Lecture Notes: Matrix Algebra
Part B: Introduction to Matrices

Peter J. Hammond

revised 2023 September 15th;
typeset from `matrixAlgB23.tex`

Outline

Matrices as Rectangular Arrays

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

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

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
$\mathbf{a}_j = (a_{ij})_{i=1}^m \in \mathbb{R}^m$ for $j = 1, 2, \dots, n$;

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

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

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

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

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 $\mathbf{A}$ and $\mathbf{B}$
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 $\mathbf{A}$;
- the jth column vector $\mathbf{b}_j$ of the second matrix $\mathbf{B}$.

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

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

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

The data is to be fitted to the linear regression model

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

whose scalar constants b_i $(i = 1, 2, \ldots, k)$ *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 $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ for suitable column vectors **y, b, 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 $\mathbf{y} = \mathbf{b}\mathbf{X} + \mathbf{e}$?
4. How can the equation $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ 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 $\mathbf{A}$ and $\mathbf{B}$ commute just in case $\mathbf{AB} = \mathbf{BA}$.

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

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

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

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

But then $\mathbf{AB}$ is $n \times n$, whereas $\mathbf{BA}$ is $m \times m$.

Evidently $\mathbf{AB} \neq \mathbf{BA}$ unless $m = n$.

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

To summarize, we must be in the special case
where $\mathbf{A}$ and $\mathbf{B}$ are two square matrices of the same dimension.

Matrix Multiplication Does Not Commute II

Even if both $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices,
implying that both $\mathbf{AB}$ and $\mathbf{BA}$ 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 $\mathbf{A}$ 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., $i \neq j \implies a_{ij} = 0$.

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 $\mathbf{D}$ 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 $\mathbb{N}_n \times \mathbb{N}_n \ni (i, j) \mapsto \delta_{ij} \in \{0, 1\}$.

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. $\square$

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 $\mathbf{A}$ *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 $(\mathbf{AB})\mathbf{X} = \mathbf{X}(\mathbf{AB}) = \mathbf{I}$. $\square$

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

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

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

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

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

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

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

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

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

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

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

Outline

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

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

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

$$\begin{aligned} & \mathbf{diag}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(k)})\ \mathbf{diag}(\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(k)}) \\ \textit{and} \quad & \mathbf{diag}(\mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(k)})\ \mathbf{diag}(\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(k)}) \end{aligned}$$

both exist, and that they equal

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

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 (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

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 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, 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. $\square$

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. $\square$

Definition

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

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

Adjacency Transpositions and Their Products, II

Exercise

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

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

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

Reduction to the Product of Adjacency Transpositions

Lemma

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

Proof.

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

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

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

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

Outline

The Inversions of a Permutation

Definition

1. Let $\mathbb{N}_{n,2} = \{S \subseteq \mathbb{N}_n \mid \#S = 2\}$ denote the set of all (unordered) pair subsets of $\mathbb{N}_n$.
2. Obviously, if $\{i, j\} \in \mathbb{N}_{n,2}$, then $i \neq j$.
3. Given any pair $\{i, j\} \in \mathbb{N}_{n,2}$,
   define $i \vee j := \max\{i, j\}$ and $i \wedge j := \min\{i, j\}$.
   For all $\{i, j\} \in \mathbb{N}_{n,2}$, because $i \neq j$, one has $i \vee j > i \wedge j$.
4. Given any permutation $\pi \in \Pi_n$,
   the pair $\{i, j\} \in \mathbb{N}_{n,2}$ is an inversion of π
   just in case π "reorders" $\{i, j\}$
   in the sense that $\pi(i \vee j) < \pi(i \wedge j)$.
5. Denote the set of inversions of π by

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

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

The Sign of a Permutation

Definition

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

The Sign of an Adjacency Transposition

Theorem

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

Proof.

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

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

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

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

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

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

A Multi-Part Exercise

Exercise

Show that:

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

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

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

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

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

Outline

Double Products

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

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

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

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

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

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

Preliminary Example and Definition

Example

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

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

Then one has

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

Definition

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

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

Basic Lemma

Lemma

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

Proof.

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

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

$\square$

The Product Rule for Signs of Permutations

Theorem

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

Proof.

The basic lemma implies that

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

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

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

The Sign of any Inverse Permutation

Corollary

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

Proof.

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

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

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