Lecture Notes 1: Matrix Algebra
Part B: Determinants and Inverses

Peter J. Hammond

revised 2020 September 14th

Outline

Square Matrices

A square matrix has an equal number of rows and columns,
this number being called its dimension.

The (principal, or main) diagonal
of a square matrix $\mathbf{A} = (a_{ij})_{n \times n}$ of dimension n
is the list $(a_{ii})_{i=1}^n = (a_{11}, a_{22}, \ldots, a_{nn})$ of its n diagonal elements.

The other elements a_{ij} with $i \neq j$ are the off-diagonal elements.

A square matrix is often expressed in the form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

with some extra dots along the diagonal.

Symmetric Matrices

Definition

A square matrix $\mathbf{A}$ is symmetric just in case
it is equal to its transpose — i.e., if $\mathbf{A}^\top = \mathbf{A}$.

Example

The product of two symmetric matrices need not be symmetric.

Using again our example of non-commuting 2×2 matrices,
here are two examples
where the product of two symmetric matrices is asymmetric:

- $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$;
- $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

Two Exercises with Symmetric Matrices

Exercise

Let $\mathbf{x}$ be a column n-vector.

1. *Find the dimensions of $\mathbf{x}^\top \mathbf{x}$ and of $\mathbf{x}\mathbf{x}^\top$.*
2. *Show that one is a non-negative number which is positive unless $\mathbf{x} = \mathbf{0}$, and that the other is an $n \times n$ symmetric matrix.*

Exercise

Let $\mathbf{A}$ be an $m \times n$-matrix.

1. *Find the dimensions of $\mathbf{A}^\top \mathbf{A}$ and of $\mathbf{A}\mathbf{A}^\top$.*
2. *Show that both $\mathbf{A}^\top \mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ are symmetric matrices.*
3. *Show that $m = n$ is a necessary condition for $\mathbf{A}^\top \mathbf{A} = \mathbf{A}\mathbf{A}^\top$.*
4. *Show that $m = n$ with $\mathbf{A}$ symmetric is a sufficient condition for $\mathbf{A}^\top \mathbf{A} = \mathbf{A}\mathbf{A}^\top$.*

Diagonal Matrices

A square matrix $\mathbf{A} = (a_{ij})^{n \times n}$ is diagonal
just in case all of its off diagonal elements are 0
— i.e., $i \neq j \Longrightarrow a_{ij} = 0$.

A diagonal matrix of dimension n can be written in the form

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{pmatrix} = \mathbf{diag}(d_1, d_2, d_3, \dots, d_n) = \mathbf{diag}\,\mathbf{d}$$

where the n-vector $\mathbf{d} = (d_1, d_2, d_3, \dots, d_n) = (d_i)_{i=1}^n$
consists of the diagonal elements of $\mathbf{D}$.

Note that $\mathbf{diag}\,\mathbf{d} = (d_{ij})_{n \times n}$ where each $d_{ij} = \delta_{ij} d_{ii} = \delta_{ij} d_{jj}$.

Obviously, any diagonal matrix is symmetric.

Multiplying by Diagonal Matrices

Example

Let $\mathbf{D}$ be a diagonal matrix of dimension n.

Suppose that $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ and $n \times m$ matrices, respectively.

Then $\mathbf{E} := \mathbf{AD}$ and $\mathbf{F} := \mathbf{DB}$ are well defined matrices of dimensions $m \times n$ and $n \times m$, respectively.

By the law of matrix multiplication, their elements are

$$e_{ij} = \sum\nolimits_{k=1}^{n} a_{ik} \delta_{kj} d_{jj} = a_{ij} d_{jj} \text{ and } f_{ij} = \sum\nolimits_{k=1}^{n} \delta_{ik} d_{ii} b_{kj} = d_{ii} b_{ij}$$

Thus, post-multiplying $\mathbf{A}$ by $\mathbf{D}$ is the column operation of simultaneously multiplying every column $\mathbf{a}_j$ of $\mathbf{A}$ by its matching diagonal element d_{jj}.

Similarly, pre-multiplying $\mathbf{B}$ by $\mathbf{D}$ is the row operation of simultaneously multiplying every row $\mathbf{b}_i^\top$ of $\mathbf{B}$ by its matching diagonal element d_{ii}.

Two Exercises with Diagonal Matrices

Exercise

Let $\mathbf{D}$ *be a diagonal matrix of dimension* n.
Give conditions that are both necessary and sufficient for each of the following:

1. $\mathbf{AD} = \mathbf{A}$ *for every* $m \times n$ *matrix* $\mathbf{A}$;
2. $\mathbf{DB} = \mathbf{B}$ *for every* $n \times m$ *matrix* $\mathbf{B}$.

Exercise

Let $\mathbf{D}$ *be a diagonal matrix of dimension* n,
and $\mathbf{C}$ *any* $n \times n$ *matrix.*

An earlier example shows that one can have $\mathbf{CD} \neq \mathbf{DC}$ *even if* $n = 2$.

1. *Show that* $\mathbf{C}$ *being diagonal is a sufficient condition for* $\mathbf{CD} = \mathbf{DC}$.
2. *Is this condition necessary?*

Outline

The Identity Matrix

The identity matrix of dimension n is the diagonal matrix

$$\mathbf{I}_n = \mathbf{diag}(1, 1, \ldots, 1)$$

whose n diagonal elements are all equal to 1.

Equivalently, it is the $n \times n$-matrix $\mathbf{A} = (a_{ij})^{n \times n}$
whose elements are all given by $a_{ij} = \delta_{ij}$
for the Kronecker delta function $(i, j) \mapsto \delta_{ij}$
defined on $\{1, 2, \ldots, n\}^2$.

Exercise

Given any $m \times n$ matrix $\mathbf{A}$, verify that $\mathbf{I}_m\mathbf{A} = \mathbf{A}\mathbf{I}_n = \mathbf{A}$.

Uniqueness of the Identity Matrix

Exercise

Suppose that the two $n \times n$ matrices $\mathbf{X}$ and $\mathbf{Y}$ respectively satisfy:

1. $\mathbf{AX} = \mathbf{A}$ *for every $m \times n$ matrix $\mathbf{A}$;*
2. $\mathbf{YB} = \mathbf{B}$ *for every $n \times m$ matrix $\mathbf{B}$.*

Prove that $\mathbf{X} = \mathbf{Y} = \mathbf{I}_n$.

(Hint: Consider each of the mn different cases where $\mathbf{A}$ (resp. $\mathbf{B}$) has exactly one non-zero element that is equal to 1.)

The results of the last two exercises together serve to prove:

Theorem

The identity matrix $\mathbf{I}_n$ is the unique $n \times n$-matrix such that:

- $\mathbf{I}_n\mathbf{B} = \mathbf{B}$ *for each $n \times m$ matrix $\mathbf{B}$;*
- $\mathbf{AI}_n = \mathbf{A}$ *for each $m \times n$ matrix $\mathbf{A}$.*

How the Identity Matrix Earns its Name

Remark

The identity matrix $\mathbf{I}_n$ *earns its name because it represents a multiplicative identity on the "algebra" of all* $n \times n$ *matrices.*

That is, $\mathbf{I}_n$ *is the unique* $n \times n$*-matrix with the property that* $\mathbf{I}_n\mathbf{A} = \mathbf{A}\mathbf{I}_n = \mathbf{A}$ *for every* $n \times n$*-matrix* $\mathbf{A}$.

Typical notation suppresses the subscript n in $\mathbf{I}_n$ that indicates the dimension of the identity matrix.

Outline

Left and Right Inverse Matrices

Definition

Let $\mathbf{A}$ denote any $n \times n$ matrix.

1. The $n \times n$ matrix $\mathbf{X}$ is a left inverse of $\mathbf{A}$ just in case $\mathbf{XA} = \mathbf{I}_n$.
2. The $n \times n$ matrix $\mathbf{Y}$ is a right inverse of $\mathbf{A}$ just in case $\mathbf{AY} = \mathbf{I}_n$.
3. The $n \times n$ matrix $\mathbf{Z}$ is an inverse of $\mathbf{A}$ just in case it is both a left and a right inverse — i.e., $\mathbf{ZA} = \mathbf{AZ} = \mathbf{I}_n$.

The Unique Inverse Matrix

Theorem

Suppose that the $n \times n$ matrix $\mathbf{A}$ has both a left and a right inverse.

Then both left and right inverses are unique,
and both are equal to a unique inverse matrix denoted by $\mathbf{A}^{-1}$.

Proof.

If $\mathbf{XA} = \mathbf{AY} = \mathbf{I}$, then $\mathbf{XAY} = \mathbf{XI} = \mathbf{X}$ and $\mathbf{XAY} = \mathbf{IY} = \mathbf{Y}$,
implying that $\mathbf{X} = \mathbf{XAY} = \mathbf{Y}$.

Now, if $\tilde{\mathbf{X}}$ is any alternative left inverse,
then $\tilde{\mathbf{X}}\mathbf{A} = \mathbf{I}$ and so $\tilde{\mathbf{X}} = \tilde{\mathbf{X}}\mathbf{AY} = \mathbf{Y} = \mathbf{X}$.

Similarly, if $\tilde{\mathbf{Y}}$ is any alternative right inverse,
then $\mathbf{A}\tilde{\mathbf{Y}} = \mathbf{I}$ and so $\tilde{\mathbf{Y}} = \mathbf{XA}\tilde{\mathbf{Y}} = \mathbf{X} = \mathbf{Y}$.

It follows that $\tilde{\mathbf{X}} = \mathbf{X} = \mathbf{Y} = \tilde{\mathbf{Y}}$, so we can define $\mathbf{A}^{-1}$
as the unique common value of all these four matrices. □

Big question: when does the inverse exist?
Answer: if and only if the determinant is non-zero.

Rule for Inverting Products

Theorem

Suppose that $\mathbf{A}$ *and* $\mathbf{B}$ *are two invertible* $n \times n$ *matrices.*

Then the inverse of the matrix product $\mathbf{AB}$ *exists, and is the reverse product* $\mathbf{B}^{-1}\mathbf{A}^{-1}$ *of the inverses.*

Proof.

Using the associative law for matrix multiplication repeatedly gives:

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}(\mathbf{I})\mathbf{B} = \mathbf{B}^{-1}(\mathbf{IB}) = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

and

$$(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{A}(\mathbf{I})\mathbf{A}^{-1} = (\mathbf{AI})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}.$$

These equations confirm that $\mathbf{X} := \mathbf{B}^{-1}\mathbf{A}^{-1}$ is the unique matrix satisfying the double equality $(\mathbf{AB})\mathbf{X} = \mathbf{X}(\mathbf{AB}) = \mathbf{I}$. □

Rule for Inverting Chain Products and Transposes

Exercise

Prove that, if $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *are three invertible* $n \times n$ *matrices, then* $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

Then use mathematical induction
to extend the rule for inverting any product $\mathbf{BC}$
in order to find the inverse of the product $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k$
of any finite chain of invertible $n \times n$ *matrices.*

Theorem

Suppose that $\mathbf{A}$ *is an invertible* $n \times n$ *matrix.*

Then the inverse $(\mathbf{A}^\top)^{-1}$ *of its transpose*
is $(\mathbf{A}^{-1})^\top$, *the transpose of its inverse.*

Proof.

By the rule for transposing products, one has

$$\mathbf{A}^\top(\mathbf{A}^{-1})^\top = (\mathbf{A}^{-1}\mathbf{A})^\top = \mathbf{I}^\top = \mathbf{I} \qquad \square$$

Orthogonal and Orthonormal Sets of Vectors

Definition

A set of k vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^n$ is said to be:

- pairwise orthogonal just in case $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ whenever $j \neq i$;
- orthonormal just in case, in addition, each $\|\mathbf{x}_i\| = 1$ — i.e., all k elements of the set are vectors of unit length.

The set of k vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^n$ is orthonormal just in case $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$ for all pairs $i, j \in \{1, 2, \ldots, k\}$.

Orthogonal Matrices

Definition

Any $n \times n$ matrix is orthogonal
just in case its n columns form an orthonormal set.

Theorem

Given any $n \times n$ matrix $\mathbf{P}$, the following are equivalent:

1. $\mathbf{P}$ *is orthogonal;*
2. $\mathbf{P}\mathbf{P}^\top = \mathbf{P}^\top\mathbf{P} = \mathbf{I}$;
3. $\mathbf{P}^{-1} = \mathbf{P}^\top$;
4. $\mathbf{P}^\top$ *is orthogonal.*

The proof follows from the definitions, and is left as an exercise.

Outline

Partitioned Matrices: Definition

A partitioned matrix is a rectangular array of different matrices.

Example

Consider the $(m + \ell) \times (n + k)$ matrix

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{m \times n} & \mathbf{B}_{m \times k} \\ \mathbf{C}_{\ell \times n} & \mathbf{D}_{\ell \times k} \end{pmatrix}$$

where, as indicated, the four submatrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$
are of dimension $m \times n$, $m \times k$, $\ell \times n$ and $\ell \times k$ respectively.

Note: Here matrix $\mathbf{D}$ may not be diagonal, or even square.

For any scalar $\alpha \in \mathbb{R}$,
the scalar multiple of a partitioned matrix is

$$\alpha \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{A} & \alpha\mathbf{B} \\ \alpha\mathbf{C} & \alpha\mathbf{D} \end{pmatrix}$$

Partitioned Matrices: Addition

Suppose the two partitioned matrices

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}$$

have the property that the following four pairs
of corresponding matrices have equal dimensions:
(i) $\mathbf{A}$ and $\mathbf{E}$; (ii) $\mathbf{B}$ and $\mathbf{F}$; (iii) $\mathbf{C}$ and $\mathbf{G}$; (iv) $\mathbf{D}$ and $\mathbf{H}$.

Then the sum of the two matrices is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} + \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A} + \mathbf{E} & \mathbf{B} + \mathbf{F} \\ \mathbf{C} + \mathbf{G} & \mathbf{D} + \mathbf{H} \end{pmatrix}$$

Partitioned Matrices: Multiplication

Suppose that the two partitioned matrices

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}$$

along with all the relevant pairs of their sub-matrices,
are compatible for multiplication.

Then their product is defined as

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{pmatrix}$$

This extends the usual multiplication rule for matrices:
multiply the rows of sub-matrices in the first partitioned matrix
by the columns of sub-matrices in the second partitioned matrix.

Transposes and Some Special Matrices

The rule for transposing a partitioned matrix is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^\top = \begin{pmatrix} \mathbf{A}^\top & \mathbf{C}^\top \\ \mathbf{B}^\top & \mathbf{D}^\top \end{pmatrix}$$

So the original matrix is symmetric
iff $\mathbf{A} = \mathbf{A}^\top$, $\mathbf{D} = \mathbf{D}^\top$, and $\mathbf{B} = \mathbf{C}^\top \iff \mathbf{C} = \mathbf{B}^\top$.

It is diagonal iff $\mathbf{A}, \mathbf{D}$ are both diagonal,
while also $\mathbf{B} = \mathbf{0}$ and $\mathbf{C} = \mathbf{0}$.

The identity matrix is diagonal with $\mathbf{A} = \mathbf{I}$, $\mathbf{D} = \mathbf{I}$,
possibly identity matrices of different dimensions.

Partitioned Matrices: Inverses, I

For an $(m+n) \times (m+n)$ partitioned matrix to have an inverse, the equation

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{I}_n \end{pmatrix}$$

should have a solution for the matrices $\mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}$, given $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$.

Assuming that the $m \times m$ matrix $\mathbf{A}$ has an inverse, we can:

1. construct new first m equations
   by premultiplying the old ones by $\mathbf{A}^{-1}$;
2. construct new second n equations by:
   - premultiplying the new first m equations
     by the $n \times m$ matrix $\mathbf{C}$;
   - then subtracting this product from the old second n equations.

The result is

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n \times m} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0}_{m \times n} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_n \end{pmatrix}$$

Partitioned Matrices: Inverses, II

For the next step,
assume the $n \times n$ matrix $\mathbf{X} := \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$
also has an inverse $\mathbf{X}^{-1} = (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}$.

Given $\begin{pmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n\times m} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0}_{m\times n} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_n \end{pmatrix}$,
we first premultiply the last n equations by $\mathbf{X}^{-1}$ to get

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0}_{m\times n} \\ -\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X}^{-1} \end{pmatrix}$$

Next, we subtract $\mathbf{A}^{-1}\mathbf{B}$ times the last n equations
from the first m equations to obtain

$$\begin{aligned} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} &= \begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X}^{-1} \\ -\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X}^{-1} \end{pmatrix} \end{aligned}$$

Final Exercises

Exercise

1. *Assume that* $\mathbf{A}^{-1}$ *and* $\mathbf{X}^{-1} = (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}$ *exist.*

   Given $\mathbf{Z} := \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X}^{-1} \\ -\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X}^{-1} \end{pmatrix}$,
   use direct multiplication twice in order to verify that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \mathbf{Z} = \mathbf{Z} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{I}_n \end{pmatrix}$$

2. *Let* $\mathbf{A}$ *be any invertible* $m \times m$ *matrix.*

   Show that the bordered $(m+1) \times (m+1)$ *matrix* $\begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d \end{pmatrix}$
   is invertible provided that $d \neq \mathbf{c}^\top \mathbf{A}^{-1} \mathbf{b}$,
   and find its inverse in this case.

Partitioned Matrices: Extension

Exercise

Suppose that the two partitioned matrices

$$\mathbf{A} = (\mathbf{A}_{ij})^{k \times \ell} \quad \text{and} \quad \mathbf{B} = (\mathbf{B}_{ij})^{k \times \ell}$$

are both $k \times \ell$ arrays of respective $m_i \times n_j$ matrices $\mathbf{A}_{ij}, \mathbf{B}_{ij}$, for $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, \ell$.

1. *Under what conditions can the product* $\mathbf{AB}$ *be defined as a $k \times \ell$ array of matrices?*
2. *Under what conditions can the product* $\mathbf{BA}$ *be defined as a $k \times \ell$ array of matrices?*
3. *When either* $\mathbf{AB}$ *or* $\mathbf{BA}$ *can be so defined, give a formula for its product, using summation notation.*
4. *Express* $\mathbf{A}^\top$ *as a partitioned matrix.*
5. *Under what conditions is the matrix* $\mathbf{A}$ *symmetric?*

Outline

Permutations

Definition

Given $\mathbb{N}_n = \{1, \ldots, n\}$ for any $n \in \mathbb{N}$ with $n \geq 2$,
a permutation of $\mathbb{N}_n$ is a bijective mapping $\mathbb{N}_n \ni i \mapsto \pi(i) \in \mathbb{N}_n$.

That is, the mapping $\mathbb{N}_n \ni i \mapsto \pi(i) \in \mathbb{N}_n$ is both:

1. a surjection, or mapping of $\mathbb{N}_n$ onto $\mathbb{N}_n$,
   in the sense that the range set satisfies
   $\pi(\mathbb{N}_n) := \{j \in \mathbb{N}_n \mid \exists i \in \mathbb{N}_n : j = \pi(i)\} = \mathbb{N}_n$;
2. an injection, or a one to one mapping,
   in the sense that $\pi(i) = \pi(j) \implies i = j$ or,
   equivalently, $i \neq j \implies \pi(i) \neq \pi(j)$.

Exercise

Prove that the mapping $\mathbb{N}_n \ni i \mapsto f(i) \in \mathbb{N}_n$ is a bijection, and so a permutation, if and only if its range set $f(\mathbb{N}_n) := \{j \in \mathbb{N}_n \mid \exists i \in \mathbb{N}_n : j = f(i)\}$ has cardinality $\#f(\mathbb{N}_n) = \#\mathbb{N}_n = n$.

Products of Permutations

Definition

The product $\pi \circ \rho$ of two permutations $\pi, \rho \in \Pi_n$
is the composition mapping $\mathbb{N}_n \ni i \mapsto (\pi \circ \rho)(i) := \pi[\rho(i)] \in \mathbb{N}_n$.

Exercise

Prove that the product $\pi \circ \rho$ of any two permutations $\pi, \rho \in \Pi_n$ is a permutation.

Hint: Show that $\#(\pi \circ \rho)(\mathbb{N}_n) = \#\rho(\mathbb{N}_n) = \#\mathbb{N}_n = n$.

Example

1. If you shuffle a pack of 52 playing cards once,
   without dropping any on the floor,
   the result will be a permutation π of the cards.
2. If you shuffle the same pack a second time,
   the result will be a new permutation ρ of the shuffled cards.
3. Overall, the result of shuffling the cards twice
   will be the single permutation $\rho \circ \pi$.

Finite Permutation Groups

Definition

Given any $n \in \mathbb{N}$, the family Π_n of all permutations of $\mathbb{N}_n$ includes:

- the identity permutation ι defined by $\iota(h) = h$ for all $h \in \mathbb{N}_n$;
- because the mapping $\mathbb{N}_n \ni i \mapsto f(i) \in \mathbb{N}_n$ is bijective, for each $\pi \in \Pi_n$, a unique inverse permutation $\pi^{-1} \in \Pi_n$ satisfying $\pi^{-1} \circ \pi = \pi \circ \pi^{-1} = \iota$.

Definition

The associative law for functions says that, given any three functions $h : X \to Y$, $g : Y \to Z$ and $f : Z \to W$, the composite function $f \circ g \circ h : X \to W$ satisfies

$$(f \circ g \circ h)(x) \equiv f(g(h(x))) \equiv [(f \circ g) \circ h](x) \equiv [f \circ (g \circ h)](x)$$

Exercise

Given any $n \in \mathbb{N}$, show that (Π_n, π, ι) is an algebraic group — i.e., the group operation $(\pi, \rho) \mapsto \pi \circ \rho$ is well-defined, associative, with ι as the unit, and an inverse π^{-1} for every $\pi \in \Pi_n$.

Outline

Transpositions

Definition

For each disjoint pair $k, \ell \in \{1, 2, \ldots, n\}$,
the transposition mapping $i \mapsto \tau_{k\ell}(i)$ on $\{1, 2, \ldots, n\}$
is the permutation defined by

$$\tau_{k\ell}(i) := \begin{cases} \ell & \text{if } i = k; \\ k & \text{if } i = \ell; \\ i & \text{otherwise;} \end{cases}$$

That is, $\tau_{k\ell}$ transposes the order of k and ℓ,
leaving all $i \notin \{k, \ell\}$ unchanged. □

Evidently $\tau_{k\ell} = \tau_{\ell k}$ and $\tau_{k\ell} \circ \tau_{\ell k} = \iota$, the identity permutation,
and so $\tau \circ \tau = \iota$ for every transposition τ.

Transposition is Not Commutative

Any $(j_1, j_2, \ldots, j_n) \in \mathbb{N}_n^n$ whose components are all different corresponds to a unique permutation, denoted by $\pi^{j_1 j_2 \cdots j_n} \in \Pi_n$, that satisfies $\pi(i) = j_i$ for all $i \in \mathbb{N}_n^n$.

Example

Two transpositions defined
on a set containing more than two elements
may not commute because, for example,

$$\tau_{12} \circ \tau_{23} = \pi^{231} \neq \tau_{23} \circ \tau_{12} = \pi^{312}$$

Permutations are Products of Transpositions

Theorem

Any permutation $\pi \in \Pi_n$ *on* $\mathbb{N}_n := \{1, 2, \ldots, n\}$
is the product of at most $n - 1$ *transpositions.*

We will prove the result by induction on n.

As the induction hypothesis,
suppose the result holds for permutations on $\mathbb{N}_{n-1}$.

Any permutation π on $\mathbb{N}_2 := \{1, 2\}$ is either the identity,
or the transposition τ_{12}, so the result holds for $n = 2$.

Proof of Induction Step

For general n, let $j := \pi^{-1}(n)$ denote the element that π moves to the end.

By construction, the permutation $\pi \circ \tau_{jn}$ must satisfy $\pi \circ \tau_{jn}(n) = \pi(\tau_{jn}(n)) = \pi(j) = n$.

So the restriction $\tilde{\pi}$ of $\pi \circ \tau_{jn}$ to $\mathbb{N}_{n-1}$ is a permutation on $\mathbb{N}_{n-1}$.

By the induction hypothesis, for all $k \in \mathbb{N}_{n-1}$, there exist transpositions $\tau^1, \tau^2, \ldots, \tau^q$ such that $\tilde{\pi}(k) = (\pi \circ \tau_{jn})(k) = (\tau^1 \circ \tau^2 \circ \ldots \circ \tau^q)(k)$ where $q \leq n-2$ is the number of transpositions in the product.

For $p = 1, \ldots, q$, because τ^p interchanges only elements of $\mathbb{N}_{n-1}$, one can extend its domain to include n by letting $\tau^p(n) = n$.

Then $(\pi \circ \tau_{jn})(k) = (\tau^1 \circ \tau^2 \circ \ldots \circ \tau^q)(k)$ for $k = n$ as well, so $\pi = (\pi \circ \tau_{jn}) \circ \tau_{jn}^{-1} = \tau^1 \circ \tau^2 \circ \ldots \circ \tau^q \circ \tau_{jn}^{-1}$.

Hence π is the product of at most $q + 1 \leq n - 1$ transpositions.

This completes the proof by induction on n. □

Adjacency Transpositions and Their Products, I

Definition

For each $k \in \{1, 2, \ldots, n-1\}$, the transposition $\tau_{k,k+1}$ of element k with its successor is an adjacency transposition. □

Definition

For each pair $k, \ell \in \mathbb{N}_n$ with $k < \ell$, define:

1. $\pi^{k \nearrow \ell} := \tau_{\ell-1,\ell} \circ \tau_{\ell-2,\ell-1} \circ \ldots \circ \tau_{k,k+1} \in \Pi_n$
as the composition of $\ell - k$
successive adjacency transpositions in order,
starting with $\tau_{k,k+1}$ and ending with $\tau_{\ell-1,\ell}$;
2. $\pi^{\ell \searrow k} := \tau_{k,k+1} \circ \tau_{k+1,k+2} \circ \ldots \circ \tau_{\ell-1,\ell} \in \Pi_n$
as the composition of the same $\ell - k$
successive adjacency transpositions in reverse order.

Adjacency Transpositions and Their Products, II

Exercise

For each pair $k, \ell \in \mathbb{N}_n$ with $k < \ell$, prove that:

- $$\pi^{k \nearrow \ell}(i) := \begin{cases} i & \text{if } i < k \text{ or } i > \ell; \\ i - 1 & \text{if } k < i \leq \ell; \\ \ell & \text{if } i = k. \end{cases}$$
- $\pi^{k \nearrow k} = \pi^{k \searrow k} = \iota$
- $\pi^{k \nearrow \ell}$ *and* $\pi^{\ell \searrow k}$ *are inverses*
- $\pi^{k \nearrow \ell} = \pi^{1,2,\ldots,k-1,k+1,\ldots,\ell-1,\ell,k,\ell+1,\ldots,n}$
- $\pi^{\ell \searrow k} = \pi^{1,2,\ldots,k-1,\ell,k,k+1,\ldots,\ell-2,\ell-1,\ell+1,\ldots,n}$ □

1. Note that $\pi^{k \nearrow \ell}$ moves k up to the ℓth position, while moving each element between $k + 1$ and ℓ down by one.
2. By contrast, $\pi^{\ell \searrow k}$ moves ℓ down to the kth position, while moving each element between k and $\ell - 1$ up by one.

Reduction to the Product of Adjacency Transpositions

Lemma

For each pair $k, \ell \in \mathbb{N}_n$ with $k < \ell$, the transposition $\tau_{k\ell}$ equals both $\pi^{\ell-1 \searrow k} \circ \pi^{k \nearrow \ell}$ and $\pi^{k+1 \nearrow \ell} \circ \pi^{\ell \searrow k}$, the compositions of $2(\ell - k) - 1$ adjacency transpositions.

Proof.

1. As noted, $\pi^{k \nearrow \ell}$ moves k up to the ℓth position, while moving each element between $k + 1$ and ℓ down by one.

   Then $\pi^{\ell-1 \searrow k}$ moves ℓ, which $\pi^{k \nearrow \ell}$ left in position $\ell - 1$, down to the k position, and moves $k + 1, k + 2, \ldots, \ell - 1$ up by one, back to their original positions.

   This proves that $\pi^{\ell-1 \searrow k} \circ \pi^{k \nearrow \ell} = \tau_{k\ell}$.

   It also expresses $\tau_{k\ell}$ as the composition of $(\ell - k) + (\ell - 1 - k) = 2(\ell - k) - 1$ adjacency transpositions.
2. The proof that $\pi^{k+1 \nearrow \ell} \circ \pi^{\ell \searrow k} = \tau_{k\ell}$ is similar; details are left as an exercise. □

Outline

The Inversions of a Permutation

Definition

1. Let $\mathbb{N}_{n,2} = \{S \subseteq \mathbb{N}_n \mid \#S = 2\}$ denote the set of all (unordered) pair subsets of $\mathbb{N}_n$.
2. Obviously, if $\{i, j\} \in \mathbb{N}_{n,2}$, then $i \neq j$.
3. Given any pair $\{i, j\} \in \mathbb{N}_{n,2}$, define

   $$i \vee j := \max\{i, j\} \quad \text{and} \quad i \wedge j := \min\{i, j\}$$

   For all $\{i, j\} \in \mathbb{N}_{n,2}$, because $i \neq j$, one has $i \vee j > i \wedge j$.
4. Given any permutation $\pi \in \Pi_n$,
   the pair $\{i, j\} \in \mathbb{N}_{n,2}$ is an inversion of π
   just in case π "reorders" $\{i, j\}$
   in the sense that $\pi(i \vee j) < \pi(i \wedge j)$.
5. Denote the set of inversions of π by

   $$\mathfrak{N}(\pi) := \{\{i, j\} \in \mathbb{N}_{n,2} \mid \pi(i \vee j) < \pi(i \wedge j)\}$$

The Sign of a Permutation

Definition

1. Given any permutation $\pi : \mathbb{N}_n \to \mathbb{N}_n$,
   let $\mathfrak{n}(\pi) := \#\mathfrak{N}(\pi) \in \mathbb{N} \cup \{0\}$
   denote the number of its inversions.
2. A permutation $\pi : \mathbb{N}_n \to \mathbb{N}_n$ is either even or odd
   according as $\mathfrak{n}(\pi)$ is an even or odd number.
3. The sign or signature of a permutation π,
   is defined as $\text{sgn}(\pi) := (-1)^{\mathfrak{n}(\pi)}$, which is:
   (i) $+1$ if π is even; (ii) -1 if π is odd.

The Sign of an Adjacency Transposition

Theorem

For each $k \in \mathbb{N}_{n-1}$, if π is the adjacency transposition $\tau_{k,k+1}$, then $\mathfrak{N}(\pi) = \{\{k, k+1\}\}$, so $\mathfrak{n}(\pi) = 1$ and $\operatorname{sgn}(\pi) = -1$.

Proof.

If π is the adjacency transposition $\tau_{k,k+1}$, then

$$\pi(i) = \begin{cases} i & \text{if } i \notin \{k, k+1\} \\ k+1 & \text{if } i = k \\ k & \text{if } i = k+1 \end{cases}$$

It is evident that $\{k, k+1\}$ is an inversion.

Also $\pi(i) \leq i$ for all $i \neq k$, and $\pi(j) \geq j$ for all $j \neq k+1$.

So if $i < j$, then $\pi(i) \leq i < j \leq \pi(j)$ unless $i = k$ and $j = k+1$, and so $\pi(i) > \pi(j)$ only if $(i, j) = (k, k+1)$.

Hence $\mathfrak{N}(\pi) = \{\{k, k+1\}\}$, implying that $\mathfrak{n}(\pi) = 1$. □

A Multi-Part Exercise

Exercise

Show that:

1. *For each permutation $\pi \in \Pi_n$, one has*

$$\begin{aligned} \mathfrak{N}(\pi) &:= \{\{i,j\} \in \mathbb{N}_{n,2} \mid (i-j)[\pi(i)-\pi(j)] < 0\} \\ &= \left\{\{i,j\} \in \mathbb{N}_{n,2} \mid \frac{\pi(i)-\pi(j)}{i-j} < 0\right\} \end{aligned}$$

2. $\mathfrak{n}(\pi) = 0 \iff \pi = \iota$, *the identity permutation;*
3. $\mathfrak{n}(\pi) \leq \frac{1}{2}n(n-1)$, *with equality if and only if π is the reversal permutation defined by $\pi(i) = n - i + 1$ for all $i \in \mathbb{N}_n$ — i.e.,*

$$(\pi(1), \pi(2), \ldots, \pi(n-1), \pi(n)) = (n, n-1, \ldots, 2, 1)$$

Hint: *Consider the number of ordered pairs $(i,j) \in \mathbb{N}_n \times \mathbb{N}_n$ that satisfy $i < j$.*

Outline

Double Products

Let $\mathbf{X} = \langle x_{ij} \rangle_{(i,j) \in \mathbb{N}_n \times \mathbb{N}_n}$ denote an $n \times n$ matrix.
We introduce the notation

$$\prod_{i>j}^{n} x_{ij} := \prod_{i=1}^{n} \prod_{j=1}^{n-1} x_{ij} := \prod_{j=1}^{n} \prod_{i=j+1}^{n} x_{ij}$$

for the product of all the elements in the lower triangular matrix $\mathbf{L}$ with elements $\ell_{ij} := \begin{cases} x_{ij} & \text{if } i > j \\ 0 & \text{if } i \leq j \end{cases}$

In case the matrix $\mathbf{X}$ is symmetric, one has

$$\prod_{i>j}^{n} x_{ij} = \prod_{i>j}^{n} x_{ji} = \prod_{i<j}^{n} x_{ij}$$

This can be rewritten as $\prod_{i>j}^{n} x_{ij} = \prod_{\{i,j\} \in \mathbb{N}_{n,2}} x_{ij}$,
which is the product over all unordered pairs of elements in $\mathbb{N}_n$.

Preliminary Example and Definition

Example

For every $n \in \mathbb{N}$, define the double product

$$\mathbb{P}_{n,2} := \prod_{\{i,j\} \in \mathbb{N}_{n,2}} |i - j| = \prod_{i>j}^{n} |i - j| = \prod_{i<j}^{n} |i - j|$$

Then one has

$$\begin{aligned} \mathbb{P}_{n,2} &= (n-1)\,(n-2)^2\,(n-3)^3 \cdots 3^{n-3}\,2^{n-2}\,1^{n-1} \\ &= \prod_{k=1}^{n-1} k^{n-k} \\ &= (n-1)!\,(n-2)!\,(n-3)! \cdots 3!\,2! \;=\; \prod_{k=1}^{n-1} k! \end{aligned}$$

Definition

For every permutation $\pi \in \Pi_n$, define the symmetric matrix $\mathbf{X}^\pi$

so that $x_{ij}^{\pi} := \begin{cases} \dfrac{\pi(i) - \pi(j)}{i - j} & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$

Basic Lemma

Lemma

For every permutation $\pi \in \Pi_n$*, one has* $\text{sgn}(\pi) = \prod_{\{i,j\} \in \mathbb{N}_{n,2}} x_{ij}^\pi$.

Proof.

- Because π is a permutation,
  the mapping $\mathbb{N}_{n,2} \ni \{i,j\} \mapsto \{\pi(i), \pi(j)\} \in \mathbb{N}_{n,2}$
  has inverse $\mathbb{N}_{n,2} \ni \{i,j\} \mapsto \{\pi^{-1}(i), \pi^{-1}(j)\} \in \mathbb{N}_{n,2}$.

  In fact it is a bijection between $\mathbb{N}_{n,2}$ and itself.
- Hence $\mathbb{P}_{n,2} := \prod_{\{i,j\} \in \mathbb{N}_{n,2}} |i-j| = \prod_{\{i,j\} \in \mathbb{N}_{n,2}} |\pi(i) - \pi(j)|$.
- So $\prod_{\{i,j\} \in \mathbb{N}_{n,2}} \dfrac{|\pi(i) - \pi(j)|}{|i-j|} = \prod_{\{i,j\} \in \mathbb{N}_{n,2}} |x_{ij}^\pi| = 1$.
- Also $x_{ij}^\pi = \mp 1$ according as $\{i,j\}$ is or is not a reversal of π.
- It follows that
  $\prod_{\{i,j\} \in \mathbb{N}_{n,2}} x_{ij}^\pi = (-1)^{\mathfrak{n}(\pi)} \prod_{\{i,j\} \in \mathbb{N}_{n,2}} |x_{ij}^\pi| = (-1)^{\mathfrak{n}(\pi)} = \text{sgn}(\pi)$

□

The Product Rule for Signs of Permutations

Theorem

For all permutations $\rho, \pi \in \Pi_n$ *one has* $\operatorname{sgn}(\rho \circ \pi) = \operatorname{sgn}(\rho) \operatorname{sgn}(\pi)$.

Proof.

The basic lemma implies that

$$\begin{aligned}
\frac{\operatorname{sgn}(\rho \circ \pi)}{\operatorname{sgn}(\pi)} &= \prod_{\{i,j\} \in \mathbb{N}_{n,2}} \frac{\rho(\pi(i)) - \rho(\pi(j))}{i - j} \prod_{\{k,\ell\} \in \mathbb{N}_{n,2}} \frac{k - \ell}{\pi(k) - \pi(\ell)} \\
&= \prod_{\{i,j\} \in \mathbb{N}_{n,2}} \frac{\rho(\pi(i)) - \rho(\pi(j))}{i - j} \prod_{\{i,j\} \in \mathbb{N}_{n,2}} \frac{i - j}{\pi(i) - \pi(j)}
\end{aligned}$$

After cancelling the product $\prod_{\{i,j\} \in \mathbb{N}_{n,2}} (i - j)$
and then replacing $\pi(i)$ by k and $\pi(j)$ by ℓ,
because π and ρ are permutations, one obtains

$$\frac{\operatorname{sgn}(\rho \circ \pi)}{\operatorname{sgn}(\pi)} = \prod_{\{k,\ell\} \in \mathbb{N}_{n,2}} \frac{\rho(k) - \rho(\ell)}{k - \ell} = \operatorname{sgn}(\rho) \qquad \square$$

The Sign of the Inverse Permutations

Corollary

Given any permutation $\pi \in \Pi_n$*, one has* $\text{sgn}(\pi^{-1}) = \text{sgn}(\pi)$.

Proof.

Because the identity permutation satisfies $\iota = \pi \circ \pi^{-1}$,
the product rule implies that

$$1 = \text{sgn}(\iota) = \text{sgn}(\pi \circ \pi^{-1}) = \text{sgn}(\pi)\,\text{sgn}(\pi^{-1})$$

Because $\text{sgn}(\pi), \text{sgn}(\pi^{-1}) \in \{-1, 1\}$,
they must both have the same sign, and the result follows. □

Outline

Determinants of Order 2: Definition

Consider again the pair of linear equations

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

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

$$x_1 = \frac{1}{D}(b_1 a_{22} - b_2 a_{12}), \quad x_2 = \frac{1}{D}(b_2 a_{11} - b_1 a_{21})$$

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

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 the two rows or columns of $\mathbf{A}$
changes the sign of $|\mathbf{A}|$.

Sign Corrected 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) $\hat{\mathbf{T}}^\top = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

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

Cramer's Rule in the 2×2 Case

Using determinant notation, the solution to the equations

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

can be written in the alternative form

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

This accords with Cramer's rule,
which says that the solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ 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 ith column,
   then replacing it with the $\mathbf{b}$ vector of right-hand side elements,
   while keeping all the columns in their original order.

Outline

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

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

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

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

is 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 rth row (a_{r1}, a_{r2}, a_{r3})
- the sth 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 $|\mathbf{A}| = \sum_{\pi \in \Pi_3} \operatorname{sgn}(\pi) \prod_{i=1}^{3} a_{i\pi(i)}$
for the sign function $\Pi_3 \ni \pi \mapsto \operatorname{sgn}(\pi) \in \{-1, +1\}$.

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

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

Exercise

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

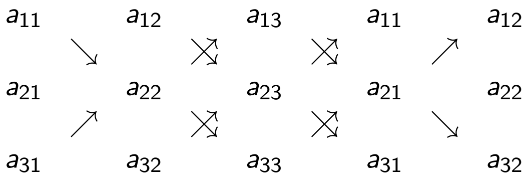
1. *calculating the number of inversions directly;*
2. *expressing each π as the product of transpositions, and then counting these.*

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:

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

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



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

The Determinant Mapping

Let $\mathcal{D}_n$ denote the domain $\mathbb{R}^{n \times n}$ of $n \times n$ matrices.

Definition

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

$$\mathcal{D}_n \ni \mathbf{A} \mapsto |\mathbf{A}| := \sum\nolimits_{\pi \in \Pi_n} \operatorname{sgn}(\pi) \prod\nolimits_{i=1}^{n} a_{i\pi(i)} \in \mathbb{R}$$

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

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

Row Mappings

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

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

defined on the domain $\mathcal{D}_n$ of $n \times n$ matrices $\mathbf{A}$
with row vectors $\langle \mathbf{a}_i^\top \rangle_{i=1}^n$.

Notation: For each fixed $r \in \mathbb{N}_n$, let $D(\mathbf{A}/\mathbf{b}_r^\top)$
denote the new value $D(\mathbf{a}_1^\top, \ldots, \mathbf{a}_{r-1}^\top, \mathbf{b}_r^\top, \mathbf{a}_{r+1}^\top, \ldots, \mathbf{a}_n^\top)$
of the row mapping D after the rth row $\mathbf{a}_r^\top$ of the matrix $\mathbf{A}$
has been replaced by the new row vector $\mathbf{b}_r^\top \in \mathbb{R}^n$.

Row Multilinearity

Definition

The function $\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A})$ 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 $i \in \{1, 2, \ldots, n\}$, each pair $\mathbf{b}_i^\top, \mathbf{c}_i^\top \in \mathbb{R}^n$ of new versions of row i, and 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 ith 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 determinant mapping

$$\mathcal{D}_n \ni \mathbf{A} \mapsto |\mathbf{A}| := \sum\nolimits_{\pi \in \Pi_n} \operatorname{sgn}(\pi) \prod\nolimits_{i=1}^{n} a_{i\pi(i)} \in \mathbb{R}$$

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

Proof.

For each fixed row $r \in \mathbb{N}$, we have

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

as required for multilinearity. □

Outline

Permutation Matrices: Definition

Definition

Given any permutation $\pi \in \Pi_n$ on $\{1, 2, \ldots, n\}$,
define $\mathbf{P}^\pi$ as the $n \times n$ permutation matrix
whose elements satisfy $p^\pi_{\pi(i),j} = \delta_{i,j}$ or equivalently $p^\pi_{i,j} = \delta_{\pi^{-1}(i),j}$.

That is, the rows of the identity matrix $\mathbf{I}_n$ are permuted
so that for each $i = 1, 2, \ldots, n$,
its ith row vector is moved to become row $\pi(i)$ of $\mathbf{P}^\pi$. □

Lemma

For each permutation matrix $\mathbf{P}^\pi$ *one has* $(\mathbf{P}^\pi)^\top = \mathbf{P}^{\pi^{-1}}$.

Proof.

Because π is a permutation, $i = \pi(j) \iff j = \pi^{-1}(i)$.

Then the definitions imply that for all $(i, j) \in \mathbb{N}_n^2$ one has

$$(\mathbf{P}^\pi)^\top_{i,j} = p^\pi_{j,i} = \delta_{\pi(j),i} = \delta_{\pi^{-1}(i),j} = p^{\pi^{-1}}(i, j)$$ □

Permutation Matrices: 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 are respectively $+1$ and -1.

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 are respectively $+1$, -1, -1, $+1$, $+1$ and -1.

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 $\mathbf{P}^\pi$ 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

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

and also

$$(\mathbf{A}\mathbf{P}^\pi)_{i,\pi(j)} = \sum\nolimits_{k=1}^{n} a_{ik} p^\pi_{k,\pi(j)} = \sum\nolimits_{k=1}^{n} a_{ik} \delta_{kj} = a_{ij} \qquad \square$$

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

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

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

□

Corollary

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

Proof.

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

Any Permutation Matrix Is Orthogonal

Proposition

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

Proof.

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

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

□

Transposition Matrices

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

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

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

Exercise

Let $\mathbf{A}$ be any $n \times n$ matrix. Prove that:
1) any transposition matrix $\mathbf{T}_{rs}$ is symmetric and orthogonal;
2) $\mathbf{T}_{rs} = \mathbf{T}_{sr}$; 3) $\mathbf{T}_{rs}\mathbf{T}_{sr} = \mathbf{T}_{sr}\mathbf{T}_{rs} = \mathbf{I}$;
4) $\mathbf{T}_{rs}\mathbf{A}$ is $\mathbf{A}$ with rows r and s interchanged;
5) $\mathbf{A}\mathbf{T}_{rs}$ 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 $\pi \in \mathbb{N}_n$, one has $|\mathbf{P}^\pi \mathbf{A}| = |\mathbf{A}\mathbf{P}^\pi| = \operatorname{sgn}(\pi)\,|\mathbf{A}|$.

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 Transpositions

We define the sign adjusted transposition matrix $\hat{\mathbf{T}}_{rs}$
as either one of the two matrices that:
(i) swaps rows or columns r and s;
(ii) then multiplies one, but only one,
of the two swapped rows or columns by -1.

As the matrix $\mathbf{I}$ with rows r and s transposed,
and then one sign changed, it satisfies

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

where $\alpha_r, \alpha_s \in \{-1, +1\}$ with $\alpha_r = -\alpha_s$.

It evidently satisfies $|\hat{\mathbf{T}}_{rs}\mathbf{A}| = |\mathbf{A}\hat{\mathbf{T}}_{rs}| = |\mathbf{A}|$.

Sign Adjusted Permutations

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.

Outline

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

Products of Upper Triangular Matrices

Theorem

The product $\mathbf{W} = \mathbf{U}\mathbf{V}$ *of any two upper triangular matrices* $\mathbf{U}, \mathbf{V}$ *is upper triangular,*
with diagonal elements $w_{ii} = u_{ii} v_{ii}$ $(i = 1, \ldots, n)$ *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 $\mathbf{W} = \mathbf{U}\mathbf{V}$ 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 $\mathbf{W} = \mathbf{U}\mathbf{V}$ is upper triangular.

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

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) iff $q < r$; (ii) iff $q > r$.

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 $(\mathbf{L}\mathbf{M})^\top = \mathbf{M}^\top \mathbf{L}^\top = \mathbf{U}$, where the product $\mathbf{U}$ is upper triangular, as the product of upper triangular matrices.

Hence $\mathbf{L}\mathbf{M} = \mathbf{U}^\top$ 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 $\{1, 2, \ldots, n\}$.

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 $i = 1, 2, \ldots, n$.

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) is invertible if and only if no element on its principal diagonal is 0.