Lecture Notes: Matrix Algebra
Part B: Introduction to Matrices

Peter J. Hammond

revised 2025 September 17th;
typeset from `matrixAlgB25.tex`

Outline

Matrices as Rectangular Arrays

An $m \times n$ matrix $\mathbf{A} = (a_{ij})_{m \times n}$ is a (rectangular) array, such as

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} = ((a_{ij})_{i=1}^m)_{j=1}^n = ((a_{ij})_{j=1}^n)_{i=1}^m$$

Note that in a_{ij}, we write the row number i
before the column number j.

An $m \times 1$ matrix is a column vector with m rows and 1 column.

A $1 \times n$ matrix is a row vector with 1 row and n columns.

The $m \times n$ matrix $\mathbf{A}$ consists of:

n columns in the form of m-vectors
$$\mathbf{a}_j = (a_{ij})_{i=1}^m \in \mathbb{R}^m \text{ for } j = 1, 2, \ldots, n;$$
m rows in the form of n-vectors
$$\mathbf{a}_i^\top = (a_{ij})_{j=1}^n \in \mathbb{R}^n \text{ for } i = 1, 2, \ldots, m.$$

The Transpose of a Matrix

The transpose $\mathbf{A}^\top$ of the $m \times n$ matrix $\mathbf{A} = (a_{ij})_{m \times n}$
is defined as the $n \times m$ matrix

$$\mathbf{A}^\top = (a_{ij}^\top)_{n \times m} = (a_{ji})_{n \times m} = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \ldots & a_{nm} \end{pmatrix}$$

Thus the transposed matrix $\mathbf{A}^\top$ results from transforming
each column m-vector $\mathbf{a}_j = (a_{ij})_{i=1}^m$ $(j = 1, 2, \ldots, n)$ of $\mathbf{A}$
into the corresponding row m-vector $\mathbf{a}_j^\top = (a_{ji}^\top)_{i=1}^m$ of $\mathbf{A}^\top$.

Equivalently, for each $i = 1, 2, \ldots, m$,
the ith row n-vector $\mathbf{a}_i^\top = (a_{ij})_{j=1}^n$ of $\mathbf{A}$
is transformed into the ith column n-vector $\mathbf{a}_i = (a_{ji})_{j=1}^n$ of $\mathbf{A}^\top$.

Either way, one has $a_{ij}^\top = a_{ji}$ for all relevant pairs i, j.

Rows Before Columns

VERY Important Rule: Rows before columns!

This order really matters.

Reversing it gives a transposed matrix.

Exercise
Verify that the double transpose of any $m \times n$ matrix $\mathbf{A}$
satisfies $(\mathbf{A}^\top)^\top = \mathbf{A}$
— i.e., transposing a matrix twice recovers the original matrix.

Outline

Matrices: Introduction
Matrices and Their Transposes
Matrix Multiplication: Definition

Special Matrices
Square, Symmetric, and Diagonal Matrices
The Identity Matrix
The Inverse Matrix
Partitioned Matrices
Block Diagonal Matrices

Permutations and Their Signs
Permutations
Transpositions
Adjacency Transpositions
The Inversions and Sign of a Permutation
The Product Rule

Multiplying a Matrix by a Scalar

A scalar, usually denoted by a Greek letter,
is simply a member $\alpha \in \mathbb{F}$ of the algebraic field $\mathbb{F}$
over which the vector space is defined.

So when $\mathbb{F} = \mathbb{R}$, a scalar is a real number $\alpha \in \mathbb{R}$.

The product of any $m \times n$ matrix $\mathbf{A} = (a_{ij})_{m \times n}$
and any scalar $\alpha \in \mathbb{R}$
is the new $m \times n$ matrix denoted by $\alpha \mathbf{A} = (\alpha a_{ij})_{m \times n}$,
each of whose elements αa_{ij} results
from multiplying the corresponding element a_{ij} of $\mathbf{A}$ by α.

Matrix Multiplication

The matrix product of two matrices **A** and **B**
is defined (whenever possible) as the matrix $\mathbf{C} = \mathbf{AB} = (c_{ij})_{m \times n}$
whose element c_{ij} in row i and column j
is the inner product $c_{ij} = \mathbf{a}_i^\top \mathbf{b}_j$ of:

- the ith row vector $\mathbf{a}_i^\top$ of the first matrix **A**;
- the jth column vector $\mathbf{b}_j$ of the second matrix **B**.

$$
\begin{pmatrix}
a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\
\vdots & & \vdots & & \vdots \\
a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\
\vdots & & \vdots & & \vdots \\
a_{m1} & \cdots & a_{mj} & \cdots & a_{mn}
\end{pmatrix}
\begin{pmatrix}
b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\
\vdots & & \vdots & & \vdots \\
b_{j1} & \cdots & b_{2j} & \cdots & b_{1p} \\
\vdots & & \vdots & & \vdots \\
b_{n1} & \cdots & b_{nj} & \cdots & b_{np}
\end{pmatrix}
=
\begin{pmatrix}
c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\
\vdots & & \vdots & & \vdots \\
c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\
\vdots & & \vdots & & \vdots \\
c_{m1} & \cdots & c_{mj} & \cdots & c_{mp}
\end{pmatrix}
$$

$$\mathbf{a}_i^\top \qquad \cdot \qquad \mathbf{b}_j \qquad = \qquad c_{ij}$$

Again: rows before columns!

Compatibility for Matrix Multiplication, I

Note that the resulting matrix product **C** must have:

▶ as many rows as the first matrix **A**;

▶ as many columns as the second matrix **B**.

Yet again: rows before columns!

Compatibility for Matrix Multiplication, II

Question: when is this definition
of the matrix product $\mathbf{C} = \mathbf{AB}$ possible?

Answer: if and only if $\mathbf{A}$ has as many columns as $\mathbf{B}$ has rows.

This condition ensures that every inner product $\mathbf{a}_i^\top \mathbf{b}_j$ is defined,
which is true iff (if and only if) every row of $\mathbf{A}$
has exactly the same number of elements as every column of $\mathbf{B}$.

In this case, the two matrices $\mathbf{A}$ and $\mathbf{B}$
are compatible for multiplication.

Specifically, if $\mathbf{A}$ is $m \times \ell$ for some m,
then $\mathbf{B}$ must be $\ell \times n$ for some n.

Then the product $\mathbf{C} = \mathbf{AB}$ is $m \times n$,
with elements $c_{ij} = \mathbf{a}_i^\top \mathbf{b}_j = \sum_{k=1}^{\ell} a_{ik} b_{kj}$
for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Laws of Matrix Multiplication

Exercise

*Verify that, whenever the relevant matrix products are defined,
the following laws of matrix multiplication hold:*

associative law for matrices: $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$;

distributive: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ *and* $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$;

transpose: $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$.

associative law for scalars: $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$ *(all $\alpha \in \mathbb{R}$).*

Exercise

Let $\mathbf{X}$ be any $m \times n$ matrix, and $\mathbf{z}$ any column n-vector.

1. *Show that the matrix product $\mathbf{z}^\top \mathbf{X}^\top \mathbf{X} \mathbf{z}$ is well-defined,
   and that its value is a scalar.*

2. *By putting $\mathbf{w} = \mathbf{X}\mathbf{z}$ in the previous exercise regarding the sign
   of the quadratic form $\mathbf{w}^\top \mathbf{w}$, what can you conclude
   about the value of the scalar $\mathbf{z}^\top \mathbf{X}^\top \mathbf{X} \mathbf{z}$?*

Exercise for Econometricians I

Exercise

An econometrician has access to data
in the form of the real-valued time series:

- y_t *(t = 1, 2, . . . , T) of one endogenous variable;*
- x_{ti} *(t = 1, 2, . . . , T and i = 1, 2, . . . , k)*
  of k different exogenous variables
  — sometimes called explanatory variables or regressors.

The data is to be fitted to the linear regression model

$$y_t = \sum_{i=1}^{k} b_i x_{ti} + e_t$$

whose scalar constants b_i (i = 1, 2, . . . , k)
are unknown regression coefficients,
and each scalar e_t is the error term or residual.

Exercise for Econometricians II

1. Discuss how the regression model with $y_t = \sum_{i=1}^{k} b_i x_{ti} + e_t$
   can be written in the form $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$
   for suitable column vectors $\mathbf{y}, \mathbf{b}, \mathbf{e}$.

2. What are the dimensions of these vectors,
   and of the exogenous data matrix $\mathbf{X}$?

3. Why do you think econometricians use this matrix equation,
   rather than the alternative $\mathbf{y} = \mathbf{bX} + \mathbf{e}$?

4. How can the equation $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$
   accommodate the constant term α
   in the alternative equation $y_t = \alpha + \sum_{i=1}^{k} b_i x_{ti} + e_t$?

Matrix Multiplication Does Not Commute I

The two matrices **A** and **B** commute just in case **AB** = **BA**.

Note that typical pairs of matrices DO NOT commute,
meaning that **AB** $\neq$ **BA** — i.e., the order of multiplication matters.

Indeed, suppose that **A** is $\ell \times m$ and **B** is $m \times n$,
as is needed for **AB** to be defined as an $\ell \times n$ matrix.

Then the reverse product **BA** is undefined
except in the special case when $n = \ell$.

Hence, for both **AB** and **BA** to be defined,
where **B** is $m \times n$, the matrix **A** must be $n \times m$.

But then **AB** is $n \times n$, whereas **BA** is $m \times m$.

Evidently **AB** $\neq$ **BA** unless $m = n$.

Then all four matrices **A**, **B**, **AB** and **BA** are $m \times m = n \times n$.

To summarize, we must be in the special case
where **A** and **B** are two square matrices of the same dimension.

Matrix Multiplication Does Not Commute II

Even if both **A** and **B** are $n \times n$ matrices,
implying that both **AB** and **BA** are also $n \times n$,
one can still have $\mathbf{AB} \neq \mathbf{BA}$.

Example

Here is a 2×2 example:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Exercise

*For matrix multiplication, explain why there are two different
versions of the distributive law — namely*

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \text{ and } (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

More Warnings Regarding Matrix Multiplication

Exercise

Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ *denote three general matrices.*

Give examples showing that:

1. *The matrix* $\mathbf{AB}$ *might be defined, even if* $\mathbf{BA}$ *is not.*
2. *One can have* $\mathbf{AB} = \mathbf{0}$ *even though* $\mathbf{A} \neq \mathbf{0}$ *and* $\mathbf{B} \neq \mathbf{0}$.
3. *If* $\mathbf{AB} = \mathbf{AC}$ *and* $\mathbf{A} \neq \mathbf{0}$, *it does not follow that* $\mathbf{B} = \mathbf{C}$.

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

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

The extra dots indicate omitted elements along the diagonal.

Symmetric Matrices

Definition
A square matrix **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 **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 **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 **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., $i \neq j \implies a_{ij} = 0$.

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

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

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 $\mathbf{diag\,d} = (a_{ij})_{n \times n}$ where each $a_{ij} = \delta_{ij} d_i = \delta_{ij} d_j$,
using Kronecker delta notation.

Obviously, any diagonal matrix is symmetric.

Multiplying by Diagonal Matrices

Example

Let **D** be a diagonal matrix of dimension n.

Suppose that **A** and **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 **A** by **D** is the column operation
of simultaneously multiplying every column $\mathbf{a}_j$ of **A**
by its matching diagonal element d_{jj}.

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

Two Exercises with Diagonal Matrices

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

Exercise
*Let **D** be a diagonal matrix of dimension n,*
*and **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 **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 $\mathbb{N}_n \times \mathbb{N}_n \ni (i, j) \mapsto \delta_{ij} \in \{0, 1\}$.

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 of a square matrix exist?
Answer discussed later: if and only if its determinant is non-zero.

Rule for Inverting Products

Theorem

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

Then the inverse of the matrix product **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 $(\mathbf{AB})\mathbf{X} = \mathbf{X}(\mathbf{AB}) = \mathbf{I}$. $\qquad\square$

Rule for Inverting Chain Products

Exercise
Prove that, if **A**, **B** *and* **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* **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.*

Rule for Inverting Transposes

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

$$\text{both} \quad \mathbf{A}^\top(\mathbf{A}^{-1})^\top = (\mathbf{A}^{-1}\mathbf{A})^\top = \mathbf{I}^\top = \mathbf{I}$$
$$\text{and} \quad (\mathbf{A}^{-1})^\top\mathbf{A}^\top = (\mathbf{A}\mathbf{A}^{-1})^\top = \mathbf{I}^\top = \mathbf{I}$$

This proves that $(\mathbf{A}^{-1})^\top$
is both a left and a right inverse of $\mathbf{A}^\top$. □

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,
  one has $\mathbf{x}_i \cdot \mathbf{x}_i = \|\mathbf{x}_i\|^2 = 1$ and so $\|\mathbf{x}_i\| = 1$ for each $i \in \mathbb{N}_k$
  — i.e., all k elements of the set are vectors of unit length.

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

Proof.
The result is immediate from the definitions of the norm in $\mathbb{R}^n$,
as well as of an orthonormal set of vectors
and of the Kronecker delta function. □

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.

(The answer will come later
in the section on eigenvalues and eigenvectors.)

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

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

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

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

Inverting Partitioned Matrices: Two 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

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

*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

Block Diagonal Matrices: Definition

Definition
A block diagonal matrix is a partitioned square matrix
which is a diagonal $k \times k$ square array of "blocks"
in the form of $n_i \times n_i$ square matrices $\mathbf{A}^{(i)}_{n_i \times n_i}$, for $i \in \mathbb{N}_k$.

The array can be written as

$$\begin{pmatrix} \mathbf{A}^{(1)}_{n_1 \times n_1} & \mathbf{0}_{n_1 \times n_2} & \cdots & \mathbf{0}_{n_1 \times n_k} \\ \mathbf{0}_{n_2 \times n_1} & \mathbf{A}^{(2)}_{n_2 \times n_2} & \cdots & \mathbf{0}_{n_2 \times n_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_k \times n_1} & \mathbf{0}_{n_k \times n_2} & \cdots & \mathbf{A}^{(k)}_{n_k \times n_k} \end{pmatrix}$$

or, more succinctly, as $\mathbf{diag}(\mathbf{A}^{(1)}_{n_1 \times n_1}, \ldots, \mathbf{A}^{(k)}_{n_k \times n_k})$.

Products of Block Diagonal Matrices

Exercise

Suppose that the two block diagonal matrices

$$\mathbf{diag}(\mathbf{A}^{(1)}_{m_1 \times m_1}, \ldots, \mathbf{A}^{(k)}_{m_k \times m_k}) \quad and \quad \mathbf{diag}(\mathbf{B}^{(1)}_{n_1 \times n_1}, \ldots, \mathbf{B}^{(\ell)}_{n_\ell \times n_\ell})$$

*have compatible dimensions in the sense that $k = \ell$
and $m_i = n_i$ for all $i \in \mathbb{N}_k = \mathbb{N}_\ell$.*

Verify that then the two matrix products

$$\mathbf{diag}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(k)}) \; \mathbf{diag}(\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(k)})$$
$$and \quad \mathbf{diag}(\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(k)}) \; \mathbf{diag}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(k)})$$

both exist, and that they equal

$$\mathbf{diag}(\mathbf{A}^{(1)}\mathbf{B}^{(1)}, \ldots, \mathbf{A}^{(k)}\mathbf{B}^{(k)}) \; and \; \mathbf{diag}(\mathbf{B}^{(1)}\mathbf{A}^{(1)}, \ldots, \mathbf{B}^{(k)}\mathbf{A}^{(k)})$$

respectively.

The Inverse of a Block Diagonal Matrix

Exercise

Suppose the block diagonal matrix $\mathbf{diag}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(k)})$
has the property that each block $\mathbf{A}^{(i)}_{n_i \times n_i}$ ($i \in \mathbb{N}_k$) is invertible.

Show that then the block diagonal matrix is invertible, with inverse

$$\left[\mathbf{diag}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(k)})\right]^{-1} = \mathbf{diag}\left(\left[\mathbf{A}^{(1)}\right]^{-1}, \ldots, \left[\mathbf{A}^{(k)}\right]^{-1}\right)$$

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 $\pi(i) = \pi(j) \Longrightarrow i = j$ or,
   equivalently, $i \neq j \Longrightarrow \pi(i) \neq \pi(j)$.

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 $(\Pi_n, \circ, \iota)$ is an algebraic group
— i.e., the operation $\Pi_n \times \Pi_n \ni (\pi, \rho) \mapsto \pi \circ \rho \in \Pi_n$
is associative, with ι as 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

Let $(j_1, j_2, \ldots, j_n) = (j_k)_{k \in \mathbb{N}_n} \in \mathbb{N}_n^n$ denote
any list of n integers $(j_k)_{k \in \mathbb{N}_n}$ in $\mathbb{N}_n$,
or equivalently, any mapping $\mathbb{N}_n \ni k \mapsto j_k \in \mathbb{N}_n$.

Then any list $(j_k)_{k \in \mathbb{N}_n}$ whose components in $\mathbb{N}_n$ are all different
corresponds to a unique permutation, denoted by $\pi^{j_1 j_2 \cdots j_n} \in \Pi_n$,
that satisfies $\pi(k) = j_k$ for all $k \in \mathbb{N}_n^n$.

Example

Two transpositions defined
on a set containing more than two elements may not commute.

For example, one has

$$\tau_{12} \circ \tau_{23} = \tau_{12}(\pi^{132}) = \pi^{312} \quad \text{and} \quad \tau_{23} \circ \tau_{12} = \tau_{23}(\pi^{213}) = \pi^{231}$$

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 the Induction Step

For general n, let $j := \pi^{-1}(n)$ be the element 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, the permutation $\tilde{\pi}$ on $\mathbb{N}_{n-1}$
is the product $\tau^1 \circ \tau^2 \circ \ldots \circ \tau^q$ of $q \leq n-2$ transpositions.

Hence, for all $k \in \mathbb{N}_{n-1}$,
one has $\tilde{\pi}(k) = (\pi \circ \tau_{jn})(k) = (\tau^1 \circ \tau^2 \circ \ldots \circ \tau^q)(k)$.

Also, for each $p = 1, \ldots, q$,
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.

It follows that $\pi = (\pi \circ \tau_{jn}) \circ \tau_{jn}^{-1} = \tau^1 \circ \tau^2 \circ \ldots \circ \tau^q \circ \tau_{jn}^{-1}$.

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

Example Illustrating the Proof

Example

We find a set of transpositions
whose product is the particular permutation $\pi^{4132} \in \Pi_4$.

Using compressed notation, start with the transposition τ_{14}
which moves the 4 in 4132 to its right place in the end.

In fact τ_{14} applied to π^{4132} is π^{2134},
which we abbreviate to 2134, with 3 and 4 in the right place.

Last, we use τ_{12} to get 1234.

Reversing the order of these transpositions
gives us $\pi^{4132} = \tau_{14} \circ \tau_{12}$
because $(\tau_{14} \circ \tau_{12})(1234) = \tau_{14}(2134) = 4132$.

Outline

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 the right 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*

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

▶ $\pi^{\ell \searrow k} = \pi^{1,2,\ldots,k-1,\ell,k,k+1,\ldots,\ell-2,\ell-1,\ell+1,\ldots,n}$ $\qquad \square$

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 $\pi^{k+1 \nearrow \ell} \circ \pi^{\ell \searrow k}$.

Both are 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 kth 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. $\qquad \square$

Example Illustrating the Proof

Example

We illustrate how $\tau_{2,5} = \pi^{4\searrow 2} \circ \pi^{2\nearrow 5}$.
Indeed, applying $\pi^{2\nearrow 5}$ to 123456 gives successively

$$132456$$
$$134256$$
$$134526$$

Then, applying $\pi^{4\searrow 2}$ to 134526 gives successively

$$134526$$
$$135426$$
$$153426$$

The result 153426 is indeed $\tau_{2,5}$ applied to 123456.

Outline

The Inversions of a Permutation

Definition

1. Let $\mathbb{N}_{n,2}^< = \{(i,j) \in \mathbb{N}_n \times \mathbb{N}_n \mid i < j\}$ denote the set of all ordered pair subsets of $\mathbb{N}_n$.

2. Given any permutation $\pi \in \Pi_n$,
   the ordered pair $(i,j) \in \mathbb{N}_{n,2}^<$ is an inversion of π
   just in case π "reorders" $\{i,j\}$ in the sense that $\pi(i) > \pi(j)$.

   This implies that $(i - j)[\pi(i) - \pi(j)] < 0$
   because $i - j$ and $\pi(i) - \pi(j)$ have opposite signs.

3. Denote the set of inversions of π by

$$\mathcal{I}(\pi) := \{\{i,j\} \in \mathbb{N}_{n,2}^< \mid \pi(i) > \pi(j)\}$$

Note that an inversion of π is very different from its inverse!

The Sign or Parity of a Permutation

Definition

1. Given any integer $n \in \mathbb{Z}$, define the parity of n
   so that $\text{parity}(n) = \begin{cases} \text{even} & \text{if } \frac{1}{2}n \in \mathbb{Z} \\ \text{odd} & \text{if } \frac{1}{2}n \notin \mathbb{Z} \end{cases}$

2. Given any permutation $\pi : \mathbb{N}_n \to \mathbb{N}_n$,
   let $\mathfrak{n}(\pi) := \#\mathcal{I}(\pi) \in \mathbb{N} \cup \{0\}$
   denote the number of its inversions.

3. 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,
   so the parity of π is $\text{parity}(\mathfrak{n}(\pi))$,
   which is the parity of the natural number $\mathfrak{n}(\pi)$.

4. 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 a General Transposition

Theorem

For each transposition $\tau_{k,\ell}$ with $k, \ell \in \mathbb{N}_n$ and $k < \ell$:

1. *the set of inversions is*

$$\mathcal{I}(\tau_{k,\ell}) = \cup_{r \in \mathbb{N} \cap (k,\ell)} \{(k,r), (r,\ell)\} \cup \{(k,\ell)\}$$

*where $\mathbb{N} \cap (k,\ell)$ denotes
the intersection of $\mathbb{N}$ with the open interval (k,ℓ);*

2. $\mathfrak{n}(\tau_{k,\ell}) = 2(\ell - k) - 1$ *and* $\text{sgn}(\tau_{k,\ell}) = -1$.

Proof.

If π is the transposition $\tau_{k,\ell}$, then (k, ℓ) is evidently an inversion.

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

So if $i < j$, then $\pi(i) \leq i < j \leq \pi(j)$ unless $i = k$ or $j = \ell$.

But then (k, r) and (r, ℓ) are both inversions iff $k < r < \ell$.

The rest of the proof is straightforward. $\qquad\square$

A Two-Part Exercise and a Fundamental Lemma

Exercise

For each permutation $\pi \in \Pi_n$, show that:

1. $\mathfrak{n}(\pi) = 0 \Longleftrightarrow \pi = \iota$, *the identity permutation;*

2. $\mathfrak{n}(\pi) = 1$ *if and only if π is an adjacency transposition;*

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.,

$$(\pi(1), \pi(2), \ldots, \pi(n-1), \pi(n)) = (n, n-1, \ldots, 2, 1)$$
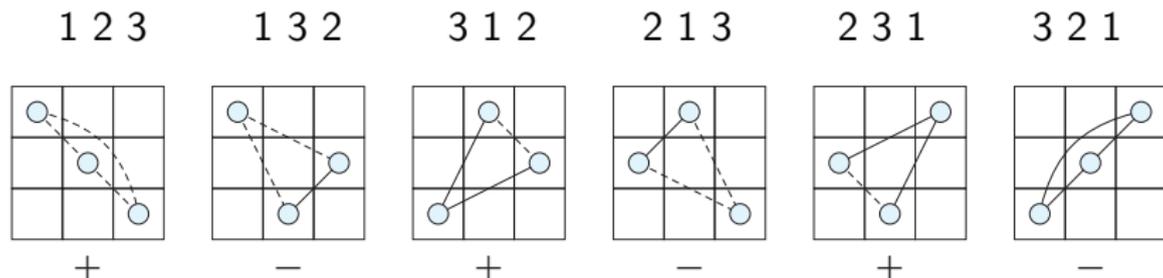
   Hint: *Consider the number*
   of ordered pairs $(i, j) \in \mathbb{N}_n \times \mathbb{N}_n$ that satisfy $i < j$.

Lemma

For each adjacency transposition $\tau_{r,r+1}$ with $r \in \mathbb{N}_{n-1}$
and each permutation $\pi \in \Pi_n$,
one has $\mathrm{sgn}(\pi \circ \tau_{r,r+1}) = -\,\mathrm{sgn}(\pi)$.

Examples of the Fundamental Lemma, I

Figure 13.3.1 of EMEA6 and 7, shown below, illustrates
the six possible permutations in Π_3, and links between pairs.

Each inversion is marked by a solid link; each other link is dashed.

Below each permutation is a $+$ or $-$ sign, indicating the parity,
which equals the parity of the number of inversions.

| 1 2 3 | 1 3 2 | 3 1 2 | 2 1 3 | 2 3 1 | 3 2 1 |
| $+$ | $-$ | $+$ | $-$ | $+$ | $-$ |

Interchanging two adjacent rows replaces
any positive dashed link with a negative solid link, and vice versa.

This changes the number of inversions by ± 1, reversing the parity.

Examples of the Fundamental Lemma, II

$$
\begin{array}{cccc}
a_{11} & a_{12} & \boxed{a_{13}} & a_{14} \\
\boxed{a_{21}} & a_{22} & a_{23} & a_{24} \\
a_{31} & \boxed{a_{32}} & a_{33} & a_{34} \\
a_{41} & a_{42} & a_{43} & \boxed{a_{44}}
\end{array}
$$

Consider $\pi \in \Pi_4$ defined so that each boxed element a_{ij}
in the above diagram corresponds to the pair $(i, \pi(i))$, for $i \in \mathbb{N}_4$.

So the diagram represents the permutation π^{3124}.

It has two inversions $(1, 2)$ and $(1, 3)$ where $j > i$ but $\pi(j) < \pi(i)$,
implying that $\mathrm{sgn}(\pi^{3124}) = (-1)^2 = +1$.

The adjacency transposition $\tau_{1,2}$ removes the inversion $(1, 2)$.

The adjacency transposition $\tau_{2,3}$ adds the inversion $(2, 3)$.

The adjacency transposition $\tau_{3,4}$ adds the inversion $(3, 4)$.

In each case the new matrix $\pi' = \pi \circ \tau_{r,r+1}$ has sign -1.

$$
\begin{array}{ccccc}
a_{11} & \boxed{a_{12}} & a_{13} & a_{14} & a_{15} \\
a_{21} & a_{22} & \boxed{a_{23}} & a_{24} & a_{25} \\
a_{31} & a_{32} & a_{33} & a_{34} & \boxed{a_{35}} \\
\boxed{a_{41}} & a_{42} & a_{43} & a_{44} & a_{45} \\
a_{51} & a_{52} & a_{53} & \boxed{a_{54}} & a_{55}
\end{array}
$$

Consider $\pi \in \Pi_5$ defined so that each boxed element a_{ij}
in the above diagram corresponds to the pair $(i, \pi(i))$, for $i \in \mathbb{N}_5$.

So the diagram represents the permutation π^{23514}.

It has four inversions $(1,4)$, $(2,4)$, $(3,4)$ and $(3,5)$
where $j > i$ but $\pi(j) < \pi(i)$, so $\operatorname{sgn}(\pi^{23514}) = (-1)^4 = +1$.

The three adjacency transpositions $\tau_{1,2}$, $\tau_{2,3}$, and $\tau_{3,4}$
each remove one inversion — namely $(1,2)$, $(2,3)$, and $(3,4)$.

The adjacency transposition $\tau_{4,5}$ adds the inversion $(4,5)$.

In each case the new matrix $\pi' = \pi \circ \tau_{r,r+1}$ has sign -1.

Restatement and Proof of Fundamental Lemma, I

Given permutation $\pi \in \Pi_n$ and adjacency transposition $\tau_{r,r+1}$,
we introduce the notation $\pi' = \pi \circ \tau_{r,r+1}$.

This implies that $\pi'(r) = \pi(r+1)$ and $\pi'(r+1) = \pi(r)$
but $\pi'(p) = \pi(p)$ and $\pi'(q) = \pi(q)$ for all $p < r$ and $q > r+1$.

Lemma

For each adjacency transposition $\tau_{r,r+1}$ with $r \in \mathbb{N}_{n-1}$
and each permutation $\pi \in \Pi_n$, one has $\mathrm{sgn}(\pi') = - \,\mathrm{sgn}(\pi)$.

Proof.

First, for all $r \in \mathbb{N}_{n-1}$ and all $(p,q) \in \mathbb{N}_n \setminus \{r, r+1\}$ with $p < q$,
one has $\tau_{r,r+1}(p) = p$ and $\tau_{r,r+1}(q) = q$.

This implies that $(p,q) \in \mathcal{I}(\pi) \Longleftrightarrow (p,q) \in \mathcal{I}(\pi')$.

Second, for the particular pair $(r, r+1)$ one has

$$(r, r+1) \in \mathcal{I}(\pi) \Longleftrightarrow \pi(r) > \pi(r+1)$$
$$\Longleftrightarrow \pi'(r+1) > \pi'(r) \Longleftrightarrow (r, r+1) \notin \mathcal{I}(\pi')$$

Proof of Fundamental Lemma, II

Proof.

To continue the proof, we consider all the four or fewer pairs
in $\{p, q\} \times \{r, r+1\}$ satisfying $p < r$ and $r + 1 < q$.

1. In case $p < r$, one has $(p, r) \in \mathcal{I}(\pi)$ if and only if

$$\pi(p) = \pi'(p) > \pi(r) = \pi'(r+1) \Longleftrightarrow (p, r+1) \in \mathcal{I}(\pi')$$

   and $(p, r+1) \in \mathcal{I}(\pi)$ if and only if

$$\pi(p) = \pi'(p) > \pi(r+1) = \pi'(r) \Longleftrightarrow (p, r) \in \mathcal{I}(\pi')$$

2. In case $r + 1 < q$, one has $(r, q) \in \mathcal{I}(\pi)$ if and only if

$$\pi(r) = \pi'(r+1) > \pi(q) = \pi'(q) \Longleftrightarrow (r+1, q) \in \mathcal{I}(\pi')$$

   and $(r + 1, q) \in \mathcal{I}(\pi)$ if and only if

$$\pi(r+1) = \pi'(r) > \pi(q) = \pi'(q) \Longleftrightarrow (r, q) \in \mathcal{I}(\pi') \quad \square$$

Proof of Fundamental Lemma, III

Proof.

To summarize, given any $\tau_{r,r+1}$ and any $\pi \in \Pi_n$,
as well as $\pi' = \pi \circ \tau_{r,r+1}$, we have proved that:

1. for each pair $(p,q) \in \mathbb{N}_n \setminus \{r, r+1\}$,
   one has $(p,q) \in \mathcal{I}(\pi) \Longleftrightarrow (p,q) \in \mathcal{I}(\pi')$;

2. for each $p < r$ and $q > r+1$, one has:
   (i) $(p,r) \in \mathcal{I}(\pi) \Longleftrightarrow (p, r+1) \in \mathcal{I}(\pi')$;
   (ii) $(p, r+1) \in \mathcal{I}(\pi) \Longleftrightarrow (p, r) \in \mathcal{I}(\pi')$;
   (iii) $(r, q) \in \mathcal{I}(\pi) \Longleftrightarrow (r+1, q) \in \mathcal{I}(\pi')$;
   (iv) $(r+1, q) \in \mathcal{I}(\pi) \Longleftrightarrow (r, q) \in \mathcal{I}(\pi')$;

3. $(r, r+1) \in \mathcal{I}(\pi) \Longleftrightarrow (r, r+1) \notin \mathcal{I}(\pi')$.

So, apart from the pair $(r, r+1)$, any other pair in $\mathcal{I}(\pi)$
is matched with a unique corresponding pair in $\mathcal{I}(\pi')$.

It follows that $|\mathfrak{n}(\pi) - \mathfrak{n}(\pi')| = |\#\mathcal{I}(\pi) - \#\mathcal{I}(\pi')| = 1$,
implying that $\mathrm{sgn}(\pi') \neq \mathrm{sgn}(\pi)$ and so $\mathrm{sgn}(\pi') = -\mathrm{sgn}(\pi)$. $\qquad\square$

Outline

The Parity of a Product of Adjacency Transpositions

Theorem

The sign $\text{sgn}(\pi^{(k)})$ *of the product* $\pi^{(k)} = \prod_{i\in\mathbb{N}_k} \tau_{r_i,r_i+1}$
of any finite collection of k adjacency transpositions satisfies

$$\text{sgn}(\pi^{(k)}) = \text{sgn}\left(\prod_{i\in\mathbb{N}_k} \tau_{r_i,r_i+1}\right) = (-1)^k$$

Proof.

We prove the result by induction on k.

Indeed, when $k = 1$ one has $\text{sgn}(\tau_{r_1,r_1+1}) = -1$.

As the induction hypothesis, suppose $\text{sgn}\left(\pi^{(k-1)}\right) = (-1)^{(k-1)}$
for $\pi^{(k-1)} = \prod_{i\in\mathbb{N}_{k-1}} \tau_{r_i,r_i+1}$.

Then the fundamental lemma implies that

$$\text{sgn}(\pi^{(k)}) = \text{sgn}\left(\pi^{(k-1)} \circ \tau_{r_k,r_k+1}\right) = -\text{sgn}\left(\pi^{(k-1)}\right) = (-1)^k \quad \square$$

The Parity of a General Transposition is Odd

Theorem
For each pair $k, \ell \in \mathbb{N}_n$ with $k < \ell$,
the sign of the transposition $\tau_{k\ell}$ is $\mathrm{sgn}(\tau_{k\ell}) = -1$.

Proof.
A previous lemma established that the transposition $\tau_{k\ell}$
equals both $\pi^{\ell-1\searrow k} \circ \pi^{k\nearrow \ell}$ and $\pi^{k+1\nearrow \ell} \circ \pi^{\ell\searrow k}$.

Both of these are the compositions
of $2(\ell - k) - 1$ adjacency transpositions.

Because $2(\ell - k) - 1$ is an odd integer,
it follows that $\mathrm{sgn}(\tau_{k\ell}) = (-1)^{2(\ell-k)-1} = -1$. $\qquad\square$

Products of General Transpositions

Theorem

For each $r \in \mathbb{N}$, the sign $\mathrm{sgn}\left(\prod_{i \in \mathbb{N}_r} \tau_{k_i, \ell_i}\right)$
of the product of r general transpositions is $(-1)^r$.

Proof.

For each $i \in \mathbb{N}_r$, the transposition τ_{k_i, ℓ_i} can be expressed
as the product of $2(\ell_i - k_i) - 1$ adjacency transpositions.

So the product $\prod_{i \in \mathbb{N}_r} \tau_{k_i, \ell_i}$ can be expanded into the product
of $\sum_{i \in \mathbb{N}_r}[2(\ell_i - k_i) - 1]$ adjacency transpositions, whose sign is

$$(-1)^{\sum_{i \in \mathbb{N}_r}[2(\ell_i - k_i) - 1]} = (-1)^{-r} = (-1)^r \qquad \square$$

Products of Permutations

Theorem
For all pairs of permutations $\rho, \pi \in \Pi_n$
one has $\mathrm{sgn}(\rho \circ \pi) = \mathrm{sgn}(\rho)\,\mathrm{sgn}(\pi)$.

Proof.
Every permutation can be expressed
as a product of finitely many transpositions.

So suppose that the permutations ρ and π can be expressed
as the finite products $\rho = \prod_{i=1}^{k} \tau_{r_i, r_{i+1}}$ and $\pi = \prod_{j=1}^{\ell} \tau_{s_j, s_{j+1}}$
of k and ℓ transpositions respectively.

It follows that $\mathrm{sgn}(\rho) = (-1)^k$ and $\mathrm{sgn}(\pi) = (-1)^\ell$.

But $\rho \circ \pi = \left(\prod_{i=1}^{k} \tau_{r_i, r_{i+1}} \right) \circ \left(\prod_{j=1}^{\ell} \tau_{s_j, s_{j+1}} \right)$
is the product of $k + \ell$ transpositions,
so $\mathrm{sgn}(\rho \circ \pi) = (-1)^{k+\ell} = (-1)^k \cdot (-1)^\ell = \mathrm{sgn}(\rho)\,\mathrm{sgn}(\pi)$. $\qquad \square$

The Sign of an Inverse Permutation

Corollary

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

Proof.

The identity permutation ι has 0 reversals,
so $\text{sgn}(\iota) = (-1)^0 = +1$.

Then, because $\iota = \pi \circ \pi^{-1}$, the product rule implies that

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

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

A Final Remark

Example

The same permutation $\tau_{1,2} \in \Pi_2$ is both:

1. the product $\tau_{1,2}$ of one transposition;
2. the product $\tau_{1,2} \circ \tau_{1,2} \circ \tau_{1,2}$ of three transpositions.

Remark

The sign of any permutation $\pi \in \Pi_n$
equals the parity of the number of permutations
in any set whose product is π.

It is the parity of this number that is unique;
not the set of permutations,
nor even the number of permutations in this set.