula on the previous slide to calculate $|\mathbf{A}|$ when \mathbf{A} is:

1. the general 2×2 matrix $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$;

2. any 3×3 matrix of the form $\begin{pmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$
with at most one non-zero term off the diagonal;

3. any 3×3 matrix of the form $\begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$
with at most two non-zero terms off the diagonal;

4. any $n \times n$ diagonal matrix $\mathbf{D} = \mathbf{diag}(d_1, d_2, \dots, d_n)$.

Functions of the Rows of a Matrix

For a general natural number $n \in \mathbb{N}$, consider any function

$$\mathcal{M}_{n \times n} \ni \mathbf{A} \mapsto D(\mathbf{A}) = D\left(\langle \mathbf{a}_i^\top \rangle_{i=1}^n\right) \in \mathbb{R}$$

whose domain $\mathcal{M}_{n \times n}$ is the set of all $n \times n$ matrices \mathbf{A} , regarded as a collection of n row vectors $\langle \mathbf{a}_i^\top \rangle_{i=1}^n$.

Notation: For each fixed $r \in \mathbb{N}_n$, let $D(\mathbf{A}_{-r}/\mathbf{b}_r^\top)$

denote the new value $D(\mathbf{a}_1^\top, \dots, \mathbf{a}_{r-1}^\top, \mathbf{b}_r^\top, \mathbf{a}_{r+1}^\top, \dots, \mathbf{a}_n^\top)$

of the function $\mathbf{A} \mapsto D(\mathbf{A})$ after the r th row \mathbf{a}_r^\top of the matrix \mathbf{A} has been replaced by the new row vector $\mathbf{b}_r^\top \in \mathbb{R}^n$, with all the other $n - 1$ rows remaining fixed.

Row Multilinearity

Definition

The function $\mathcal{M}_{n \times n} \ni \mathbf{A} \mapsto D(\mathbf{A})$ of the n rows $\langle \mathbf{a}_i^\top \rangle_{i=1}^n$ of \mathbf{A} is **(row) multilinear** just in case, for each row number $i \in \mathbb{N}_n$, for each pair $\mathbf{b}_i^\top, \mathbf{c}_i^\top \in \mathbb{R}^n$ of new versions of row i , and for each pair of scalars $\lambda, \mu \in \mathbb{R}$, one has

$$D(\mathbf{A}_{-i}/\lambda\mathbf{b}_i^\top + \mu\mathbf{c}_i^\top) = \lambda D(\mathbf{A}_{-i}/\mathbf{b}_i^\top) + \mu D(\mathbf{A}_{-i}/\mathbf{c}_i^\top) \quad \square$$

Formally, the mapping $\mathbb{R}^n \ni \mathbf{a}_i^\top \mapsto D(\mathbf{A}_{-i}/\mathbf{a}_i^\top) \in \mathbb{R}$ is required to be linear, for fixed each row $i \in \mathbb{N}_n$.

That is, D is a linear function of the i th row vector \mathbf{a}_i^\top on its own, when all the other rows \mathbf{a}_h^\top ($h \neq i$) are fixed.

Determinants are Row Multilinear

Theorem

For all $n \in \mathbb{N}$, the earlier definition implies that the determinant mapping

$$\mathcal{M}_{n \times n} \ni \mathbf{A} \mapsto |\mathbf{A}| := \sum_{\pi \in \Pi_n} \text{sgn}(\pi) \prod_{i=1}^n a_{i\pi(i)} \in \mathbb{R}$$

is a row multilinear function of its n row vectors $(\mathbf{a}_i^\top)_{i=1}^n$.

Proof.

For each fixed row $r \in \mathbb{N}_n$, the determinant mapping satisfies

$$\begin{aligned} & \det(\mathbf{A}_{-r} / \lambda \mathbf{b}_r^\top + \mu \mathbf{c}_r^\top) \\ &= \sum_{\pi \in \Pi_n} \text{sgn}(\pi) (\lambda b_{r\pi(r)} + \mu c_{r\pi(r)}) \prod_{i \neq r} a_{i\pi(i)} \\ &= \sum_{\pi \in \Pi_n} \text{sgn}(\pi) \left[\lambda b_{r\pi(r)} \prod_{i \neq r} a_{i\pi(i)} + \mu c_{r\pi(r)} \prod_{i \neq r} a_{i\pi(i)} \right] \\ &= \lambda \det(\mathbf{A}_{-r} / \mathbf{b}_r^\top) + \mu \det(\mathbf{A}_{-r} / \mathbf{c}_r^\top) \end{aligned}$$

This confirms multilinearity. □

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

The Eight Basic Rules: Background and Explanation

EMEA is an acronym for our undergraduate textbook *Essential Mathematics for Economic Analysis*.

EMEA n is an abbreviation for the n edition.

Some of you may have used EMEA5, but EMEA6 did appear in 2021.

The eight rules labelled 1–8 here appear as Rules A–H in:

- ▶ Section 16.4 of EMEA5
— see Theorem 16.4.1 on page 636;
- ▶ Section 13.4 of EMEA6
— see Theorem 13.4.1 on page 509.

Of the eight rules:

- ▶ Rule 6 plays a key role when discussing pivoting subsequently;
- ▶ Rules 1–6 and Rule 8 will be confirmed here;
- ▶ a proof of Rule 7, which uses pivoting in a key way, is deferred until the next Segment D.

The Eight Basic Rules: Statement

Let $|\mathbf{A}|$ denote the determinant of any $n \times n$ matrix \mathbf{A} .

1. $|\mathbf{A}| = 0$ if all the elements in a row (or column) of \mathbf{A} are 0.
2. $|\mathbf{A}^\top| = |\mathbf{A}|$, where \mathbf{A}^\top is the transpose of \mathbf{A} .
3. If all the elements in a single row (or column) of \mathbf{A} are multiplied by a scalar α , so is its determinant.
4. If two rows (or two columns) of \mathbf{A} are interchanged, the determinant changes sign, but not its absolute value.
5. If two of the rows (or columns) of \mathbf{A} are proportional, then $|\mathbf{A}| = 0$.
6. The value of the determinant of \mathbf{A} is unchanged if any multiple of one row (or one column) is added to a **different** row (or column) of \mathbf{A} .
7. The determinant of the product $|\mathbf{AB}|$ of two $n \times n$ matrices equals the product $|\mathbf{A}| \cdot |\mathbf{B}|$ of their determinants.
8. If α is any scalar, then $|\alpha\mathbf{A}| = \alpha^n |\mathbf{A}|$.

Verifying the Transpose Rule 2

The transpose rule 2 is very useful: it implies that for any statement \mathcal{S} about how $|\mathbf{A}|$ depends on the **rows** of \mathbf{A} , there is an equivalent “transpose” statement \mathcal{S}^\top about how $|\mathbf{A}|$ depends on the **columns** of \mathbf{A} .

Exercise

Verify Rule 2 directly for 2×2 and then for 3×3 matrices.

Proof of Rule 2 The expansion formula implies that

$$|\mathbf{A}| = \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{i\pi(i)} = \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{j=1}^n a_{\pi^{-1}(j)j}$$

But we proved earlier that $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$.

Also $a_{\pi^{-1}(j)j} = a_{j\pi^{-1}(j)}^\top$ by definition of transpose.

Hence, because $\pi \leftrightarrow \pi^{-1}$ is a bijection on the set Π , the expansion formula with π replaced by π^{-1}

implies that $|\mathbf{A}| = \sum_{\pi^{-1} \in \Pi} \operatorname{sgn}(\pi^{-1}) \prod_{j=1}^n a_{j\pi^{-1}(j)}^\top = |\mathbf{A}^\top|$. □

Verifying the Alternation Rule 4

Recall the notation $\tau_{r,s}$ for the transposition of $r, s \in \mathbb{N}_n$.

Let $\mathbf{A}_{r \leftrightarrow s}$ denote the matrix that results from applying $\tau_{r,s}$ to the rows of the matrix \mathbf{A} — i.e., interchanging rows r and s .

Theorem

Given any $n \times n$ matrix \mathbf{A} and any transposition $\tau_{r,s}$, one has $\det \mathbf{A}_{r \leftrightarrow s} = -\det \mathbf{A}$.

Proof.

Write τ for $\tau_{r,s}$. Then, because $\pi \leftrightarrow \tau^{-1} \circ \pi$ is a bijection on Π_n and $\text{sgn}(\tau^{-1} \circ \pi) = -\text{sgn}(\pi)$ for all $\pi \in \Pi_n$, we have

$$\begin{aligned} \det \mathbf{A}_{r \leftrightarrow s} &= \sum_{\pi \in \Pi_n} \text{sgn}(\pi) \prod_{i=1}^n a_{\tau(i), \pi(i)} \\ &= \sum_{\pi \in \Pi_n} \text{sgn}(\pi) \prod_{i=1}^n a_{i, (\tau^{-1} \circ \pi)(i)} \\ &= - \sum_{\pi \in \Pi_n} \text{sgn}(\tau^{-1} \circ \pi) \prod_{i=1}^n a_{i, (\tau^{-1} \circ \pi)(i)} \\ &= - \sum_{\pi \in \Pi_n} \text{sgn}(\pi) \prod_{i=1}^n a_{i, \pi(i)} = -\det \mathbf{A} \quad \square \end{aligned}$$

The Duplication Rule, and Rule 8

The following duplication rule is a special case of Rule 5.

Proposition

If two different rows r and s of \mathbf{A} are equal, then $|\mathbf{A}| = 0$.

Proof.

Suppose that rows r and s of \mathbf{A} are equal.

Then $\mathbf{A}_{r \leftrightarrow s} = \mathbf{A}$, and so $|\mathbf{A}_{r \leftrightarrow s}| = |\mathbf{A}|$.

Yet the alternation Rule 4 implies that $|\mathbf{A}_{r \leftrightarrow s}| = -|\mathbf{A}|$.

Hence $|\mathbf{A}| = -|\mathbf{A}|$, implying that $|\mathbf{A}| = 0$. □

Rule 8: $|\alpha \mathbf{A}| = \alpha^n |\mathbf{A}|$ for any $\alpha \in \mathbb{R}$.

Proof.

The expansion formula implies that

$$\begin{aligned} |\alpha \mathbf{A}| &= \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^n (\alpha a_{i\pi(i)}) \\ &= \alpha^n \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{i\pi(i)} = \alpha^n |\mathbf{A}| \quad \square \end{aligned}$$

First Implications of Multilinearity: Rules 1 and 3

Recall the notation $\mathbf{A}_{-r}/\mathbf{b}_r^\top$ for the matrix that results after the r th row \mathbf{a}_r^\top of \mathbf{A} has been replaced by \mathbf{b}_r^\top .

With this notation, the matrix $\mathbf{A}_{-r}/\alpha\mathbf{a}_r^\top$ is the result of replacing the r th row \mathbf{a}_r^\top of \mathbf{A} by $\alpha\mathbf{a}_r^\top$.

That is, it is the result of multiplying the r th row \mathbf{a}_r^\top of \mathbf{A} by the scalar α .

Rule 3: If all the elements in a single row of \mathbf{A} are multiplied by a scalar α , so is its determinant.

Proof.

By multilinearity one has $|\mathbf{A}_{-r}/\alpha\mathbf{a}_r^\top| = \alpha|\mathbf{A}_{-r}/\mathbf{a}_r^\top| = \alpha|\mathbf{A}|$. □

Rule 1: $|\mathbf{A}| = 0$ if all the elements in a row of \mathbf{A} are 0.

Proof.

This follows from putting $\alpha = 0$ in Rule 3. □

More Implications of Multilinearity: Rule 5

Rule 5: If two rows of \mathbf{A} are proportional, then $|\mathbf{A}| = 0$.

Proof.

Suppose that $\mathbf{a}_r^\top = \alpha \mathbf{a}_s^\top$ where $r \neq s$.

Then $|\mathbf{A}| = |\mathbf{A}_{-r}/(\alpha \mathbf{a}_s^\top)_r| = \alpha |\mathbf{A}_{-r}/(\mathbf{a}_s^\top)_r| = 0$ by duplication. \square

More Implications of Multilinearity: Rule 6

Rule 6: $|\mathbf{A}|$ is unchanged if any multiple of one row is added to a different row of \mathbf{A} .

Proof.

For the matrix $\mathbf{A}_{-r}/(\mathbf{a}_r^\top + \alpha\mathbf{a}_s^\top)_r$, where α times row s of \mathbf{A} has been added to row r , row multilinearity implies that

$$|\mathbf{A}_{-r}/(\mathbf{a}_r^\top + \alpha\mathbf{a}_s^\top)_r| = |\mathbf{A}_{-r}/(\mathbf{a}_r^\top)_r| + \alpha|\mathbf{A}_{-r}/(\mathbf{a}_s^\top)_r|$$

But $\mathbf{A}_{-r}/(\mathbf{a}_r^\top)_r = \mathbf{A}$ and $\mathbf{A}_{-r}/(\mathbf{a}_s^\top)_r$ has a copy of row s in row r . By the duplication rule, it follows that $|\mathbf{A}_{-r}/(\mathbf{a}_s^\top)_r| = 0$ and so

$$\begin{aligned} |\mathbf{A}_{-r}/(\mathbf{a}_r^\top + \alpha\mathbf{a}_s^\top)_r| &= |\mathbf{A}_{-r}/(\mathbf{a}_r^\top)_r| + \alpha|\mathbf{A}_{-r}/(\mathbf{a}_s^\top)_r| \\ &= |\mathbf{A}| + 0 = |\mathbf{A}| \end{aligned}$$

□

Verification of the Product Rule 7: Diagonal Case

Recall that Rule 7 is the **product rule** stating that $|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|$.

Later we will use pivoting to verify this rule for general matrices.

Here we consider the special case when the first matrix \mathbf{A} is the $n \times n$ diagonal matrix $\mathbf{D} = \mathbf{diag}(d_1, d_2, \dots, d_n)$.

Proposition

For any $n \times n$ matrix \mathbf{B} , one has $|\mathbf{DB}| = |\mathbf{D}| \cdot |\mathbf{B}| = (\prod_{k=1}^n d_k) |\mathbf{B}|$.

Proof.

First, note that \mathbf{DB} is the matrix that results from simultaneously multiplying each row $i = 1, 2, \dots, n$ of \mathbf{B} by the corresponding diagonal element d_i of \mathbf{D} .

By Rule 3 applied n times,

the result of all these n simultaneous multiplications

is that the determinant is multiplied by the n -fold product $\prod_{i=1}^n d_i$.

So $|\mathbf{DB}| = \prod_{i=1}^n d_i \cdot |\mathbf{B}|$.

But \mathbf{D} is diagonal, so $|\mathbf{D}| = \prod_{i=1}^n d_i$, and $|\mathbf{DB}| = |\mathbf{D}| \cdot |\mathbf{B}|$. □

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

Triangular Matrices: Definition

Definition

A square matrix is **upper** (resp. **lower**) **triangular** if all its non-zero off diagonal elements are above and to the right (resp. below and to the left) of the diagonal — i.e., in the upper (resp. lower) triangle bounded by the principal diagonal.

- ▶ The elements of an upper triangular matrix **U** satisfy $(\mathbf{U})_{ij} = 0$ whenever $i > j$.
- ▶ The elements of a lower triangular matrix **L** satisfy $(\mathbf{L})_{ij} = 0$ whenever $i < j$.

Products of Upper Triangular Matrices

Theorem

The product $\mathbf{W} = \mathbf{UV}$ of any two upper triangular matrices \mathbf{U}, \mathbf{V} is upper triangular,

with diagonal elements $w_{ii} = u_{ii}v_{ii}$ ($i = 1, \dots, n$) equal to the product of the corresponding diagonal elements of \mathbf{U}, \mathbf{V} .

Proof.

Given any two upper triangular $n \times n$ matrices \mathbf{U} and \mathbf{V} , one has $u_{ik}v_{kj} = 0$ unless both $i \leq k$ and $k \leq j$.

So the elements $(w_{ij})^{n \times n}$ of their product $\mathbf{W} = \mathbf{UV}$ satisfy

$$w_{ij} = \begin{cases} \sum_{k=i}^j u_{ik}v_{kj} & \text{if } i \leq j \\ 0 & \text{if } i > j \end{cases}$$

Hence $\mathbf{W} = \mathbf{UV}$ is upper triangular.

Finally, when $j = i$ the above sum collapses to just one term, and $w_{ii} = u_{ii}v_{ii}$ for $i = 1, \dots, n$. □

Triangular Matrices: Exercises

Exercise

Prove that the transpose:

1. \mathbf{U}^\top of any upper triangular matrix \mathbf{U} is lower triangular;
2. \mathbf{L}^\top of any lower triangular matrix \mathbf{L} is upper triangular.

Exercise

Consider the matrix $\mathbf{E}_{r+\alpha q}$ that represents the elementary row operation of adding a multiple of α times row q to row r , with $r \neq q$.

Under what conditions is $\mathbf{E}_{r+\alpha q}$ (i) upper triangular? (ii) lower triangular?

Hint: Apply the row operation to the identity matrix \mathbf{I} .

Answer: (i) if and only if $q > r$; (ii) if and only if $q < r$.

Products of Lower Triangular Matrices

Theorem

The product of any two lower triangular matrices is lower triangular.

Proof.

Given any two lower triangular matrices \mathbf{L} , \mathbf{M} , taking transposes shows that $(\mathbf{LM})^\top = \mathbf{M}^\top \mathbf{L}^\top = \mathbf{U}$, where the product \mathbf{U} is upper triangular, as the product of upper triangular matrices.

Hence $\mathbf{LM} = \mathbf{U}^\top$ is lower triangular, as the transpose of an upper triangular matrix. □

Determinants of Triangular Matrices

Theorem

The determinant of any $n \times n$ upper triangular matrix \mathbf{U} equals the product of all the elements on its principal diagonal.

Proof.

Recall the expansion formula $|\mathbf{U}| = \sum_{\pi \in \Pi} \text{sgn}(\pi) \prod_{i=1}^n u_{i\pi(i)}$ where Π denotes the set of permutations on $\{1, 2, \dots, n\}$.

Because \mathbf{U} is upper triangular, one has $u_{i\pi(i)} = 0$ unless $i \leq \pi(i)$.

So $\prod_{i=1}^n u_{i\pi(i)} = 0$ unless $i \leq \pi(i)$ for all $i = 1, 2, \dots, n$.

But the identity ι is the only permutation $\pi \in \Pi$ that satisfies $i \leq \pi(i)$ for all $i \in \mathbb{N}_n$.

Because $\text{sgn}(\iota) = +1$, the expansion reduces to the single term

$$|\mathbf{U}| = \text{sgn}(\iota) \prod_{i=1}^n u_{i\iota(i)} = \prod_{i=1}^n u_{ii}$$

This is the product of the n diagonal elements, as claimed. □

Invertible Triangular Matrices

Similarly $|\mathbf{L}| = \prod_{i=1}^n \ell_{ii}$ for any lower triangular matrix \mathbf{L} .

Evidently:

Corollary

A triangular matrix (upper or lower) has a non-zero determinant, and so is invertible, if and only if no element on its principal diagonal is 0.

The Product Rule 7 for Triangular Determinants

Example

Let \mathbf{A} and \mathbf{B} be $n \times n$ matrices where:

(i) either both are upper triangular; or (ii) both are lower triangular.

We showed earlier that the product $\mathbf{C} = \mathbf{AB}$ is also triangular.

We also showed that diagonal elements $c_{ii} = a_{ii}b_{ii}$ of the product equal the product of the diagonal elements of \mathbf{A} and \mathbf{B} .

Also, recall that the determinant of a triangular matrix, either upper or lower, equals the product of its diagonal elements.

It follows that

$$\begin{aligned} |\mathbf{C}| &= \prod_{i=1}^n c_{ii} = \prod_{i=1}^n a_{ii}b_{ii} \\ &= \left(\prod_{i=1}^n a_{ii} \right) \left(\prod_{i=1}^n b_{ii} \right) = |\mathbf{A}| \cdot |\mathbf{B}| \end{aligned}$$

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

Three Simultaneous Equations

Consider the following system of three simultaneous equations in three unknowns, which depends upon two “exogenous” constants a and b :

$$\begin{aligned}x + y - z &= 1 \\x - y + 2z &= 2 \\x + 2y + az &= b\end{aligned}$$

It can be expressed, using an augmented 3×4 matrix, as :

$$\begin{array}{ccc|c}1 & 1 & -1 & 1 \\1 & -1 & 2 & 2 \\1 & 2 & a & b\end{array}$$

Perhaps even more useful is the doubly augmented 3×7 matrix:

$$\begin{array}{ccc|c|ccc}1 & 1 & -1 & 1 & 1 & 0 & 0 \\1 & -1 & 2 & 2 & 0 & 1 & 0 \\1 & 2 & a & b & 0 & 0 & 1\end{array}$$

whose last 3 columns are those of the 3×3 identity matrix \mathbf{I}_3 .

Pivoting: First Step

Start with the doubly augmented 3×7 matrix:

$$\begin{array}{ccc|c|ccc} 1 & 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 2 & a & b & 0 & 0 & 1 \end{array}$$

First, **pivot** about the element in row 1 and column 1 to eliminate or “zeroize” the other elements of column 1.

This **elementary row operation** requires us to subtract row 1 from both rows 2 and 3.

It is equivalent to pre-multiplying by $\mathbf{E}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$.

Note: this is the result of applying the same row operations to \mathbf{I}_3 .

The resulting 3×7 matrix is:

$$\begin{array}{ccc|c|ccc} 1 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 3 & 1 & -1 & 1 & 0 \\ 0 & 1 & a+1 & b-1 & -1 & 0 & 1 \end{array}$$

Pivoting: Second Step

Including another copy of the identity matrix at the end gives:

$$\begin{array}{ccc|c|ccc|ccc} 1 & 1 & -1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 3 & 1 & -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & a+1 & b-1 & -1 & 0 & 1 & 0 & 0 & 1 \end{array}$$

Next, we pivot about the element in row 2 and column 2.

Specifically, add half the second row

to both the first and third rows to obtain:

$$\begin{array}{ccc|c|ccc|ccc} 1 & 0 & \frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 1 & \frac{1}{2} & 0 \\ 0 & -2 & 3 & 1 & -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & a + \frac{5}{2} & b - \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} & 1 & 0 & \frac{1}{2} & 1 \end{array}$$

Again, the pivot operation is equivalent to pre-multiplying,

here by the matrix $\mathbf{E}_2 = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$,

which is the result of applying the same row operation to \mathbf{I}_3 .

The Augmented Matrix After Downward Pivoting

The resulting augmented matrix is

$$\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & -2 & 3 & 1 \\ 0 & 0 & a + \frac{5}{2} & b - \frac{1}{2} \end{array}$$

whose top two rows and columns form a 2×2 diagonal matrix.

Thus, the two steps of pivoting have eliminated:

- ▶ x , the 1st variable, from both the 2nd and 3rd equations;
- ▶ y , the 2nd variable, from both the 1st and 3rd equations.

To conclude, we need to treat two different cases:

Case 1: if $a + \frac{5}{2} \neq 0$, the 3×3 coefficient matrix is upper triangular, with a non-zero diagonal;

Case 2: if $a + \frac{5}{2} = 0$, the 3×3 coefficient matrix takes the partitioned form $\begin{pmatrix} \mathbf{D}_{2 \times 2} & \mathbf{B}_{2 \times 1} \\ \mathbf{0}_{1 \times 2} & 0 \end{pmatrix}$ where $\mathbf{D}_{2 \times 2}$ is a 2×2 diagonal matrix.

Case 1: Third Pivoting Step

In case 1 when $a + \frac{5}{2} \neq 0$, we will complete solving the equation by pivoting a third time about the 3, 3 element to reach a diagonal matrix whose diagonal terms are non-zero.

Starting with the augmented matrix

$$\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & -2 & 3 & 1 \\ 0 & 0 & a + \frac{5}{2} & b - \frac{1}{2} \end{array}$$

and with $c = 1/(a + \frac{5}{2})$, we pivot about the 3, 3 element by adding:
(i) $-\frac{1}{2} \cdot c$ times row 3 to row 1; (ii) $-3 \cdot c$ times row 3 to row 2.

The final augmented matrix that results

from this last pivot operation is

$$\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{2} - \frac{1}{2}c(b - \frac{1}{2}) \\ 0 & -2 & 0 & 1 - 3c(b - \frac{1}{2}) \\ 0 & 0 & a + \frac{5}{2} & b - \frac{1}{2} \end{array}$$

The coefficient matrix has become diagonal, with all its diagonal elements non-zero.

This makes the resulting equations easy to solve.

Case 1: Solution of the Equation System

The three pivoting operations we have completed have reduced the equation system to

$$\begin{aligned}x &= \frac{3}{2} - \frac{1}{2}c(b - \frac{1}{2}) \\-2y &= 1 - 3c(b - \frac{1}{2}) \\(a + \frac{5}{2})z &= b - \frac{1}{2}\end{aligned}$$

Because $c = 1/(a + \frac{5}{2})$, this gives the unique solution

$$x = \frac{3}{2} - \frac{1}{2}c(b - \frac{1}{2}), \quad y = -\frac{1}{2} + \frac{3}{2}c(b - \frac{1}{2}), \quad z = c(b - \frac{1}{2})$$

Case 2: Pivoting Concludes after Two Steps

In **case 2**, when $a + \frac{5}{2} = 0$, after two steps of pivoting,

the augmented matrix has been reduced to
$$\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & -2 & 3 & 1 \\ 0 & 0 & 0 & b - \frac{1}{2} \end{array}$$

This takes the partitioned form $\begin{pmatrix} \mathbf{D}_{2 \times 2} & \mathbf{B}_{2 \times 1} \\ \mathbf{0}_{1 \times 2} & 0 \end{pmatrix}$ where:

$\mathbf{D}_{2 \times 2}$ is a 2×2 diagonal matrix with non-zero diagonal elements;

$\mathbf{B}_{2 \times 1}$ is a 2×1 matrix, or a 2×1 column vector $\mathbf{b}_{2 \times 1}$.

Case 2: Dependent Equations

In **case 2A**, when $b \neq \frac{1}{2}$, neither the last equation, nor the system as a whole, has any solution.

In **case 2B**, when $b = \frac{1}{2}$, the third equation is redundant.

Then the augmented matrix for the remaining two equations

reduces to
$$\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & -2 & 3 & 1 \end{array}$$

The associated equation system has a general solution

$$x = \frac{3}{2} - \frac{1}{2}z \quad \text{and} \quad y = \frac{3}{2}z - \frac{1}{2}$$

where z is an arbitrary scalar.

In particular, there is a one-dimensional set of solutions along the unique straight line in \mathbb{R}^3 that passes through both:

(i) $(\frac{3}{2}, -\frac{1}{2}, 0)$, when $z = 0$; (ii) $(1, 1, 1)$, when $z = 1$.

Outline

Determinants: Introduction

Determinants of Orders 2 and 3
The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants
Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations
Permutation and Transposition Matrices

Row and Column Operations

Definition

For each $m, n \in \mathbb{N}$, let $\mathcal{M}_{m \times n}$ denote the family of all $m \times n$ matrices.

▶ A **row operation**

is a mapping $\mathcal{M}_{m \times n} \ni \mathbf{X} \mapsto \mathbf{E}\mathbf{X} \in \mathcal{M}_{m \times n}$ represented by an $m \times m$ matrix \mathbf{E} that **pre-multiplies** (or multiplies on the **left**) any $\mathbf{X} \in \mathcal{M}_{m \times n}$.

▶ Similarly, a **column operation**

is a mapping $\mathcal{M}_{m \times n} \ni \mathbf{Y} \mapsto \mathbf{Y}\mathbf{E} \in \mathcal{M}_{m \times n}$ represented by an $n \times n$ matrix \mathbf{E} that **post-multiplies** (or multiplies on the **right**) any $\mathbf{Y} \in \mathcal{M}_{m \times n}$. □

Given any $k \in \mathbb{N}$, note that \mathbf{E} is the $k \times k$ matrix which results from applying either the row or the column operation represented by \mathbf{E} to the identity matrix \mathbf{I}_k .

Three Kinds of Elementary Row Operation

The pivoting operations used in the previous example are examples of row operations that belong to a special category of **elementary row operation**.

Textbooks (including ours) usually specify the following three kinds of elementary row operation $\mathbf{A} \mapsto \mathbf{EA}$:

1. **rescale** one row $r \in \mathbb{N}_m$ by multiplying it by a scalar $\alpha \in \mathbb{R} \setminus \{0\}$;
2. **swap** two rows $r, s \in \mathbb{N}_m$ with $r \neq s$;
3. **pivot** by adding one rescaled row s to another row r .

In the next few slides we will describe each of these in detail.

There are obviously similar elementary column operations.

Type 1: Rescaling One Row

For each $r \in \mathbb{N}_m$ and each scalar $\alpha \in \mathbb{R} \setminus \{0\}$, let the $m \times m$ matrix $\mathbf{E}_{r \times \alpha}$ represent the **rescaling** operation that, when applied to any $m \times n$ matrix \mathbf{A} , multiplies row r of \mathbf{A} by α . The elements of $\mathbf{E}_{r \times \alpha}$, which are those of $\mathbf{E}_{r \times \alpha} \mathbf{I}_m$, are given by

$$(\mathbf{E}_{r \times \alpha})_{ij} = \begin{cases} \delta_{ij} & \text{if } i \neq r \\ \alpha \delta_{ij} & \text{if } i = r \end{cases} \quad \text{for all } (i, j) \in \mathbb{N}_m \times \mathbb{N}_m$$

This implies that $\mathbf{E}_{r \times \alpha} = \mathbf{diag}(1, \dots, 1, \alpha, 1, \dots, 1)$, which differs from \mathbf{I}_m in at most the (r, r) element.

Suppose $m = n$, so the determinant $|\mathbf{A}|$ is well defined.

Then Rule 3 for determinants implies that $|\mathbf{E}_{r \times \alpha} \mathbf{A}| = \alpha |\mathbf{A}|$.

Putting $\mathbf{A} = \mathbf{I}_m$ in this equality implies that

$$|\mathbf{E}_{r \times \alpha}| = |\mathbf{E}_{r \times \alpha} \mathbf{I}_m| = \alpha |\mathbf{I}_m| = \alpha$$

Only in the trivial case when $\alpha = 1$ and so $\mathbf{E}_{r \times \alpha} = \mathbf{I}_m$ does $\mathbf{E}_{r \times \alpha}$ “preserve the determinant” in the sense that $|\mathbf{E}_{r \times \alpha} \mathbf{A}| = |\mathbf{A}|$.

Type 2: Swapping Two Rows

For each distinct pair $r, s \in \mathbb{N}_m$,

let the $m \times m$ matrix $\mathbf{E}_{r \leftrightarrow s}$ represent the **swap** operation

that, when applied to any $m \times n$ matrix \mathbf{A} ,

results in row r of \mathbf{A} becoming row s of $\mathbf{E}_{r \leftrightarrow s} \mathbf{A}$, and vice versa.

The elements of $\mathbf{E}_{r \leftrightarrow s}$, which are those of $\mathbf{E}_{r \leftrightarrow s} \mathbf{I}_m$, are given by

$$(\mathbf{E}_{r \leftrightarrow s})_{ij} = \begin{cases} \delta_{ij} & \text{if } i \notin \{r, s\} \\ \delta_{sj} & \text{if } i = r \\ \delta_{rj} & \text{if } i = s \end{cases} \quad \text{for all } (i, j) \in \mathbb{N}_m \times \mathbb{N}_m$$

Suppose $m = n$, so the determinant $|\mathbf{A}|$ is well defined.

Then Rule 4 for determinants implies that $|\mathbf{E}_{r \leftrightarrow s} \mathbf{A}| = -|\mathbf{A}|$.

Putting $\mathbf{A} = \mathbf{I}_m$ in this equality implies that

$$|\mathbf{E}_{r \leftrightarrow s}| = |\mathbf{E}_{r \leftrightarrow s} \mathbf{I}_m| = -|\mathbf{I}_m| = -1$$

Because $|\mathbf{E}_{r \leftrightarrow s} \mathbf{A}| = |\mathbf{A}|$ only if $|\mathbf{A}| = 0$,

this matrix is not “determinant preserving”.

Type 3: Pivoting by Adding One Rescaled Row to Another

For each distinct pair $r, s \in \mathbb{N}_m$ and each scalar $\alpha \in \mathbb{R}$, let the $m \times m$ matrix $\mathbf{E}_{r+\alpha s}$ represent the **elementary row pivot operation** which, when applied to any $m \times n$ matrix \mathbf{A} , adds α times its row s to its row r , without affecting any other row.

The elements of $\mathbf{E}_{r+\alpha s}$, which are those of $\mathbf{E}_{r+\alpha s} \mathbf{I}_m$, are given for all $(i, j) \in \mathbb{N}_m \times \mathbb{N}_m$ by

$$(\mathbf{E}_{r+\alpha s})_{ij} = \left\{ \begin{array}{ll} \delta_{ij} & \text{if } i \neq r \\ \delta_{ij} + \alpha \delta_{sj} & \text{if } i = r \end{array} \right\} = \delta_{ij} + \alpha \delta_{ir} \delta_{sj}$$

Thus $\mathbf{E}_{r+\alpha s} = \mathbf{I}_m + \alpha \mathbf{1}_{rs}$ where $\mathbf{1}_{rs}$ denotes the $m \times m$ matrix whose only non-zero element is 1 in row r and column s .

In particular $\mathbf{E}_{r+\alpha s}$ is upper or lower triangular according as $r < s$ or $r > s$, or equivalently, according as row r is above or below row s .

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

Determinant Preserving Operations: Definition

Definition

For each $m, n \in \mathbb{N}$,

let $\mathcal{M}_{m \times n}$ denote the family of all $m \times n$ matrices.

The row operation $\mathcal{M}_{m \times n} \ni \mathbf{X} \mapsto \mathbf{EX} \in \mathcal{M}_{m \times n}$

that is represented by the $m \times m$ matrix \mathbf{E}

is **determinant preserving** just in case,

given any $m \times m$ matrix \mathbf{A} , one has $|\mathbf{EA}| = |\mathbf{A}|$.

Similarly, the column operation $\mathcal{M}_{m \times n} \ni \mathbf{Y} \mapsto \mathbf{YE} \in \mathcal{M}_{m \times n}$

that is represented by the $n \times n$ matrix \mathbf{E}

is **determinant preserving** just in case,

given any $n \times n$ matrix \mathbf{A} , one has $|\mathbf{AE}| = |\mathbf{A}|$. □

Properties of Determinant Preserving Operations, I

Lemma

If a square matrix \mathbf{E} represents either a row or column operation that is determinant preserving, then $|\mathbf{E}| = 1$.

Proof.

Because \mathbf{I} is a diagonal matrix, putting $\mathbf{X} = \mathbf{I}$ or $\mathbf{Y} = \mathbf{I}$ in the definition of determinant preservation gives:

1. $|\mathbf{E}| = |\mathbf{E}\mathbf{I}| = |\mathbf{I}| = 1$ in the case of a row operation;
2. $|\mathbf{E}| = |\mathbf{I}\mathbf{E}| = |\mathbf{I}| = 1$ in the case of a column operation. □

Properties of Determinant Preserving Operations, II

Proposition

Suppose that the two $k \times k$ matrices \mathbf{E}_1 and \mathbf{E}_2 both represent determinant preserving row (resp. column) operations.

Then the $k \times k$ product matrix $\mathbf{E}_1\mathbf{E}_2$ also represents a determinant preserving row (resp. column) operation.

Proof.

Given any $k \times n$ matrix \mathbf{X} , because $\mathbf{E}_2\mathbf{X}$ is a $k \times n$ matrix, determinant preservation of both \mathbf{E}_1 and \mathbf{E}_2 implies that

$$|(\mathbf{E}_1\mathbf{E}_2)\mathbf{X}| = |\mathbf{E}_1(\mathbf{E}_2\mathbf{X})| = |\mathbf{E}_2\mathbf{X}| = |\mathbf{X}|$$

Similarly, given any $m \times k$ matrix \mathbf{Y} , because $\mathbf{Y}\mathbf{E}_1$ is an $m \times k$ matrix, determinant preservation of both \mathbf{E}_1 and \mathbf{E}_2 implies that

$$|\mathbf{Y}(\mathbf{E}_1\mathbf{E}_2)| = |(\mathbf{Y}\mathbf{E}_1)\mathbf{E}_2| = |\mathbf{Y}\mathbf{E}_1| = |\mathbf{Y}|$$

□

Elementary Pivoting Is Determinant Preserving

Given any triple $(r, s, \alpha) \in \mathbb{N}_m \times \mathbb{N}_m \times \mathbb{R}$ with $r \neq s$, the $m \times m$ matrix $\mathbf{E}_{r+\alpha s}$ represents the elementary pivot row operation $\mathcal{M}_{m \times n} \ni \mathbf{X} \mapsto \mathbf{E}_{r+\alpha s} \mathbf{X} \in \mathcal{M}_{m \times n}$ of adding α times row s of the matrix \mathbf{X} to its row r .

Similarly, the $n \times n$ matrix $(\mathbf{E}_{r+\alpha s})^\top$ represents the elementary pivot column operation $\mathcal{M}_{m \times n} \ni \mathbf{Y} \mapsto \mathbf{Y}(\mathbf{E}_{r+\alpha s})^\top \in \mathcal{M}_{m \times n}$ of adding α times column s of the matrix \mathbf{Y} to its column r .

Consider any $m \times n$ matrix \mathbf{A} with $n = m$, so that \mathbf{A} has a well defined determinant $|\mathbf{A}|$.

Then Rule 6 for determinants implies that

$$|\mathbf{E}_{r+\alpha s} \mathbf{A}| = |\mathbf{A}(\mathbf{E}_{r+\alpha s})^\top| = |\mathbf{A}|$$

In this sense, both the row operation represented by $\mathbf{E}_{r+\alpha s}$ and the column operation represented by $(\mathbf{E}_{r+\alpha s})^\top$ are determinant preserving.

Determinant Preserving Row Swaps

The second elementary row operation $\mathbf{E}_{r \leftrightarrow s}$ of swapping is not determinant preserving without a key modification.

Let $\hat{\mathbf{T}}_{rs} = \mathbf{E}_{s \times (-1)} \mathbf{E}_{r \leftrightarrow s}$ denote the $m \times m$ matrix that describes the combined row operation of:

1. first interchanging rows r and s , as in $\mathbf{E}_{r \leftrightarrow s}$;
2. but then adjusting or correcting the sign of row s by multiplying it by -1 , as in $\mathbf{E}_{s \times (-1)}$.

From Rules 3 and 4 for determinants, given any $m \times m$ matrix \mathbf{X} , we have $|\mathbf{E}_{r \leftrightarrow s} \mathbf{X}| = -|\mathbf{X}|$ and then

$$|\hat{\mathbf{T}}_{rs} \mathbf{X}| = |\mathbf{E}_{s \times (-1)}(\mathbf{E}_{r \leftrightarrow s} \mathbf{X})| = (-1)|\mathbf{E}_{r \leftrightarrow s} \mathbf{X}| = |\mathbf{X}|$$

So $\mathbf{X} \mapsto \hat{\mathbf{T}}_{rs} \mathbf{X}$ is a determinant preserving row operation.

Determinant Preserving Column Swaps

Note that, if the $m \times m$ matrix \mathbf{R} represents a row operation $\mathbf{X} \mapsto \mathbf{RX}$ on $m \times n$ matrices \mathbf{X} , then its transpose \mathbf{R}^\top represents a column operation $\mathbf{Y} \mapsto \mathbf{YA}^\top$ on $n \times m$ matrices \mathbf{Y} .

In particular, because $\mathbf{X} \mapsto \hat{\mathbf{T}}_{rs}\mathbf{X}$ is a determinant preserving row operation, it follows that $\mathbf{Y} \mapsto \mathbf{Y}(\hat{\mathbf{T}}_{rs})^\top$ is a determinant preserving column operation.

Outline

Determinants: Introduction

Determinants of Orders 2 and 3

The Determinant Function

More Properties of Determinants

Eight Basic Rules for Determinants

Triangular Matrices

Pivoting

Motivating Example

Elementary Row Operations

Determinant Preserving Row Operations

Permutation and Transposition Matrices

Permutation Matrices: Definition

Definition

Given any permutation $\pi \in \Pi_n$ on $\mathbb{N}_n = \{1, 2, \dots, n\}$,
define \mathbf{P}^π as the $n \times n$ **permutation matrix**
whose elements satisfy $p_{\pi(i),j}^\pi = \delta_{i,j}$ or equivalently $p_{i,j}^\pi = \delta_{\pi^{-1}(i),j}$.

That is, the rows of the identity matrix \mathbf{I}_n are permuted
so that for each $i = 1, 2, \dots, n$, its i th row vector $(\mathbf{e}_i)^\top$,
whose j th element is δ_{ij} for each $j \in \mathbb{N}_n$,
is moved to become row $\pi(i)$ of \mathbf{P}^π ,
whose j th element is $\delta_{ij} = p_{\pi(i),j}^\pi$ for each $j \in \mathbb{N}_n$. □

Permutation Matrices: 2×2 Examples

Example

There are two 2×2 permutation matrices, which are given by:

$$\mathbf{P}^{12} = \mathbf{I}_2; \quad \mathbf{P}^{21} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Their signs, and their determinants, are respectively $+1$ and -1 .

Permutation Matrices: 3×3 Examples

Example

There are $3! = 6$ permutation matrices in 3 dimensions given by:

$$\mathbf{P}^{123} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{P}^{132} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \mathbf{P}^{213} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{P}^{231} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{P}^{312} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \mathbf{P}^{321} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Their signs equal their determinants, which satisfy

$$\begin{aligned} |\mathbf{P}^{123}| &= |\mathbf{P}^{231}| = |\mathbf{P}^{312}| = +1 \\ \text{and } |\mathbf{P}^{132}| &= |\mathbf{P}^{213}| = |\mathbf{P}^{321}| = -1 \end{aligned}$$

Multiplying a Matrix by a Permutation Matrix

Lemma

Given any $n \times n$ matrix \mathbf{A} , for each permutation $\pi \in \Pi_n$ the corresponding permutation matrix \mathbf{P}^π satisfies

$$(\mathbf{P}^\pi \mathbf{A})_{\pi(i),j} = a_{ij} = (\mathbf{A} \mathbf{P}^\pi)_{i,\pi(j)}$$

Proof.

For each pair $(i, j) \in \mathbb{N}_n^2$, one has

$$(\mathbf{P}^\pi \mathbf{A})_{\pi(i),j} = \sum_{k=1}^n p_{\pi(i),k}^\pi a_{kj} = \sum_{k=1}^n \delta_{ik} a_{kj} = a_{ij}$$

and also

$$(\mathbf{A} \mathbf{P}^\pi)_{i,\pi(j)} = \sum_{k=1}^n a_{ik} p_{k,\pi(j)}^\pi = \sum_{k=1}^n a_{ik} \delta_{kj} = a_{ij} \quad \square$$

So $\left\{ \begin{array}{l} \text{premultiplying} \\ \text{postmultiplying} \end{array} \right\} \mathbf{A}$ by \mathbf{P}^π applies π to \mathbf{A} 's $\left\{ \begin{array}{l} \text{rows} \\ \text{columns} \end{array} \right\}$.

Multiplying Permutation Matrices

Theorem

Given the composition $\pi \circ \rho$ of two permutations $\pi, \rho \in \Pi_n$, the associated permutation matrices satisfy $\mathbf{P}^\pi \mathbf{P}^\rho = \mathbf{P}^{\pi \circ \rho}$.

Proof.

For each pair $(i, j) \in \mathbb{N}_n^2$, one has

$$\begin{aligned}(\mathbf{P}^\pi \mathbf{P}^\rho)_{ij} &= \sum_{k=1}^n p_{ik}^\pi p_{kj}^\rho = \sum_{k=1}^n \delta_{\pi^{-1}(i), k} \delta_{\rho^{-1}(k), j} \\ &= \sum_{k=1}^n \delta_{(\rho^{-1} \circ \pi^{-1})(i), \rho^{-1}(k)} \delta_{\rho^{-1}(k), j} \\ &= \sum_{\ell=1}^n \delta_{(\pi \circ \rho)^{-1}(i), \ell} \delta_{\ell, j} = \delta_{(\pi \circ \rho)^{-1}(i), j} \\ &= p_{ij}^{\pi \circ \rho} = (\mathbf{P}^{\pi \circ \rho})_{ij} \quad \square\end{aligned}$$

Corollary

If $\pi = \pi^1 \circ \pi^2 \circ \dots \circ \pi^q$, then $\mathbf{P}^\pi = \mathbf{P}^{\pi^1} \mathbf{P}^{\pi^2} \dots \mathbf{P}^{\pi^q}$.

Proof.

By induction on q , using the result of the Theorem. □

Any Permutation Matrix Is Orthogonal

Proposition

Any permutation matrix \mathbf{P}^π satisfies $\mathbf{P}^\pi (\mathbf{P}^\pi)^\top = (\mathbf{P}^\pi)^\top \mathbf{P}^\pi = \mathbf{I}_n$, so is orthogonal.

Proof.

Because π is a permutation on \mathbb{N}_n , for each pair $(i, j) \in \mathbb{N}_n^2$, one has

$$\begin{aligned} [\mathbf{P}^\pi (\mathbf{P}^\pi)^\top]_{ij} &= \sum_{k=1}^n p_{ik}^\pi p_{jk}^\pi = \sum_{k=1}^n \delta_{\pi^{-1}(i), k} \delta_{\pi^{-1}(j), k} \\ &= \delta_{\pi^{-1}(i), \pi^{-1}(j)} = \delta_{ij} \end{aligned}$$

and also

$$\begin{aligned} [(\mathbf{P}^\pi)^\top \mathbf{P}^\pi]_{ij} &= \sum_{k=1}^n p_{ki}^\pi p_{kj}^\pi = \sum_{k=1}^n \delta_{\pi^{-1}(k), i} \delta_{\pi^{-1}(k), j} \\ &= \sum_{\ell=1}^n \delta_{\ell, i} \delta_{\ell, j} = \delta_{ij} \quad \square \end{aligned}$$

Transposition Matrices

A special case of a permutation matrix is a **transposition** or **swap** \mathbf{T}_{rs} of rows r and s .

As the matrix \mathbf{I} with rows r and s transposed, it satisfies

$$(\mathbf{T}_{rs})_{ij} = \begin{cases} \delta_{ij} & \text{if } i \notin \{r, s\} \\ \delta_{sj} & \text{if } i = r \\ \delta_{rj} & \text{if } i = s \end{cases}$$

Remark

Distinguish carefully between the two operations of:

1. **swapping** the two particular rows or columns r and s of a matrix \mathbf{A} , which results from applying \mathbf{T}_{rs} or \mathbf{T}_{rs}^T to \mathbf{A} ;
2. **transposing** an entire matrix from \mathbf{A} to \mathbf{A}^T , which results from converting each row vector of \mathbf{A} into a column vector of \mathbf{A}^T , and vice versa.

Transposition Matrices: Exercise

Exercise

1. Prove that: (i) \mathbf{T}_{rs} is symmetric and orthogonal;
(ii) $\mathbf{T}_{rs} = \mathbf{T}_{sr}$; (iii) $\mathbf{T}_{rs}\mathbf{T}_{sr} = \mathbf{T}_{sr}\mathbf{T}_{rs} = \mathbf{I}$.
2. Prove that, if \mathbf{A} is any $m \times n$ matrix, then:
 - (i) if \mathbf{T}_{rs} is $m \times m$,
then $\mathbf{T}_{rs}\mathbf{A}$ is \mathbf{A} with rows r and s interchanged;
 - (ii) if \mathbf{T}_{rs} is $n \times n$,
then $\mathbf{A}\mathbf{T}_{rs}$ is \mathbf{A} with columns r and s interchanged.

Determinants with Permuted Rows: Theorem

Theorem

Given any $n \times n$ matrix \mathbf{A} and any permutation $\pi \in \Pi_n$, one has $|\mathbf{P}^\pi \mathbf{A}| = |\mathbf{A} \mathbf{P}^\pi| = \text{sgn}(\pi) |\mathbf{A}|$.

The proof appears on the next slide.

Meanwhile, putting $\mathbf{A} = \mathbf{I}$ in the theorem gives immediately:

Corollary

Given any permutation $\pi \in \Pi_n$, the associated permutation matrix \mathbf{P}^π satisfies $|\mathbf{P}^\pi| = \text{sgn}(\pi)$.

Determinants with Permuted Rows: Proof

Proof.

The expansion formula for determinants gives

$$|\mathbf{P}^\pi \mathbf{A}| = \sum_{\rho \in \Pi_n} \operatorname{sgn}(\rho) \prod_{i=1}^n (\mathbf{P}^\pi \mathbf{A})_{i,\rho(i)}$$

But for each $i \in \mathbb{N}_n$, $\rho \in \Pi_n$, one has $(\mathbf{P}^\pi \mathbf{A})_{i,\rho(i)} = a_{\pi^{-1}(i),\rho(i)}$, so

$$\begin{aligned} |\mathbf{P}^\pi \mathbf{A}| &= \sum_{\rho \in \Pi_n} \operatorname{sgn}(\rho) \prod_{i=1}^n a_{\pi^{-1}(i),\rho(i)} \\ &= [1/\operatorname{sgn}(\pi)] \sum_{\pi \circ \rho \in \Pi_n} \operatorname{sgn}(\pi \circ \rho) \prod_{i=1}^n a_{i,(\pi \circ \rho)(i)} \\ &= \operatorname{sgn}(\pi) \sum_{\sigma \in \Pi_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} = \operatorname{sgn}(\pi) |\mathbf{A}| \end{aligned}$$

because $\operatorname{sgn}(\pi \circ \rho) = \operatorname{sgn}(\pi) \operatorname{sgn}(\rho)$ and $1/\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi)$, whereas there is an obvious bijection $\Pi_n \ni \rho \leftrightarrow \pi \circ \rho = \sigma \in \Pi_n$ on the set of permutations Π_n .

The proof that $|\mathbf{A}\mathbf{P}^\pi| = \operatorname{sgn}(\pi) |\mathbf{A}|$ is sufficiently similar to be left as an exercise. □

The Alternation Rule for Determinants

Corollary

Given any $n \times n$ matrix \mathbf{A}

and any transposition τ_{rs} with associated transposition matrix \mathbf{T}_{rs} ,
one has $|\mathbf{T}_{rs}\mathbf{A}| = |\mathbf{A}\mathbf{T}_{rs}| = -|\mathbf{A}|$.

Proof.

Apply the previous theorem in the special case
when $\pi = \tau_{rs}$ and so $\mathbf{P}^\pi = \mathbf{T}_{rs}$.

Then, because $\text{sgn}(\pi) = \text{sgn}(\tau_{rs}) = -1$,

the equality $|\mathbf{P}^\pi\mathbf{A}| = \text{sgn}(\pi)|\mathbf{A}|$ implies that $|\mathbf{T}_{rs}\mathbf{A}| = -|\mathbf{A}|$. □

We have shown that, for any $n \times n$ matrix \mathbf{A} , given any:

1. permutation $\pi \in \mathbb{N}_n$, one has $|\mathbf{P}^\pi\mathbf{A}| = |\mathbf{A}\mathbf{P}^\pi| = \text{sgn}(\pi)|\mathbf{A}|$;
2. transposition τ_{rs} , one has $|\mathbf{T}_{rs}\mathbf{A}| = |\mathbf{A}\mathbf{T}_{rs}| = -|\mathbf{A}|$.

Sign Adjusted Transposition Matrices

We define the **sign adjusted** $m \times m$ transposition matrix $\hat{\mathbf{T}}_{rs}$ so that, given any $m \times n$ matrix \mathbf{A} , the matrix $\hat{\mathbf{T}}_{rs}\mathbf{A}$ is the result of:

- first swapping rows r and s of the matrix \mathbf{A} ;
- then multiplying row s in the result by -1 .

Because it is the matrix \mathbf{I} with rows r and s transposed, and then row s multiplied by -1 , the matrix $\hat{\mathbf{T}}_{rs}$ has elements that satisfy

$$(\hat{\mathbf{T}}_{rs})_{ij} = \begin{cases} \delta_{ij} & \text{if } i \notin \{r, s\} \\ \delta_{sj} & \text{if } i = r \\ -\delta_{rj} & \text{if } i = s \end{cases}$$

Rules 3 and 4 together imply that $|\hat{\mathbf{T}}_{rs}| = |(-1)\mathbf{T}_{rs}| = 1$.

In the special case of any $m \times m$ matrix \mathbf{A} , this implies that the determinants satisfy $|\hat{\mathbf{T}}_{rs}\mathbf{A}| = |(-1)\mathbf{T}_{rs}\mathbf{A}| = |\mathbf{A}|$.

2×2 and 3×3 Sign Adjusted Transposition Matrices

Example

1. The two different 2×2 sign adjusted transposition matrices are $\hat{\mathbf{T}}_{12} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\hat{\mathbf{T}}_{21} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = (\hat{\mathbf{T}}_{12})^\top = -\hat{\mathbf{T}}_{12}$.
2. There are six 3×3 sign adjusted transposition matrices.

The first two satisfy $\hat{\mathbf{T}}_{12} = (\hat{\mathbf{T}}_{21})^\top = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Two others satisfy $\hat{\mathbf{T}}_{13} = (\hat{\mathbf{T}}_{31})^\top = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$.

The last two satisfy $\hat{\mathbf{T}}_{23} = (\hat{\mathbf{T}}_{32})^\top = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

Inverses of Sign Adjusted Transposition Matrices

Exercise

1. Verify that, because $\hat{\mathbf{T}}_{12}\hat{\mathbf{T}}_{21} = \hat{\mathbf{T}}_{21}\hat{\mathbf{T}}_{12} = \mathbf{I}_2$, the two 2×2 matrices $\hat{\mathbf{T}}_{12}$ and $\hat{\mathbf{T}}_{21}$ are inverses.
2. Verify that whenever $r, s \in \mathbb{N}_3$ with $r \neq s$, the two 3×3 matrices $\hat{\mathbf{T}}_{rs}$ and $\hat{\mathbf{T}}_{sr}$ are inverses.

Harder: Verify directly that whenever $r, s \in \mathbb{N}_m$ with $r \neq s$, the two $m \times m$ matrices $\hat{\mathbf{T}}_{rs}$ and $\hat{\mathbf{T}}_{sr}$ satisfy $\hat{\mathbf{T}}_{rs} = (\hat{\mathbf{T}}_{sr})^\top$ and are inverses.

Sign Adjusted Permutation Matrices

Given any permutation matrix \mathbf{P} ,
there is a unique permutation π such that $\mathbf{P} = \mathbf{P}^\pi$.

Suppose that $\pi = \tau_{r_1 s_1} \circ \cdots \circ \tau_{r_\ell s_\ell}$ is any one of the several ways in which the permutation π can be decomposed into a composition of transpositions.

Then $\mathbf{P} = \prod_{k=1}^{\ell} \mathbf{T}_{r_k s_k}$ and $|\mathbf{PA}| = (-1)^\ell |\mathbf{A}|$ for any \mathbf{A} .

Definition

Say that $\hat{\mathbf{P}}$ is a **sign adjusted** version of $\mathbf{P} = \mathbf{P}^\pi$
just in case it can be expressed as the product $\hat{\mathbf{P}} = \prod_{k=1}^{\ell} \hat{\mathbf{T}}_{r_k s_k}$
of sign adjusted transpositions satisfying $\mathbf{P} = \prod_{k=1}^{\ell} \mathbf{T}_{r_k s_k}$.

Then it is easy to prove by induction on ℓ
that for every $n \times n$ matrix \mathbf{A} one has $|\hat{\mathbf{P}}\mathbf{A}| = |\mathbf{A}\hat{\mathbf{P}}| = |\mathbf{A}|$.

Recall that all the elements of a permutation matrix \mathbf{P} are 0 or 1.

A sign adjustment of \mathbf{P} involves changing some of the 1 elements into -1 elements, while leaving all the 0 elements unchanged.