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

Peter J. Hammond

revised 2018 September 16th

Lecture Outline

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 $\mathbf{B}$;*
- $\mathbf{A} \mathbf{I}_n = \mathbf{A}$ *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 $\mathbf{X} = \mathbf{XAY} = \mathbf{Y}$.

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

Also, if $\mathbf{Y}'$ is any alternative right inverse,
then $\mathbf{AY}' = \mathbf{I}$ and so $\mathbf{Y}' = \mathbf{XAY}' = \mathbf{X} = \mathbf{Y}$.

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

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

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 this result 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.*

Outline

Partitioned Matrices: Definition

A partitioned matrix is a rectangular array of different matrices.

Example

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

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

where 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: Matrix $\mathbf{D}$ may no longer 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

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

are both $k \times \ell$ arrays of respective $m_i \times n_j$ matrices $\mathbf{A}_{ij}, \mathbf{B}_{ij}$.

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. $\square$

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,

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

Adjacency Transpositions

Definition

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

Reduction to Adjacency Transpositions

Lemma

The transposition $\tau_{k\ell}$ *(with* $k < \ell$*)*
is the composition of $2(\ell - k) - 1$ *adjacency transpositions.*

Proof.

First, we apply in order the $\ell - k$
successive adjacency transpositions $\tau_{k,k+1}, \tau_{k+1,k+2}, \ldots, \tau_{\ell-1,\ell}$.
The result is the permutation $\pi^{1,2,\ldots,k-1,k+1,k+2,\ldots,\ell,k,\ell+1,\ldots,n}$
that moves k to the ℓth position,
while moving each element between $k+1$ and ℓ down one.

Second, we apply in order the $\ell - k - 1$
successive adjacency transpositions $\tau_{\ell-1,\ell}, \tau_{\ell-2,\ell-1} \cdots, \tau_{k,k+1}$
that moves ℓ to the kth position,
while moving each element $k+1, k+2, \ldots, \ell-1$
back up one position to where it was originally.

The final result is the transposition $\tau_{k\ell}$. □

Permutations are Products of Transpositions

Theorem

Any permutation $\pi \in \Pi_n$ on $\mathbb{N}_n := \{1, 2, \dots, n\}$ 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. $\square$

Outline

Inversions and Signs of Permutations

Definition

1. Given any permutation $\pi \in \Pi_n$, an inversion of π is a pair $(i, j) \in \mathbb{N}_n$ that π "reorders" in the sense that $i < j$ and $\pi(i) > \pi(j)$.
2. For each permutation $\pi \in \Pi_n$, let

$$\mathfrak{N}(\pi) := \{(i, j) \in \mathbb{N}_n \times \mathbb{N}_n \mid (i - j)[\pi(i) - \pi(j)] < 0\}$$

   denote the set of inversions of π.
3. Define $\mathfrak{n}(\pi) := \#\mathfrak{N}(\pi) \in \mathbb{N} \cup \{0\}$ as the number of inversions.
4. 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.
5. The sign or signature of a permutation π, is defined as $\operatorname{sgn}(\pi) := (-1)^{\mathfrak{n}(\pi)}$, which is: (i) $+1$ if π is even; (ii) -1 if π is odd.

Three Exercises

Exercise

Show that:

1. $\mathfrak{n}(\pi) = 0 \iff \pi = \iota$, *the identity permutation;*
2. *if* π *is an adjacency transposition* $\tau_{k,k+1}$*,*
   then $\mathfrak{n}(\pi) = 1$ *and so* $\operatorname{sgn}(\pi) = -1$*;*
3. $\mathfrak{n}(\pi) \leq \frac{1}{2}n(n-1)$, *with equality*
   if and only if $\pi(i) = n - i + 1$ *for all* $i \in \mathbb{N}_n$.

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

The Sign of the Inverse Permutation

Theorem

For the inverse π^{-1} of any permutation π,
one has: (i) $\mathfrak{n}(\pi^{-1}) = \mathfrak{n}(\pi)$*; (ii)* $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$.

Proof.

For Part (i), the number of inversions of π can be calculated as

$$\begin{aligned}
\mathfrak{n}(\pi) &= \#\{(i,j,k,\ell) \in \mathbb{N}_n^4 \mid i < j \ \&\ k = \pi(i) > \pi(j) = \ell\} \\
&= \#\{(i,j,k,\ell) \in \mathbb{N}_n^4 \mid \pi^{-1}(k) = i < j = \pi^{-1}(\ell) \ \&\ k > \ell\} \\
&= \#\{(\ell,k,j,i) \in \mathbb{N}_n^4 \mid \ell < k \ \&\ j = \pi^{-1}(\ell) > \pi^{-1}(k) = i\}
\end{aligned}$$

Interchanging i with ℓ and j with k in the last expression gives

$$\mathfrak{n}(\pi) = \#\{(i,j,k,\ell) \in \mathbb{N}_n^4 \mid i < j \ \&\ k = \pi^{-1}(i) > \pi^{-1}(j) = \ell\}$$

which equals $\mathfrak{n}(\pi^{-1})$, by definition.

For Part (ii), notice how part (i)
implies that $\operatorname{sgn}(\pi^{-1}) = (-1)^{\mathfrak{n}(\pi^{-1})} = (-1)^{\mathfrak{n}(\pi)} = \operatorname{sgn}(\pi)$. $\square$

Implication of Last Proof

Exercise

Use the argument of the last proof to establish a bijection

$$\mathbb{N}_n^2 \supset \mathfrak{N}(\pi) \ni (i,j) \mapsto (\pi(i), \pi(j)) \in \mathfrak{N}(\pi^{-1}) \subset \mathbb{N}_n^2$$

between the two sets $\mathfrak{N}(\pi)$ *and* $\mathfrak{N}(\pi^{-1})$ *of inversions, whose inverse is the bijection*

$$\mathbb{N}_n^2 \supset \mathfrak{N}(\pi^{-1}) \ni (k,\ell) \mapsto (\pi^{-1}(k), \pi^{-1}(\ell)) \in \mathfrak{N}(\pi) \subset \mathbb{N}_n^2$$

Outline

Basic Lemma

Lemma

Given any $k \in \mathbb{N}_{n-1}$*, let* τ *be any adjacency transposition* $\tau_{k,k+1}$*, and* $\pi \in \Pi_n$ *any permutation on* $\mathbb{N}_n$*.*

Then $\operatorname{sgn}(\tau \circ \pi) = -\operatorname{sgn}(\pi)$*.*

The proof follows on the next 7 slides.

Commencing the Proof of the Basic Lemma

Throughout the following proof,
suppose that the labels $i, j \in \mathbb{N}_n$ of any pair $\{i, j\} \subset \mathbb{N}_n$
are chosen to satisfy $i < j$.

Our proof will treat three different main cases,
depending on whether the set $\{i, j\} \cap \{\pi^{-1}(k), \pi^{-1}(k+1)\}$
has 0, 2, or 1 member.

Case 1: Suppose that $\#\{i, j\} \cap \{\pi^{-1}(k), \pi^{-1}(k+1)\} = 0$,
or equivalently, that $\{\pi(i), \pi(j)\} \cap \{k, k+1\} = \emptyset$.

In this case $(\tau \circ \pi)(i) = \pi(i)$ and $(\tau \circ \pi)(j) = \pi(j)$,
so $\pi(i) > \pi(j) \iff (\tau \circ \pi)(i) > (\tau \circ \pi)(j)$.

Hence π and $\tau \circ \pi$ have identical sets of inversions (i, j)
that satisfy $\{i, j\} \cap \{\pi^{-1}(k), \pi^{-1}(k+1)\} = \emptyset$.

Proof: Case 2

Case 2: Suppose that $\#\{i, j\} \cap \{\pi^{-1}(k), \pi^{-1}(k+1)\} = 2$.

In this case there are two different subcases.

Subcase (a): This subcase occurs when $\pi^{-1}(k) < \pi^{-1}(k+1)$.

Here, because $i < j$, one has $i = \pi^{-1}(k)$ and $j = \pi^{-1}(k+1)$.

Then $\pi(i) = k < k+1 = \pi(j)$, so (i, j) is not an inversion of π.

Yet $(\tau \circ \pi)(i) = k+1 > k = (\tau \circ \pi)(j)$,
so (i, j) is an inversion of $\tau \circ \pi$.

Subcase (b): This subcase occurs when $\pi^{-1}(k) > \pi^{-1}(k+1)$.

Here, because $i < j$, one has $i = \pi^{-1}(k+1)$ and $j = \pi^{-1}(k)$.

Then $\pi(i) = k+1 > k = \pi(j)$, so (i, j) is an inversion of π.

Yet $(\tau \circ \pi)(i) = k < k+1 = (\tau \circ \pi)(j)$,
so (i, j) is not an inversion of $\tau \circ \pi$.

Proof: Case 3(a)

Case 3: Suppose that $\#\{i,j\} \cap \{\pi^{-1}(k), \pi^{-1}(k+1)\} = 1$.

We will show that π and $\tau \circ \pi$ have the same set of inversions.

To show this, we consider four different subcases.

Subcase (a): $\pi(i) = k$ and $\pi(j) \notin \{k, k+1\}$.

Here $(\tau \circ \pi)(i) = k+1$ and $(\tau \circ \pi)(j) = \pi(j)$.

Now, if (i,j) is an inversion of π, then $k > \pi(j)$,
implying that $(\tau \circ \pi)(i) = k+1 > k > \pi(j) = (\tau \circ \pi)(j)$,
so (i,j) is an inversion of $\tau \circ \pi$.

Conversely, if (i,j) is an inversion of $\tau \circ \pi$,
then $k+1 = (\tau \circ \pi)(i) > (\tau \circ \pi)(j) = \pi(j)$.

Because $\pi(j) \neq k$, it follows that $k > \pi(j)$.

So $\pi(i) = k > \pi(j)$, implying that (i,j) is an inversion of π.

Proof: Case 3(b)

Subcase (b): $\pi(i) = k + 1$ and $\pi(j) \notin \{k, k + 1\}$.

Here $(\tau \circ \pi)(i) = k$ and $(\tau \circ \pi)(j) = \pi(j)$.

Now, if (i, j) is an inversion of π, then $k + 1 > \pi(j)$.
Because $\pi(j) \neq k$, it follows that $k > \pi(j)$.
But then $(\tau \circ \pi)(i) = k > \pi(j) = (\tau \circ \pi)(j)$,
so (i, j) is an inversion of $\tau \circ \pi$.

Conversely, if (i, j) is an inversion of $\tau \circ \pi$,
then $\pi(i) = k + 1 > k = (\tau \circ \pi)(i) > (\tau \circ \pi)(j) = \pi(j)$.
It follows that (i, j) is an inversion of π.

Proof: Case 3(c)

Subcase (c): $\pi(j) = k$ and $\pi(i) \notin \{k, k+1\}$.

Here $(\tau \circ \pi)(j) = k + 1$ and $(\tau \circ \pi)(i) = \pi(i)$.

Now, if (i, j) is an inversion of π, then $\pi(i) > k$.

Because $\pi(i) \neq k + 1$, it follows that $\pi(i) > k + 1$.

Hence $(\tau \circ \pi)(i) = \pi(i) > k + 1 = (\tau \circ \pi)(j)$,
implying that (i, j) is an inversion of $\tau \circ \pi$.

Conversely, if (i, j) is an inversion of $\tau \circ \pi$,
then $\pi(i) = (\tau \circ \pi)(i) > (\tau \circ \pi)(j) = k + 1 > k = \pi(j)$,
so (i, j) is an inversion of π.

Proof: Case 3(d)

Subcase (d): $\pi(j) = k + 1$ and $\pi(i) \notin \{k, k + 1\}$.

Here $(\tau \circ \pi)(i) = \pi(i)$ and $(\tau \circ \pi)(j) = k$.

Now, if (i, j) is an inversion of π, then $\pi(i) > \pi(j) = k + 1$.
This implies that $(\tau \circ \pi)(i) = \pi(i) > k + 1 > k = (\tau \circ \pi)(j)$, so (i, j) is an inversion of $\tau \circ \pi$.

Conversely, if (i, j) is an inversion of $\tau \circ \pi$, then $\pi(i) = (\tau \circ \pi)(i) > (\tau \circ \pi)(j) = k$.
Because $\pi(i) \neq k + 1$, it follows that $\pi(i) > k + 1 = \pi(j)$.
Hence (i, j) is an inversion of π.

Completing the Proof

To summarize what has been proved so far:

1. The two permutations π and $\tau \circ \pi$ have identical sets of inversions (i, j) satisfying $\{i, j\} \neq \{\pi^{-1}(k), \pi^{-1}(k+1)\}$.
2. As for the unique pair (i, j) satisfying both $i < j$ and $\{i, j\} = \{\pi^{-1}(k), \pi^{-1}(k+1)\}$:
   2.1 in case $\pi^{-1}(k) < \pi^{-1}(k+1)$ and so $i = \pi^{-1}(k)$, $j = \pi^{-1}(k+1)$, the pair (i, j) is an inversion of $\tau \circ \pi$, but not of π;
   2.2 in case $\pi^{-1}(k) > \pi^{-1}(k+1)$ and so $i = \pi^{-1}(k+1)$, $j = \pi^{-1}(k)$, the pair (i, j) is an inversion of π, but not of $\tau \circ \pi$.

Hence $\mathfrak{n}(\tau \circ \pi) = \mathfrak{n}(\pi) \pm 1$ according as $\pi^{-1}(k) \lesseqgtr \pi^{-1}(k+1)$.

In particular, this implies that $\operatorname{sgn}(\tau \circ \pi) = -\operatorname{sgn}(\pi)$. $\square$

Outline

Signing the Product of Adjacency Transpositions

Lemma

If $\rho, \pi \in \Pi_n$ *where* ρ *is the product of* m *adjacency transpositions, then* $\text{sgn}(\rho \circ \pi) = (-1)^m \text{sgn}(\pi) = \text{sgn}(\rho)\,\text{sgn}(\pi)$.

Proof.

1. In case $m = 1$, apply the basic lemma once, while noting that $\text{sgn}(\rho) = -1$.
2. In case $m > 1$, apply the basic lemma iteratively m times to obtain $\text{sgn}(\rho \circ \pi) = (-1)^m \text{sgn}(\pi)$.

   Then put $\pi = \iota$ to obtain $\text{sgn}(\rho) = \text{sgn}(\rho \circ \iota) = (-1)^m$.

   Hence $\text{sgn}(\rho \circ \pi) = (-1)^m \text{sgn}(\pi) = \text{sgn}(\rho)\,\text{sgn}(\pi)$. $\square$

Signing the Product of General Transpositions

Lemma

If $\rho, \pi \in \Pi_n$ where ρ is the product of m general transpositions, then $\operatorname{sgn}(\rho \circ \pi) = (-1)^m \operatorname{sgn}(\pi) = \operatorname{sgn}(\rho) \operatorname{sgn}(\pi)$.

Proof.

1. In case $m = 1$ and ρ is the general transposition $\tau_{k,\ell}$, recall that $\tau_{k,\ell}$ is the product of $2(\ell - k) - 1$ adjacency transpositions.

   Because $2(\ell - k) - 1$ is odd, it follows that $\operatorname{sgn}(\rho \circ \pi) = (-1)^{2(\ell-k)-1} \operatorname{sgn}(\pi) = -\operatorname{sgn}(\pi)$.

   Then put $\pi = \iota$ to obtain $\operatorname{sgn}(\rho) = \operatorname{sgn}(\rho \circ \iota) = -1$.

   Hence $\operatorname{sgn}(\rho \circ \pi) = -\operatorname{sgn}(\pi) = \operatorname{sgn}(\rho) \operatorname{sgn}(\pi)$.

2. In case ρ is the product of m general transpositions, apply the result of Part 1 iteratively m times, while noting that $\operatorname{sgn}(\rho) = (-1)^m \operatorname{sgn}(\iota) = (-1)^m$.

$\square$

Signing Product and Inverse Permutations

Theorem

For all $\rho, \pi \in \Pi_n$ *one has* $\text{sgn}(\rho \circ \pi) = \text{sgn}(\rho)\,\text{sgn}(\pi)$*,*
and also $\text{sgn}(\pi^{-1}) = \text{sgn}(\pi)$*.*

Proof.

Recall that each permutation
is the product of finitely many transpositions.

Suppose that ρ and π
are the products of r and p transpositions respectively.

Then $\rho \circ \pi$ is the product of $r + p$ transpositions.

Hence $\text{sgn}(\rho \circ \pi) = (-1)^{r+p} = (-1)^r(-1)^p = \text{sgn}(\rho)\,\text{sgn}(\pi)$.

Then putting $\rho = \pi^{-1}$ implies that

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

It follows that $\text{sgn}(\pi^{-1}) \neq -\,\text{sgn}(\pi)$, so $\text{sgn}(\pi^{-1}) = \text{sgn}(\pi)$. $\square$

Signing a Special Permutation, I

For each $k, \ell \in \mathbb{N}_n$ with $k \leq \ell$,
define $\pi^{k \nearrow \ell} \in \Pi_n$ so that for each $i \in \mathbb{N}_n$ one has

$$\pi^{k \nearrow \ell}(i) := \begin{cases} i & \text{if } i < k \text{ or } i > \ell; \\ i - 1 & \text{if } k < i \leq \ell; \\ \ell & \text{if } i = k. \end{cases}$$

That is, $\pi^{k \nearrow \ell}$ moves k "up to" ℓ, and then to compensate, shifts each i between $k+1$ and ℓ "down one".

Evidently when $k = \ell$, one has $\pi^{k \nearrow k} = \iota$.

The lemma on the next slide shows that, provided $\ell < n$, because $\pi^{k \nearrow \ell+1}$ moves k up one more step to $\ell + 1$ rather than ℓ, one could achieve it by applying $\pi^{k \nearrow \ell}$ first, followed by $\tau_{\ell,\ell+1}$.

That is $\pi^{k \nearrow \ell+1} = \tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell}$.

Signing a Special Permutation, II

Lemma

For all $k, \ell \in \mathbb{N}_n$ with $k \leq \ell < n$, one has $\pi^{k \nearrow \ell+1} = \tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell}$.

Proof.

The definitions of $\pi^{k \nearrow \ell}$ and $\pi^{k \nearrow \ell+1}$ evidently imply that:

1. $\pi^{k \nearrow \ell+1}(k) = \ell + 1 = \tau_{\ell,\ell+1}(\ell) = (\tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell})(k)$
2. $\pi^{k \nearrow \ell+1}(\ell + 1) = \ell = \tau_{\ell,\ell+1}(\ell + 1) = (\tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell})(\ell + 1)$
3. if $i < k$ or $i > \ell + 1$, then
$$\pi^{k \nearrow \ell+1}(i) = i = \pi^{k \nearrow \ell}(i) = (\tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell})(i)$$
4. if $k < i \leq \ell$, then $k \leq i - 1 = \pi^{k \nearrow \ell}(i) < \ell$ and so
$$\pi^{k \nearrow \ell+1}(i) = i - 1 = \pi^{k \nearrow \ell}(i) = (\tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell})(i)$$

This proves that $\pi^{k \nearrow \ell+1}(i) = (\tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell})(i)$ for all $i \in \mathbb{N}_n$. $\square$

Signing a Special Permutation, III

Proposition

For all $k, \ell \in \mathbb{N}_n$ with $k < \ell < n$, one has:

1. $\pi^{k \nearrow \ell}$ *is the product* $\tau_{\ell-1,\ell} \circ \tau_{\ell-2,\ell-1} \circ \ldots \circ \tau_{k+1,k+2} \circ \tau_{k,k+1}$ *of $\ell - k$ successive adjacency transpositions;*
2. $\operatorname{sgn} \pi^{k \nearrow \ell} = (-1)^{\ell-k}$.

Proof.

The proof of Part 1 will be by induction on ℓ.

When $\ell = k$, one has $\pi^{k \nearrow k} = \iota$ by definition.

Suppose that $\tau_{\ell-1,\ell} \circ \tau_{\ell-2,\ell-1} \circ \ldots \circ \tau_{k+1,k+2} \circ \tau_{k,k+1}$.

Combining the result of the Lemma with this hypothesis implies

$$\pi^{k \nearrow \ell+1} = \tau_{\ell,\ell+1} \circ \pi^{k \nearrow \ell} = \tau_{\ell,\ell+1} \circ \tau_{\ell-1,\ell} \circ \ldots \circ \tau_{k+1,k+2} \circ \tau_{k,k+1}$$

This proves Part 1 by induction; Part 2 follows immediately. □

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 $D := a_{11}a_{22} - a_{21}a_{12}$.

Provided that $D \neq 0$, the 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

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

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 replacing its ith column
   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

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

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.*

Outline

The Determinant Function

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

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

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

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

For a general natural number $n \in \mathbb{N}$, consider any 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 function 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 $\mathbf{A}$'s n rows $\langle \mathbf{a}_i^\top \rangle_{i=1}^n$
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

$$D(\mathbf{A}/\lambda \mathbf{b}_i^\top + \mu \mathbf{c}_i^\top) = \lambda D(\mathbf{A}/\mathbf{b}_i^\top) + \mu D(\mathbf{A}/\mathbf{c}_i^\top)$$

Formally, the mapping $\mathbb{R}^n \ni \mathbf{a}_i^\top \mapsto D(\mathbf{A}/\mathbf{a}_i^\top) \in \mathbb{R}$
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

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

as required for multilinearity. $\square$

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 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

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

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:

$$\mathbf{P}^{123} = \mathbf{I}_3; \qquad \mathbf{P}^{132} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \quad \mathbf{P}^{213} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

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

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

Multiplying by a Matrix Permutes Its Rows or Columns

Lemma

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

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

That is, premultiplying $\mathbf{A}$ by $\mathbf{P}^\pi$ permutes the rows of $\mathbf{A}$, whereas postmultiplying $\mathbf{A}$ by $\mathbf{P}^\pi$ permutes the columns of $\mathbf{A}$.

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

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

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

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

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. $\square$

Another Multiplication Property

Proposition

If $\mathbf{P}^\pi$ is any $n \times n$ permutation matrix,
then $\mathbf{P}^\pi (\mathbf{P}^\pi)^\top = (\mathbf{P}^\pi)^\top \mathbf{P}^\pi = \mathbf{I}_n$.

Proof.

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

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

and also

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

Transposition Matrices

A special case of a permutation matrix is a transposition $\mathbf{T}_{rs}$ 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

Prove that:

1. *any transposition matrix* $\mathbf{T} = \mathbf{T}_{rs}$ *is symmetric;*
2. $\mathbf{T}_{rs} = \mathbf{T}_{sr}$;
3. $\mathbf{T}_{rs}\mathbf{T}_{sr} = \mathbf{T}_{sr}\mathbf{T}_{rs} = \mathbf{I}$;

Determinants with Permuted Rows

Theorem

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

Proof.

The expansion formula for determinants gives

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

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

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

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

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 $|\mathbf{T}_{rs}\mathbf{A}| = -|\mathbf{A}|$.

Proof.

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

Then, because $\mathrm{sgn}(\pi) = \mathrm{sgn}(\tau_{rs}) = -1$,
the equality $|\mathbf{P}^\pi \mathbf{A}| = \mathrm{sgn}(\pi)\,|\mathbf{A}|$ implies that $|\mathbf{T}_{rs}\mathbf{A}| = -|\mathbf{A}|$. $\square$

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

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. $\square$

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 $q > r$.

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 $|\mathbf{U}| = \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n} u_{i\pi(i)}$ 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 $\operatorname{sgn}(\iota) = +1$, the expansion reduces to the single term

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

This is the product of the n diagonal elements, as claimed. $\square$

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.