Chapter 2

Matrices

Section 2.1

Introduction to Matrices

Matrices (Definition 2.1.1)

A matrix (plural matrices) is a rectangular array of numbers.

The numbers in the array are called entries in the matrix.

The size of a matrix is given by $m \times n$ where m is the number of rows and n is the number of columns.

The (i, j)-entry of a matrix is the number which is in the i^{th} row and j^{th} column of a matrix.

Examples (Example 2.1.2)

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & -1 \end{bmatrix}$$ is a 3×2 matrix.

(1, 2)-entry

(3, 1)-entry

$$\begin{bmatrix} 4 \end{bmatrix}$$ is a 1×1 matrix.

A 1×1 matrix is usually treated as a number in computation.

$$\begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$$ is a 1×3 matrix.

$$\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$ is a 3×1 matrix.

$$\begin{bmatrix} \sqrt{2} & 3.1 & -2 \\ 3 & \frac{1}{2} & 0 \\ 0 & \pi & 0 \end{bmatrix}$$ is a 3×3 matrix.

Column and row matrices (Definition 2.1.3 & Example 2.1.4)

A column matrix (or a column vector) is a matrix with only one column.

A row matrix (or a row vector) is a matrix with only one row.

$\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ is a column matrix.

$\begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$ is a row matrix.

$\begin{bmatrix} 4 \end{bmatrix}$ is both a column and row matrix.

Notation of matrices (Notation 2.1.5)

In general, an $m \times n$ matrix can be written as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

or simply $\boldsymbol{A} = (a_{ij})_{m \times n}$ where a_{ij} is the (i, j)-entry of $\boldsymbol{A}$.

If the size of the matrix is already known, we may just write $\boldsymbol{A} = (a_{ij})$.

Examples (Example 2.1.6)

Let $\boldsymbol{A} = (a_{ij})_{2\times 3}$ where $a_{ij} = i + j$.

Then $\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$.

Let $\boldsymbol{B} = (b_{ij})_{3\times 2}$ where $b_{ij} = \begin{cases} 1 & \text{if } i + j \text{ is even} \\ -1 & \text{if } i + j \text{ is odd.} \end{cases}$

Then $\boldsymbol{B} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$.

Square matrices (Definition 2.1.7.1 & Example 2.1.8.1)

A matrix is called a square matrix if it has the same number of rows and columns.

In particular, an $n \times n$ matrix is called a square matrix of order n.

The following are some examples of square matrices:

$$\begin{bmatrix} 4 \end{bmatrix} \qquad \begin{bmatrix} 0 & 4 \\ 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ -1 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 6 & 2 \\ 0 & 3 & 9 & -1 \\ 1 & 0 & 0 & 0 \\ -2 & 3 & 0 & 1 \end{bmatrix}$$

Diagonal entries (Definition 2.1.7.2)

Given a square matrix $\boldsymbol{A} = (a_{ij})$ of order n, the diagonal of $\boldsymbol{A}$ is the sequence of entries $a_{11}, a_{22}, \ldots, a_{nn}$.

Each entry a_{ii} is called a diagonal entry while a_{ij}, with $i \neq j$, is called a non-diagonal entry.

diagonal entries

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

non-diagonal entries

Diagonal matrices (Definition 2.1.7.2 & Example 2.1.8.2)

A square matrix is called a diagonal matrix if all its non-diagonal entries are zero,

i.e. $\boldsymbol{A} = (a_{ij})_{n \times n}$ is a diagonal matrix

$\Leftrightarrow$ $a_{ij} = 0$ whenever $i \neq j$.

The following are some examples of diagonal matrices:

$$\begin{pmatrix} 4 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Scalar matrices (Definition 2.1.7.3 & Example 2.1.8.3)

A diagonal matrix is called a scalar matrix if all its diagonal entries are the same,

i.e. $\boldsymbol{A} = (a_{ij})_{n\times n}$ is a scalar matrix

$$\Leftrightarrow \quad a_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ c & \text{if } i = j \end{cases} \quad \text{for a constant } c.$$

The following are some examples of scalar matrices:

$$\begin{bmatrix} 4 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Identity matrices (Definition 2.1.7.4 & Example 2.1.8.4)

A diagonal matrix is called an identity matrix if all its diagonal entries are 1.

We use $\boldsymbol{I}_n$ to denote the identity matrix of order n.

Sometimes we write $\boldsymbol{I}$ instead of $\boldsymbol{I}_n$ when there is no danger of confusion.

The following are some examples of identity matrices:

$$\boldsymbol{I}_1 = \begin{bmatrix} 1 \end{bmatrix} \quad \boldsymbol{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Zero matrices (Definition 2.1.7.5 & Example 2.1.8.5)

A matrix with all entries equal to zero is called a zero matrix.

We use $\mathbf{0}_{m\times n}$ to denote the zero matrix of size $m \times n$.

Sometimes we write $\mathbf{0}$ instead of $\mathbf{0}_{m\times n}$ when there is no danger of confusion.

The following are some examples of zero matrices:

$$\mathbf{0}_{1\times 1} = \begin{bmatrix} 0 \end{bmatrix} \qquad \mathbf{0}_{2\times 4} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{0}_{4\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Symmetric matrices (Definition 2.1.7.6 & Example 2.1.7.6)

A square matrix (a_{ij}) is called symmetric if

$$a_{ij} = a_{ji} \quad \text{for all } i, \ j.$$

The following are some examples of symmetric matrices:

$$\begin{bmatrix} 4 \end{bmatrix} \qquad \begin{bmatrix} 0 & 4 \\ 4 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 6 & -2 \\ 1 & 3 & 0 & -1 \\ 6 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 \end{bmatrix}$$

Triangular matrices (Definition 2.1.7.7 & Example 2.1.7.7)

A square matrix (a_{ij}) is called upper triangular if

$$a_{ij} = 0 \quad \text{whenever} \quad i > j.$$

A square matrix (a_{ij}) is called lower triangular if

$$a_{ij} = 0 \quad \text{whenever} \quad i < j.$$

Both upper and lower triangular matrices are called triangular matrices.

The following are some examples of triangular matrices:

lower triangular

$$\begin{bmatrix} 4 \end{bmatrix} \qquad \begin{bmatrix} 0 & 4 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 \end{bmatrix}$$

upper triangular

Section 2.2

Matrix Operations

Equal (Definition 2.2.1 & Example 2.2.2)

Two matrices are said to be equal if they have the same size and their corresponding entries are equal.

Given $\boldsymbol{A} = (a_{ij})_{m \times n}$ and $\boldsymbol{B} = (b_{ij})_{p \times q}$,

$\boldsymbol{A}$ is equal to $\boldsymbol{B}$ if $m = p$, $n = q$ and $a_{ij} = b_{ij}$ for all i, j.

Let $\boldsymbol{A} = \begin{bmatrix} 1 & x \\ 2 & 4 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ and $\boldsymbol{C} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 4 & 0 \end{bmatrix}$.

Then $\boldsymbol{A} = \boldsymbol{B}$ if and only if $x = -1$;

$\boldsymbol{B} \neq \boldsymbol{C}$

and $\boldsymbol{A} \neq \boldsymbol{C}$ for any values of x.

Matrix addition (Definition 2.2.3.1)

Given $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{m\times n}$,

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m\times n},$$

i.e. $$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

An example (Example 2.2.4.1)

Let $\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 5 & 6 \end{bmatrix}$, and $\boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & -1 \end{bmatrix}$.

Then $\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} 2+1 & 3+2 & 4+3 \\ 4+(-1) & 5+(-1) & 6+(-1) \end{bmatrix}$

$= \begin{bmatrix} 3 & 5 & 7 \\ 3 & 4 & 5 \end{bmatrix}$.

Matrix subtraction (Definition 2.2.3.2)

Given $\boldsymbol{A} = (a_{ij})_{m \times n}$ and $\boldsymbol{B} = (b_{ij})_{m \times n}$,

$$\boldsymbol{A} - \boldsymbol{B} = (a_{ij} - b_{ij})_{m \times n}.$$

i.e. $$\boldsymbol{A} - \boldsymbol{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2n} - b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \cdots & a_{mn} - b_{mn} \end{bmatrix}.$$

An example (Example 2.2.4.2)

Let $\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 5 & 6 \end{bmatrix}$, and $\boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & -1 \end{bmatrix}$.

Then $\boldsymbol{A} - \boldsymbol{B} = \begin{bmatrix} 2-1 & 3-2 & 4-3 \\ 4-(-1) & 5-(-1) & 6-(-1) \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \end{bmatrix}$.

Scalar multiplication (Definition 2.2.3.3)

Given $\boldsymbol{A} = (a_{ij})_{m\times n}$ and a real constant c,

$$c\boldsymbol{A} = (ca_{ij})_{m\times n}$$

where the constant c is usually called a scalar.

i.e. $$c\boldsymbol{A} = c\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}.$$

An example (Example 2.2.4.3)

Let $\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 5 & 6 \end{bmatrix}$.

Then $4\boldsymbol{A} = \begin{bmatrix} 4\cdot 2 & 4\cdot 3 & 4\cdot 4 \\ 4\cdot 4 & 4\cdot 5 & 4\cdot 6 \end{bmatrix}$

$$= \begin{bmatrix} 8 & 12 & 16 \\ 16 & 20 & 24 \end{bmatrix}.$$

Some remarks (Remark 2.2.5)

1. Given a matrix $\boldsymbol{A}$, we normally use $-\boldsymbol{A}$ to denote the matrix $(-1)\boldsymbol{A}$.

2. The matrix subtraction can be defined using the matrix addition:

   Given two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of the same size, $\boldsymbol{A} - \boldsymbol{B}$ is defined to be the matrix $\boldsymbol{A} + (-\boldsymbol{B})$.

Some basic properties (Theorem 2.2.6)

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ be matrices of the same size and c, d are scalars.

1. Commutative law for matrix addition: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$.
2. Associative law for matrix addition:
   $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$.
3. $c(\boldsymbol{A} + \boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}$.
4. $(c + d)\boldsymbol{A} = c\boldsymbol{A} + d\boldsymbol{A}$.
5. $(cd)\boldsymbol{A} = c(d\boldsymbol{A}) = d(c\boldsymbol{A})$.
6. $\boldsymbol{A} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{A} = \boldsymbol{A}$.
7. $\boldsymbol{A} - \boldsymbol{A} = \boldsymbol{0}$.
8. $0\boldsymbol{A} = \boldsymbol{0}$.

$\boldsymbol{0}$ is the zero matrix of the same size as $\boldsymbol{A}$.

0 is the number zero.

Proof of $A + (B + C) = (A + B) + C$ (Theorem 2.2.6.2)

To prove $A + (B + C) = (A + B) + C$:

> Recall that two matrices are equal if
> (i) they have the same size and
> (ii) their corresponding entries are equal.

(i) Since A, B, C are matrices of the same size, by the definition of the matrix addition,

$$A + (B + C) \quad \text{and} \quad (A + B) + C$$

have the same size.

Proof of $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$ (Theorem 2.2.6.2)

(ii) Let $\boldsymbol{A} = (a_{ij})_{m \times n}$, $\boldsymbol{B} = (b_{ij})_{m \times n}$ and $\boldsymbol{C} = (c_{ij})_{m \times n}$.

For any i, j,

the (i, j)-entry of $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$

= a_{ij} + [the (i, j)-entry of $\boldsymbol{B} + \boldsymbol{C}$]

= $a_{ij} + [\, b_{ij} + c_{ij} \,]$

= $[\, a_{ij} + b_{ij} \,] + c_{ij}$

by the associative law for real number addition

= [the (i, j)-entry of $\boldsymbol{A} + \boldsymbol{B}$] + c_{ij}

= the (i, j)-entry of $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$.

By (i) and (ii), $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$.

The associative law (Remark 2.2.7)

Let $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, …, $\boldsymbol{A_k}$ be matrices of the same size.

By the Associative Law for Matrix Addition, we can write

$$\boldsymbol{A_1} + \boldsymbol{A_2} + \cdots + \boldsymbol{A_k}$$

to represent the sum of the matrices without using any parentheses to indicate the order of the matrix addition.

Matrix multiplication (Definition 2.2.8)

Given $\boldsymbol{A} = (a_{ij})_{m \times p}$ and $\boldsymbol{B} = (b_{ij})_{p \times n}$,

The product $\boldsymbol{AB}$ is defined to be an $m \times n$ matrix whose (i, j)-entry is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}\,.$$

for $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$.

(Remark 2.2.10.1)

We can only compute the product $\boldsymbol{AB}$ when the number of columns of $\boldsymbol{A}$ is equal to the number of rows of $\boldsymbol{B}$.

Examples (Example 2.2.9.1)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 + 3 \cdot (-1) & \\ & \end{bmatrix}$$

$$= \begin{bmatrix} 2 & \\ & \end{bmatrix}.$$

Examples (Example 2.2.9.1)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 + 3 \cdot (-1) & 1 \cdot 1 + 2 \cdot 3 + 3 \cdot (-2) \\ & \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ & \end{bmatrix}.$$

Examples (Example 2.2.9.1)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 + 3 \cdot (-1) & 1 \cdot 1 + 2 \cdot 3 + 3 \cdot (-2) \\ 4 \cdot 1 + 5 \cdot 2 + 6 \cdot (-1) & \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ 8 & \end{bmatrix}.$$

Examples (Example 2.2.9.1)

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 1\cdot 1 + 2\cdot 2 + 3\cdot(-1) & 1\cdot 1 + 2\cdot 3 + 3\cdot(-2) \\ 4\cdot 1 + 5\cdot 2 + 6\cdot(-1) & 4\cdot 1 + 5\cdot 3 + 6\cdot(-2) \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}.$$

Examples (Example 2.2.9.2)

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \cdot 1 + 1 \cdot 4 & 1 \cdot 2 + 1 \cdot 5 & 1 \cdot 3 + 1 \cdot 6 \\ 2 \cdot 1 + 3 \cdot 4 & 2 \cdot 2 + 3 \cdot 5 & 2 \cdot 3 + 3 \cdot 6 \\ (-1) \cdot 1 + (-2) \cdot 4 & (-1) \cdot 2 + (-2) \cdot 5 & (-1) \cdot 3 + (-2) \cdot 6 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 7 & 9 \\ 14 & 19 & 24 \\ -9 & -12 & -15 \end{bmatrix}$$

Multiplication isn't commutative (Remark 2.2.10.2)

The matrix multiplication is not commutative,

i.e. in general, $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are two different matrices even the products exists.

For example, let $\boldsymbol{A} = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$.

Then $\boldsymbol{AB} = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix}$ and $\boldsymbol{BA} = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$.

Hence $\boldsymbol{AB} \neq \boldsymbol{BA}$.

Multiplication isn't commutative (Remark 2.2.10.3)

It would be ambiguous to say "the multiplication of a matrix A to another matrix B" since it could mean AB or BA.

To distinguish the two, we refer to

AB as the pre-multiplication of A to B

and BA as the post-multiplication of A to B.

When the product is a zero matrix (Remark 2.2.10.4)

When $\boldsymbol{AB} = \mathbf{0}$, it is not necessary that $\boldsymbol{A} = \mathbf{0}$ or $\boldsymbol{B} = \mathbf{0}$.

For example, let $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.

We have $\boldsymbol{A} \neq \mathbf{0}$, $\boldsymbol{B} \neq \mathbf{0}$ and

$$\boldsymbol{AB} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}.$$

Some basic properties (Theorem 2.2.11)

1. Associative law for matrix multiplication:

   If $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are $m \times p$, $p \times q$ and $q \times n$ matrices respectively, then

   $$\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}.$$

   So we can write the product as $\boldsymbol{ABC}$ without using parentheses.

2. Distribution laws for matrix addition and multiplication:

   If $\boldsymbol{A}$, $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are $m \times p$, $p \times n$ and $p \times n$ matrices respectively, then

   $$\boldsymbol{A}(\boldsymbol{B}_1 + \boldsymbol{B}_2) = \boldsymbol{AB}_1 + \boldsymbol{AB}_2.$$

   If $\boldsymbol{A}$, $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are $p \times n$, $m \times p$ and $m \times p$ matrices respectively, then

   $$(\boldsymbol{C}_1 + \boldsymbol{C}_2)\boldsymbol{A} = \boldsymbol{C}_1\boldsymbol{A} + \boldsymbol{C}_2\boldsymbol{A}.$$

Some basic properties (Theorem 2.2.11)

3. If $\boldsymbol{A}$, $\boldsymbol{B}$ are $m \times p$, $p \times n$ matrices, respectively, and c is a scalar, then

$$c(\boldsymbol{AB}) = (c\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(c\boldsymbol{B}).$$

4. If $\boldsymbol{A}$ is an $m \times n$ matrix, then

$$\boldsymbol{A}\boldsymbol{0}_{n \times q} = \boldsymbol{0}_{m \times q},$$

$$\boldsymbol{0}_{p \times m}\boldsymbol{A} = \boldsymbol{0}_{p \times n}$$

and $$\boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}.$$

Proof of $A(B_1 + B_2) = AB_1 + AB_2$ (Theorem 2.2.11.2)

To prove $A(B_1 + B_2) = AB_1 + AB_2$:

> Recall that two matrices are equal if
> (i) they have the same size and
> (ii) their corresponding entries are equal.

(i) Since the size of A is $m \times p$ and the size of $B_1 + B_2$ is $p \times n$, the size of $A(B_1 + B_2)$ is $m \times n$.

On the other hand, the sizes of both AB_1 and AB_2 are $m \times n$ and hence the size of $AB_1 + AB_2$ is $m \times n$.

Thus $A(B_1 + B_2)$ and $AB_1 + AB_2$ have the same size.

Proof of $A(B_1 + B_2) = AB_1 + AB_2$ (Theorem 2.2.11.2)

(ii) Let $A = (a_{ij})_{m \times p}$, $B_1 = (b_{ij})_{p \times n}$ and $B_2 = (b_{ij}')_{p \times n}$.

For any i, j,

the (i, j)-entry of $A(B_1 + B_2)$

$$= a_{i1}\,[\text{ the } (1, j)\text{-entry of } B_1 + B_2\,] + a_{i2}\,[\text{ the } (2, j)\text{-entry of } B_1 + B_2\,] + \cdots + a_{ip}\,[\text{ the } (p, j)\text{-entry of } B_1 + B_2\,]$$

$$= a_{i1}\,[\,b_{1j} + b_{1j}'\,] + a_{i2}\,[\,b_{2j} + b_{2j}'\,] + \cdots + a_{ip}\,[\,b_{pj} + b_{pj}'\,]$$

$$= a_{i1}b_{1j} + a_{i1}b_{1j}' + a_{i2}b_{2j} + a_{i2}b_{2j}' + \cdots + a_{ip}b_{pj} + a_{ip}b_{pj}'.$$

by the distributive law for real numbers

Proof of $A(B_1 + B_2) = AB_1 + AB_2$ (Theorem 2.2.11.2)

On the other hand,

the (i, j)-entry of $AB_1 + AB_2$

$= [$ the (i, j)-entry of AB_1 $] + [$ the (i, j)-entry of AB_2 $]$

$= [\, a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} \,]$

$+ [\, a_{i1}b_{1j}' + a_{i2}b_{2j}' + \cdots + a_{ip}b_{pj}' \,]$

$= a_{i1}b_{1j} + a_{i1}b_{1j}' + a_{i2}b_{2j} + a_{i2}b_{2j}' + \cdots + a_{ip}b_{pj} + a_{ip}b_{pj}'$

$=$ the (i, j)-entry of $A(B_1 + B_2)$.

from the previous slide

By (i) and (ii), $A(B_1 + B_2) = AB_1 + AB_2$.

Powers of square matrices (Definition 2.2.12 & Example 2.2.13)

Let $\boldsymbol{A}$ be a square matrix and n a nonnegative integer. We define $\boldsymbol{A}^n$ as follows:

the identity matrix

$$\boldsymbol{A}^n = \begin{cases} \boldsymbol{I} & \text{if } n = 0 \\ \underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}_{n \text{ times}} & \text{if } n \geq 1. \end{cases}$$

For example, let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$.

Then $\boldsymbol{A}^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$.

Powers of square matrices (Remark 2.2.14)

1. Let $\boldsymbol{A}$ be a square matrix and n, m nonnegative integers. Then $\boldsymbol{A}^m\boldsymbol{A}^n = \boldsymbol{A}^{m+n}$.
2. Since matrix multiplication is not commutative, in general, for two square matrix $\boldsymbol{A}$ and $\boldsymbol{B}$ of the same size, $(\boldsymbol{AB})^2$ and $\boldsymbol{A}^2\boldsymbol{B}^2$ may be different.

   For example, let $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

   Then $(\boldsymbol{AB})^2 = (\boldsymbol{AB})(\boldsymbol{AB}) = \boldsymbol{ABAB} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

   and $\boldsymbol{A}^2\boldsymbol{B}^2 = (\boldsymbol{AA})(\boldsymbol{BB}) = \boldsymbol{AABB} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$.

Some useful notation (Notation 2.2.15)

Given $\boldsymbol{A} = (a_{ij})_{m \times p}$, we can write

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a_1} \\ \boldsymbol{a_2} \\ \vdots \\ \boldsymbol{a_m} \end{bmatrix} \quad \text{where } \boldsymbol{a_i} = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ip} \end{bmatrix}$$

is the i^{th} row of $\boldsymbol{A}$.

Given $\boldsymbol{B} = (b_{ij})_{p \times n}$, we can write $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{b_1} & \boldsymbol{b_2} & \cdots & \boldsymbol{b_n} \end{bmatrix}$

where $\boldsymbol{b_j} = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{bmatrix}$ is the j^{th} column of $\boldsymbol{B}$.

Some useful notation (Notation 2.2.15)

Then

$$\boldsymbol{AB} = \begin{bmatrix} \boldsymbol{a}_1\boldsymbol{b}_1 & \boldsymbol{a}_1\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1\boldsymbol{b}_n \\ \boldsymbol{a}_2\boldsymbol{b}_1 & \boldsymbol{a}_2\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2\boldsymbol{b}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_m\boldsymbol{b}_1 & \boldsymbol{a}_m\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_m\boldsymbol{b}_n \end{bmatrix}$$

where $\boldsymbol{a}_i\boldsymbol{b}_j = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ip} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{bmatrix}$

The i^{th} row of $\boldsymbol{A}$ pre-multiply to the j^{th} column of $\boldsymbol{B}$.

$$= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}$$

which is the (i, j)-entry of $\boldsymbol{AB}$.

Some useful notation (Notation 2.2.15)

Also we can write

$$\boldsymbol{AB} = \boldsymbol{A}\begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{Ab}_1 & \boldsymbol{Ab}_2 & \cdots & \boldsymbol{Ab}_n \end{bmatrix},$$

where $\boldsymbol{Ab}_j$ is the j^{th} column of $\boldsymbol{AB}$;

or

$$\boldsymbol{AB} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_m \end{bmatrix} \boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}_1\boldsymbol{B} \\ \boldsymbol{a}_2\boldsymbol{B} \\ \vdots \\ \boldsymbol{a}_m\boldsymbol{B} \end{bmatrix}$$

where $\boldsymbol{a}_i\boldsymbol{B}$ is the i^{th} row of $\boldsymbol{AB}$.

The j^{th} column of $\boldsymbol{AB}$ is

$$\begin{bmatrix} a_{11}b_{1j} + a_{12}b_{2j} + \cdots + a_{1p}b_{pj} \\ a_{21}b_{1j} + a_{22}b_{2j} + \cdots + a_{2p}b_{pj} \\ \vdots \\ a_{m1}b_{1j} + a_{m2}b_{2j} + \cdots + a_{mp}b_{pj} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{bmatrix} = \boldsymbol{Ab_j}.$$

The i^{th} row of $\boldsymbol{AB}$ is

$$\begin{bmatrix} a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{ip}b_{p1} & a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{ip}b_{p2} & \cdots & a_{i1}b_{1n} + a_{i2}b_{2n} + \cdots + a_{ip}b_{pn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{bmatrix} = \boldsymbol{a_i B}.$$

An example (Example 2.2.16)

Let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ with $\boldsymbol{a}_1 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ and $\boldsymbol{a}_2 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$;

$\boldsymbol{B} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -1 & -2 \end{bmatrix}$ with $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $\boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$.

Then $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix}$ and

$$\boldsymbol{AB} = \begin{bmatrix} \boldsymbol{a}_1\boldsymbol{b}_1 & \boldsymbol{a}_1\boldsymbol{b}_2 \\ \boldsymbol{a}_2\boldsymbol{b}_1 & \boldsymbol{a}_2\boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{Ab}_1 & \boldsymbol{Ab}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1\boldsymbol{B} \\ \boldsymbol{a}_2\boldsymbol{B} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}.$$

Linear systems (Remark 2.2.17)

Consider the system of linear equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m . \end{cases}$$

Using matrix multiplication, the system can be rewritten as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Linear systems (Remark 2.2.17)

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Let $\boldsymbol{A} = (a_{ij})_{m \times n}$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$.

$\boldsymbol{b}$ is called the constant matrix.

We represent the system by the matrix equation $\boldsymbol{Ax} = \boldsymbol{b}$.

$\boldsymbol{A}$ is called the coefficient matrix.

$\boldsymbol{x}$ is called the variable matrix.

Linear systems (Remark 2.2.17)

Write $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{c_1} & \boldsymbol{c_2} & \cdots & \boldsymbol{c_n} \end{bmatrix}$ where $\boldsymbol{c_j}$ is the j^{th} column of $\boldsymbol{A}$.

The linear system can also be represented by

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

i.e. $x_1\boldsymbol{c_1} + x_2\boldsymbol{c_2} + \cdots + x_n\boldsymbol{c_n} = \boldsymbol{b}$ or $\sum_{j=1}^{n} x_j\,\boldsymbol{c_j} = \boldsymbol{b}$.

An example (Example 2.2.18)

The system of linear equations

$$\begin{cases} 4x + 5y + 6z = 5 \\ x - y = 2 \\ y - z = 3 \end{cases}$$

can be written as $\begin{bmatrix} 4 & 5 & 6 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}$

or $x \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} + z \begin{bmatrix} 6 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}.$

Solutions to linear systems

We follow the notation of Remark 2.2.16, an $n \times 1$ matrix $\boldsymbol{u}$ is said to be a solution to the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ if the equation is satisfied when we substitute $\boldsymbol{x} = \boldsymbol{u}$ into the equation, i.e. $\boldsymbol{Au} = \boldsymbol{b}$.

For example, in Example 2.2.17,

$$\boldsymbol{u} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} \text{ is a solution to } \begin{bmatrix} 4 & 5 & 6 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}$$

because $\begin{bmatrix} 4 & 5 & 6 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}$.

Transposes (Definition 2.2.19 & Example 2.2.20 & Remark 2.2.21.1)

Let $\boldsymbol{A} = (a_{ij})$ be an $m \times n$ matrix.

The transpose of $\boldsymbol{A}$, denoted by $\boldsymbol{A}^{\mathrm{T}}$ (or $\boldsymbol{A}^{\mathrm{t}}$), is an $n \times m$ matrix whose (i, j)-entry is a_{ji}.

Let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$. Then $\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$.

Note that the rows of $\boldsymbol{A}$ are the columns of $\boldsymbol{A}^{\mathrm{T}}$ and vice versa.

Symmetric matrices (Remark 2.2.21.2)

Recall that a square matrix (a_{ij}) is called symmetric if

$$a_{ij} = a_{ji} \quad \text{for all } i,\ j.$$

Thus a square matrix $\boldsymbol{A}$ is symmetric if and only if $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$.

Some basic properties (Theorem 2.2.22)

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

1. $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$.
2. If $\boldsymbol{B}$ be an $m \times n$ matrix, then $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$.
3. If c is a scalar, then $(c\boldsymbol{A})^{\mathrm{T}} = c\boldsymbol{A}^{\mathrm{T}}$.
4. If $\boldsymbol{B}$ be an $n \times p$ matrix, then $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

Proof of $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ (Theorem 2.2.22.4)

To prove $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$:

> Recall that two matrices are equal if
> (i) they have the same size and
> (ii) their corresponding entries are equal.

(i) Since the size of $\boldsymbol{AB}$ is $m \times p$, the size of $(\boldsymbol{AB})^{\mathrm{T}}$ is $p \times m$.

On the other hand, the sizes of $\boldsymbol{B}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}$ are $p \times n$ and $n \times m$, respectively, and hence the size of $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ is $p \times m$.

Thus $(\boldsymbol{AB})^{\mathrm{T}}$ and $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ have the same size.

Proof of $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ (Theorem 2.2.11.2)

(ii) Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{n\times p}$.

For any i, j,

the (i, j)-entry of $\boldsymbol{AB} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

Thus

$$
\begin{aligned}
&\text{the } (i, j)\text{-entry of } (\boldsymbol{AB})^{\mathrm{T}} \\
&= \text{the } (j, i)\text{-entry of } \boldsymbol{AB} \\
&= a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni}.
\end{aligned}
$$

Proof of $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ (Theorem 2.2.11.2)

On the other hand, write $\boldsymbol{A}^{\mathrm{T}} = (a_{ij}')_{n\times m}$ and $\boldsymbol{B}^{\mathrm{T}} = (b_{ij}')_{p\times n}$ where $a_{ij}' = a_{ji}$ and $b_{ij}' = b_{ji}$.

For any i, j,

$$
\begin{aligned}
&\text{the } (i, j)\text{-entry of } \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} \\
&= b_{i1}'a_{1j}' + b_{i2}'a_{2j}' + \cdots + b_{in}'a_{nj}' \\
&= b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ni}a_{jn} \\
&= a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni} \\
&= \text{the } (i, j)\text{-entry of } (\boldsymbol{AB})^{\mathrm{T}}.
\end{aligned}
$$

from the previous slide

By (i) and (ii), $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

Section 2.3

Inverses of Square Matrices

Inverses (Discussion 2.3.1)

Let a and b be two real number such that $a \neq 0$. Then the solution of the equation

$$ax = b$$

is $x = \frac{b}{a} = a^{-1}b$.

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrices. It is much harder to solve the matrix equation

$$\boldsymbol{AX} = \boldsymbol{B}$$

because we do not have "division" for matrices.

However, for some square matrices, we can find their "inverses" which have the similar property as a^{-1} in the computation of the solution of $ax = b$ above.

Inverses of square matrices (Definition 2.3.2)

Let $\boldsymbol{A}$ be a square matrix of order n.

Then $\boldsymbol{A}$ is said to be invertible if there exists a square matrix $\boldsymbol{B}$ of order n such that

$$\boldsymbol{AB} = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{BA} = \boldsymbol{I}.$$

The matrix $\boldsymbol{B}$ here is called an inverse of $\boldsymbol{A}$.

A square matrix is called singular if it has no inverse.

Examples (Example 2.3.3.1)

Let $\boldsymbol{A} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$.

Then

$$\boldsymbol{AB} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I},$$

$$\boldsymbol{BA} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}.$$

So $\boldsymbol{A}$ is invertible and $\boldsymbol{B}$ is an inverse of $\boldsymbol{A}$.

Examples (Example 2.3.3.2)

$$\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \boldsymbol{X} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \boldsymbol{X} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\Rightarrow \boldsymbol{IX} = \begin{bmatrix} 12 \\ 4 \end{bmatrix}$$

$$\Rightarrow \boldsymbol{X} = \begin{bmatrix} 12 \\ 4 \end{bmatrix}$$

By Theorem 2.2.11.3, $\boldsymbol{IX} = \boldsymbol{X}$.

Examples (Example 2.3.3.3)

In next section, we shall learnt a systematic method to check whether a square matrix is invertible.

Show that $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ is singular.

(Proof by contradiction)

Assume the contrary, i.e. A has an inverse $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

i.e. $BA = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

On the other hand, $BA = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ c+d & 0 \end{bmatrix}$.

It is impossible.

Hence our assumption is wrong, i.e. A is singular.

Cancellation laws (Remark 2.3.4)

1. Cancellation laws for matrix multiplication:

   Let $\boldsymbol{A}$ be an invertible $m \times m$ matrix.

   (a) If $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are $m \times n$ matrices such that $\boldsymbol{AB}_1 = \boldsymbol{AB}_2$, then $\boldsymbol{B}_1 = \boldsymbol{B}_2$.

   (b) If $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are $n \times m$ matrices such that $\boldsymbol{C}_1\boldsymbol{A} = \boldsymbol{C}_2\boldsymbol{A}$, then $\boldsymbol{C}_1 = \boldsymbol{C}_2$.

2. If $\boldsymbol{A}$ is not invertible, the cancellation laws may not hold.

   For example, let $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, $\boldsymbol{B}_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and

   $\boldsymbol{B}_2 = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$. Then $\boldsymbol{AB}_1 = \boldsymbol{AB}_2$ but $\boldsymbol{B}_1 \neq \boldsymbol{B}_2$.

Uniqueness of inverses (Theorem 2.3.5)

If $\boldsymbol{B}$ and $\boldsymbol{C}$ are inverses of a square matrix $\boldsymbol{A}$, then $\boldsymbol{B} = \boldsymbol{C}$.

Proof: By the definition of inverses (Definition 2.3.2),

$$\boxed{\boldsymbol{AB} = \boldsymbol{I},} \quad \boldsymbol{BA} = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{AC} = \boldsymbol{I}, \quad \boxed{\boldsymbol{CA} = \boldsymbol{I}.}$$

So $\boldsymbol{AB} = \boldsymbol{I} \Rightarrow \boldsymbol{CAB} = \boldsymbol{CI} \Rightarrow \boldsymbol{IB} = \boldsymbol{C} \Rightarrow \boldsymbol{B} = \boldsymbol{C}$.

(Notation 2.3.6)

Given an invertible matrix $\boldsymbol{A}$, since there is only one inverse of $\boldsymbol{A}$, we use the symbol $\boldsymbol{A}^{-1}$ to denote this unique inverse of $\boldsymbol{A}$.

2 x 2 matrices (Example 2.3.8)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then $\boldsymbol{A}$ is invertible and

$$\boldsymbol{A}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix}.$$

Proof: Let $\boldsymbol{B} = \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix}$.

(Remark 2.3.7)

To show that $\boldsymbol{A}$ is invertible and $\boldsymbol{B}$ is the inverse of $\boldsymbol{A}$, by Definition 2.3.2, we need to check $\boldsymbol{AB} = \boldsymbol{I}$ and $\boldsymbol{BA} = \boldsymbol{I}$.

(By Theorem 2.4.12 in the next section, we shall see that we only need to check one of the two conditions: $\boldsymbol{AB} = \boldsymbol{I}$ or $\boldsymbol{BA} = \boldsymbol{I}$.)

2 x 2 matrices (Example 2.3.8)

$$\boldsymbol{AB} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I},$$

$$\boldsymbol{BA} = \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}.$$

Hence $\boldsymbol{A}$ is invertible and $\boldsymbol{A}^{-1} = \boldsymbol{B}$.

Some basic properties (Theorem 2.3.9)

Let $\boldsymbol{A}$, $\boldsymbol{B}$ be two invertible matrices and c a nonzero scalar.

1. $c\boldsymbol{A}$ is invertible and $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$.
2. $\boldsymbol{A}^{\mathrm{T}}$ is invertible and $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$.
3. $\boldsymbol{A}^{-1}$ is invertible and $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.
4. $\boldsymbol{AB}$ is invertible and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

(Remark 2.3.10)

By Part 4, if $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, ..., $\boldsymbol{A}_k$ are invertible matrices, then $\boldsymbol{A}_1\boldsymbol{A}_2 \cdots \boldsymbol{A}_k$ is invertible and

$$(\boldsymbol{A}_1\boldsymbol{A}_2 \cdots \boldsymbol{A}_k)^{-1} = \boldsymbol{A}_k^{-1} \cdots \boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}.$$

A^{T} is invertible (Theorem 2.3.9.2)

To prove that A^{T} is invertible and $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$:

By Remark 2.3.7, we only need to show that

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = I \text{ and } (A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = I.$$

(By Theorem 2.4.12 in the next section, we shall see that we only need to check one of the two conditions:

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = I \text{ or } (A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = I.)$$

by Theorem 2.2.22.4

$$A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = (A^{-1}A)^{\mathrm{T}} = I^{\mathrm{T}} = I$$

and similarly $(A^{-1})^{\mathrm{T}}A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = I^{\mathrm{T}} = I$.

So A^{T} is invertible and $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$.

Powers of square matrices (Definition 2.3.11 & Example 2.3.12)

Let $\boldsymbol{A}$ be an invertible matrix and n a positive integer. We define $\boldsymbol{A}^{-n}$ as follows:

$$\boldsymbol{A}^{-n} = (\boldsymbol{A}^{-1})^{n} = \underbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1} \cdots \boldsymbol{A}^{-1}}_{n \text{ times}}.$$

For example, let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$. Then $\boldsymbol{A}^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$.

So $\boldsymbol{A}^{-3} = (\boldsymbol{A}^{-1})^{3} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}.$$

Some basic properties (Remark 2.3.13)

Let $\boldsymbol{A}$ be an invertible matrix.

1. $\boldsymbol{A}^r\boldsymbol{A}^s = \boldsymbol{A}^{r+s}$ for any integers r and s.
2. $\boldsymbol{A}^n$ is invertible and $(\boldsymbol{A}^n)^{-1} = (\boldsymbol{A}^{-1})^n$.

Section 2.4

Elementary Matrices

Some useful terms from Chapter 1 (Definition 2.4.1)

In Chapter 1, the following concepts are defined for augmented matrices:

- elementary row operations,
- row equivalent matrices,
- row-echelon forms,
- reduced row-echelon forms,
- Gaussian Elimination,
- Gauss-Jordan Elimination.

From now on, these terms will also be used for matrices.

Elementary row operations (Discussion 2.4.2.1)

Multiply a row by a constant:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} \xrightarrow{2R_2} \boldsymbol{B} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 4 & -2 & 6 & 12 \\ 1 & 4 & 4 & 0 \end{bmatrix}$$

Let $\boldsymbol{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Then $\boldsymbol{E}_1\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 4 & -2 & 6 & 12 \\ 1 & 4 & 4 & 0 \end{bmatrix} = \boldsymbol{B}$.

Elementary row operations (Discussion 2.4.2.1)

In general, let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{E}$ an $m \times m$ matrix such that

$$\boldsymbol{E} = \begin{array}{c} \begin{array}{ccccccc} & & & i^{\text{th}}\text{ column} & & & \\ & & & \downarrow & & & \end{array} \\ \left[\begin{array}{ccc|c|ccc} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ \hline & & & k & & & \\ \hline & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{array}\right] \end{array} . \leftarrow i^{\text{th}}\text{ row}$$

Then $\boldsymbol{A} \xrightarrow{kR_i} \boldsymbol{EA}$.

Elementary row operations (Discussion 2.4.2.1)

If $k \neq 0$, then $\boldsymbol{E}$ is invertible
and

$$\begin{array}{cc} & \quad\quad\quad\quad i^{\text{th}} \text{ column} \\ & \quad\quad\quad\quad\downarrow \\ \boldsymbol{E}^{-1} = \left[\begin{array}{ccc|c|ccc} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ \hline & & & \frac{1}{k} & & & \\ \hline & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{array}\right]. & \leftarrow i^{\text{th}} \text{ row} \end{array}$$

Note that $\boldsymbol{A} \xrightarrow{\frac{1}{k}R_i} \boldsymbol{E}^{-1}\boldsymbol{A}$.

Elementary row operations (Discussion 2.4.2.2)

Interchange two rows:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \boldsymbol{B} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 1 & 4 & 4 & 0 \\ 2 & -1 & 3 & 6 \end{bmatrix}$$

Let $\boldsymbol{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Then $\boldsymbol{E}_2\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 1 & 4 & 4 & 0 \\ 2 & -1 & 3 & 6 \end{bmatrix} = \boldsymbol{B}$.

Elementary row operations (Discussion 2.4.2.2)

In general, let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{E}$ an $m \times m$ matrix such that

$$\boldsymbol{E} = \begin{array}{r c} & \begin{matrix} \qquad & i^{\text{th}}\text{ column} & \qquad & j^{\text{th}}\text{ column} & \qquad \end{matrix} \\ \begin{matrix} \\ \\ \\ i^{\text{th}}\text{ row} \\ \\ \\ \\ j^{\text{th}}\text{ row} \\ \\ \\ \end{matrix} & \begin{bmatrix} \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & & & 0 \\ & 0 & & 1 & \\ & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} & & \\ & 1 & & 0 & \\ 0 & & & & \begin{matrix}1&&\\&\ddots&\\&&1\end{matrix} \end{bmatrix} \end{array}.$$

Then $\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{EA}$.

$\boldsymbol{E}$ is invertible and $\boldsymbol{E}^{-1} = \boldsymbol{E}$.

Elementary row operations (Discussion 2.4.2.3)

Add a multiple of a row to another row:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} \xrightarrow{R_3 + 2R_1} \boldsymbol{B} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 3 & 4 & 8 & 6 \end{bmatrix}$$

Let $\boldsymbol{E_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$.

Then $\boldsymbol{E_3 A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 3 & 4 & 8 & 6 \end{bmatrix} = \boldsymbol{B}$.

Elementary row operations (Discussion 2.4.2.3)

In general, let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{E}$ an $m \times m$ matrix such that if $i < j$,

$$\boldsymbol{E} = \begin{array}{cc}
\begin{array}{ccccccccc} & & & i^{\text{th}}\text{ column} & & j^{\text{th}}\text{ column} & & & \end{array} & \\
\begin{bmatrix}
1 & & & & & & & & \\
 & \ddots & & & & & & 0 & \\
 & & 1 & & & & & & \\
 & & & 1 & & & & & \\
 & & & & \ddots & & & & \\
 & & & k & & 1 & & & \\
 & & & & & & 1 & & \\
 & 0 & & & & & & \ddots & \\
 & & & & & & & & 1
\end{bmatrix} & \begin{array}{c} \\ \\ \\ \\ \\ \leftarrow j^{\text{th}}\text{ row} \\ \\ \\ \\ \end{array}
\end{array};$$

Elementary row operations (Discussion 2.4.2.3)

and if $i > j$,

$$
\begin{array}{rcl}
 & \begin{array}{ccccccccc} & & & j^{\text{th}}\text{ column} & & & i^{\text{th}}\text{ column} & & \\ & & & \downarrow & & & \downarrow & & \end{array} & \\
\boldsymbol{E} = & \begin{bmatrix}
1 & & & & & & & & \\
 & \ddots & & & & & & 0 & \\
 & & 1 & & & & & & \\
 & & & 1 & & & k & & \\
 & & & & \ddots & & & & \\
 & & & & & 1 & & & \\
 & & & & & & 1 & & \\
 & 0 & & & & & & \ddots & \\
 & & & & & & & & 1
\end{bmatrix} & \begin{array}{l} \\ \\ \\ \longleftarrow j^{\text{th}}\text{ row} \\ \\ \\ \\ \\ \\ \end{array}
\end{array}
.
$$

Then $\boldsymbol{A} \xrightarrow{R_j + kR_i} \boldsymbol{EA}$.

Elementary row operations (Discussion 2.4.2.3)

$\boldsymbol{E}$ is invertible

and if $i < j$,

$$
E^{-1} = \begin{array}{cl}
\begin{array}{ccccccccc} & & & i^{\text{th}}\text{ column} & & j^{\text{th}}\text{ column} & & & \\ & & & \downarrow & & \downarrow & & & \end{array} & \\
\left[\begin{array}{ccccccccc}
1 & & & & & & & & \\
 & \ddots & & & & & & 0 & \\
 & & 1 & & & & & & \\
 & & & 1 & & & & & \\
 & & & & \ddots & & & & \\
 & & & -k & & 1 & & & \\
 & & & & & & 1 & & \\
 & 0 & & & & & & \ddots & \\
 & & & & & & & & 1
\end{array}\right] & \begin{array}{l} \\ \\ \\ \\ \\ \leftarrow j^{\text{th}}\text{ row} \\ \\ \\ \\ \end{array}
\end{array};
$$

Elementary row operations (Discussion 2.4.2.3)

and if $i > j$,

$$
\begin{array}{rcl}
 & \begin{array}{ccccccccc} & & & j^{\text{th}}\text{ column} & & i^{\text{th}}\text{ column} & & & \\ & & & \downarrow & & \downarrow & & & \end{array} & \\
E^{-1} = & \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & 1 & & -k & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow j^{\text{th}}\text{ row} \\ \\ \\ \\ \\ \\ \end{array} .
\end{array}
$$

Note that $A \xrightarrow{R_j - kR_i} E^{-1}A$.

Elementary matrices (Definition 2.4.3 & Remark 2.4.4)

A square matrix is called an elementary matrix if it can be obtained from an identity matrix by performing a single elementary row operation.

The three types of matrices $\boldsymbol{E}$ described in the previous slides (Discussion 2.4.2) are elementary matrices and every elementary matrix belongs to one of these three types.

All elementary matrices are invertible and their inverses are also elementary matrices.

An example (Example 2.4.5)

$$\boldsymbol{A} = \begin{bmatrix} 0 & 4 & 2 \\ -2 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix} \xrightarrow[\boldsymbol{E}_1]{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 2 \end{bmatrix}$$

$$\boldsymbol{E}_1\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 4 & 2 \\ -2 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 2 \end{bmatrix}$$

$$\boldsymbol{E}_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

An example (Example 2.4.5)

$$\boldsymbol{A} \xrightarrow[\boldsymbol{E_1}]{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 2 \end{bmatrix} \xrightarrow[\boldsymbol{E_2}]{R_2 + 2R_1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 4 & 2 \end{bmatrix}$$

$$\boldsymbol{E_2 E_1 A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 4 & 2 \end{bmatrix}$$

$$\boldsymbol{E_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \boldsymbol{E_2} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

An example (Example 2.4.5)

$$A \xrightarrow[\boldsymbol{E_1}]{R_1 \leftrightarrow R_3} \xrightarrow[\boldsymbol{E_2}]{R_2 + 2R_1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 4 & 2 \end{bmatrix} \xrightarrow[\boldsymbol{E_3}]{R_3 - 4R_2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix}$$

$$\boldsymbol{E_3 E_2 E_1 A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix}$$

$$\boldsymbol{E_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \boldsymbol{E_2} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{E_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}$$

An example (Example 2.4.5)

$$\boldsymbol{A} \xrightarrow[\boldsymbol{E_1}]{R_1 \leftrightarrow R_3} \xrightarrow[\boldsymbol{E_2}]{R_2 + 2R_1} \xrightarrow[\boldsymbol{E_3}]{R_3 - 4R_2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \xrightarrow[\boldsymbol{E_4}]{-\frac{1}{2}R_3} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \boldsymbol{B}$$

$$\boldsymbol{E_4 E_3 E_2 E_1 A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{E_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \boldsymbol{E_2} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{E_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \quad \boldsymbol{E_4} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

An example (Example 2.4.5)

$$A = \begin{bmatrix} 0 & 4 & 2 \\ -2 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We have seen that

$$E_4E_3E_2E_1A = B.$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \quad E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

$$E_1^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

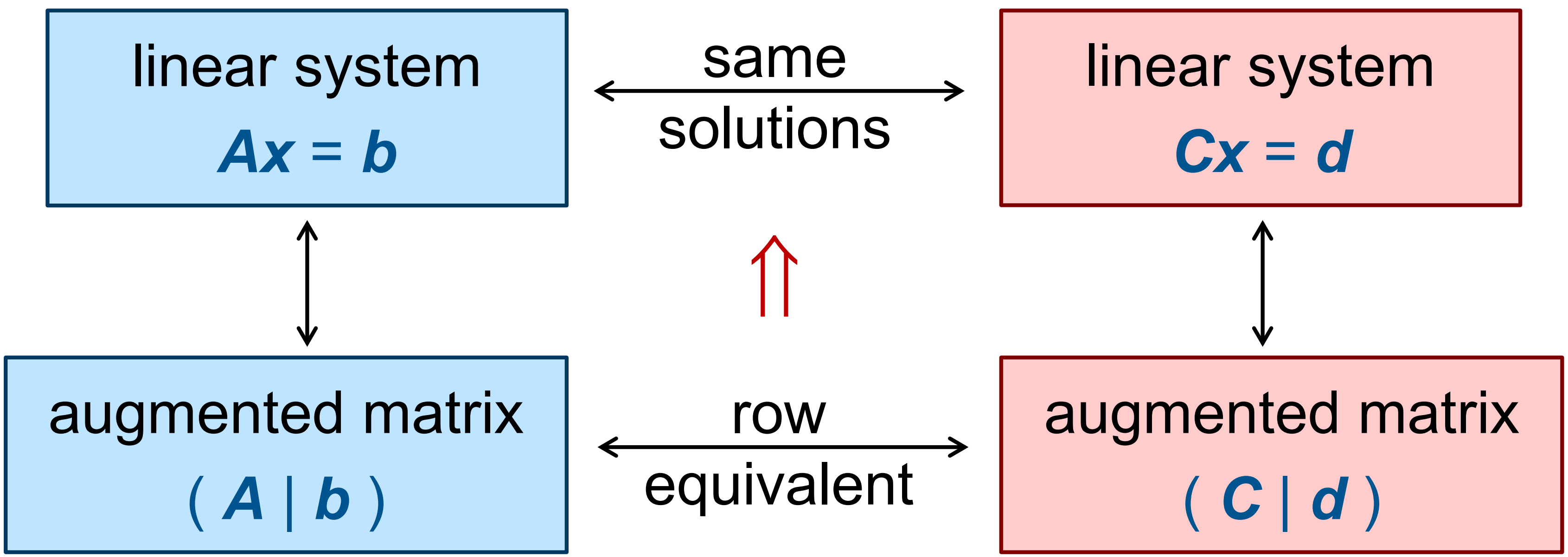
$$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \quad E_4^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Then

$$\begin{aligned} A &= (E_4E_3E_2E_1)^{-1}B \\ &= E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}B. \end{aligned}$$

System of linear equations (Remark 2.4.6)

Recall Theorem 1.2.7: If augmented matrices of two systems of linear equations are row equivalent, then the two systems have the same set of solutions.



System of linear equations (Remark 2.4.6)

Proof: Since $(\boldsymbol{A} \mid \boldsymbol{b})$ and $(\boldsymbol{C} \mid \boldsymbol{d})$ are row equivalent, one can be obtained by the other by a series of elementary row operations.

Thus (by Discussion 2.4.2) there exists elementary matrices $\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_k$ such that

$$\begin{aligned}
& \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1(\boldsymbol{A} \mid \boldsymbol{b}) = (\boldsymbol{C} \mid \boldsymbol{d}) \\
\Rightarrow \quad & (\boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A} \mid \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{b}) = (\boldsymbol{C} \mid \boldsymbol{d}) \\
\Rightarrow \quad & \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A} = \boldsymbol{C} \quad \text{and} \quad \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{b} = \boldsymbol{d}.
\end{aligned}$$

Note that $\boldsymbol{A} = (\boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1)^{-1}\boldsymbol{C} = \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1} \cdots \boldsymbol{E}_k^{-1}\boldsymbol{C}$

and $\boldsymbol{b} = (\boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1)^{-1}\boldsymbol{d} = \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1} \cdots \boldsymbol{E}_k^{-1}\boldsymbol{d}$.

System of linear equations (Remark 2.4.6)

If $\boldsymbol{x} = \boldsymbol{u}$ is a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, then

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{b} \;\Rightarrow\; \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}\boldsymbol{u} = \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{b}$$
$$\Rightarrow\; \boldsymbol{C}\boldsymbol{u} = \boldsymbol{d}$$

and hence $\boldsymbol{x} = \boldsymbol{u}$ is a solution to $\boldsymbol{C}\boldsymbol{x} = \boldsymbol{d}$.

If $\boldsymbol{x} = \boldsymbol{v}$ is a solution to $\boldsymbol{C}\boldsymbol{x} = \boldsymbol{d}$, then

$$\boldsymbol{C}\boldsymbol{v} = \boldsymbol{d} \;\Rightarrow\; \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1} \cdots \boldsymbol{E}_k^{-1}\boldsymbol{C}\boldsymbol{v} = \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1} \cdots \boldsymbol{E}_k^{-1}\boldsymbol{d}$$
$$\Rightarrow\; \boldsymbol{A}\boldsymbol{v} = \boldsymbol{b}$$

and hence $\boldsymbol{x} = \boldsymbol{v}$ is a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

So we have shown that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{C}\boldsymbol{x} = \boldsymbol{d}$ has the same set of solutions.

Invertible matrices (Theorem 2.4.7)

Let $\boldsymbol{A}$ be a square matrix.

The following statements are equivalent:

1. $\boldsymbol{A}$ is invertible.
2. The linear system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.
3. The reduced row-echelon form of $\boldsymbol{A}$ is an identity matrix.
4. $\boldsymbol{A}$ can be expressed as a product of elementary matrices.

Invertible matrices (Theorem 2.4.7)

$1 \Rightarrow 2$: If $\boldsymbol{A}$ is invertible, then

$$\begin{aligned} \boldsymbol{Ax} = \mathbf{0} \quad &\Rightarrow \quad \boldsymbol{A}^{-1}\boldsymbol{Ax} = \boldsymbol{A}^{-1}\mathbf{0} \\ &\Rightarrow \quad \boldsymbol{Ix} = \mathbf{0} \\ &\Rightarrow \quad \boldsymbol{x} = \mathbf{0}. \end{aligned}$$

and hence the system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.

$2 \Rightarrow 3$: Suppose the system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.

The augmented matrix of the system is $(\ \boldsymbol{A} \mid \mathbf{0}\ )$.

Invertible matrices (Theorem 2.4.7)

Since $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution, every column of its row echelon form is a pivot column (see Remark 1.4.8.2).

$$(\, \boldsymbol{A} \mid \mathbf{0} \,) \xrightarrow[\text{Elimination}]{\text{Gaussian}} \left[\begin{array}{cccc|c} \otimes & * & \cdots & * & 0 \\ 0 & \otimes & \cdots & * & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & \otimes & 0 \end{array}\right]$$

Suppose the size of $\boldsymbol{A}$ is $n \times n$.

the number of nonzero rows
= the number of leading entries
= the number of pivot columns
= n

Hence there is no zero row.

$$\xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{array}\right] = (\, \boldsymbol{I} \mid \mathbf{0} \,)$$

Thus the reduced row-echelon form of $\boldsymbol{A}$ is an identity matrix.

Invertible matrices (Theorem 2.4.7)

$3 \Rightarrow 4$: Since the reduced row-echelon form of $\boldsymbol{A}$ is an identity matrix $\boldsymbol{I}$, there exists elementary matrices $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_k$ such that

$$\boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1 \boldsymbol{A} = \boldsymbol{I}.$$

Then

$$\begin{aligned} \boldsymbol{A} &= (\boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1)^{-1} \boldsymbol{I} \\ &= (\boldsymbol{E}_k \cdots \boldsymbol{E}_2 \boldsymbol{E}_1)^{-1} = \boldsymbol{E}_1^{-1} \boldsymbol{E}_2^{-1} \cdots \boldsymbol{E}_k^{-1} \end{aligned}$$

where $\boldsymbol{E}_1^{-1}$, $\boldsymbol{E}_2^{-1}$, ..., $\boldsymbol{E}_k^{-1}$ are also elementary matrices.

$4 \Rightarrow 1$: Suppose $\boldsymbol{A}$ is a product of elementary matrices.

Since elementary matrices are invertible, $\boldsymbol{A}$ is invertible (see Remark 2.3.10).

Finding inverses (Discussion 2.4.8)

Let A be an invertible matrix of order n.

There exists elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2 E_1 A = I \;\Rightarrow\; E_k \cdots E_2 E_1 A A^{-1} = I A^{-1}$$
$$\Rightarrow\; E_k \cdots E_2 E_1 I = A^{-1}$$
$$\Rightarrow\; E_k \cdots E_2 E_1 = A^{-1}.$$

Construct an $n \times 2n$ matrix $(\, A \mid I \,)$.

Then $E_k \cdots E_2 E_1 (\, A \mid I \,) = (\, E_k \cdots E_2 E_1 A \mid E_k \cdots E_2 E_1 I \,)$
$= (\, I \mid A^{-1} \,)$.

Thus

$$(\, A \mid I \,) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} (\, I \mid A^{-1} \,).$$

An example (Example 2.4.9)

Let $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

So $\boldsymbol{A}^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$.

To check whether invertible (Remark 2.4.10)

Let $\boldsymbol{A}$ be a square matrix.

Recall that if $\boldsymbol{A}$ is invertible, then

$$(\boldsymbol{A} \mid \boldsymbol{0}) \xrightarrow[\text{Elimination}]{\text{Gaussian}} \left[\begin{array}{cccc|c} \otimes & * & \cdots & * & 0 \\ 0 & \otimes & * & * & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & \otimes & 0 \end{array}\right]. \quad \longleftarrow \text{no zero row}$$

Hence if a row-echelon form of $\boldsymbol{A}$ has at least one zero row, $\boldsymbol{A}$ is singular (i.e. not invertible).

Examples (Example 2.4.11)

1. $$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 4 \\ 3 & 6 & 3 \end{bmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

   So $\boldsymbol{A}$ is singular.

2. Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

   We already know that if $ad - bc \neq 0$, then $\boldsymbol{A}$ is invertible (see Example 2.3.8).

   Actually, $\boldsymbol{A}$ is invertible if and only if $ad - bc \neq 0$.

   (By using Remark 2.4.10, we can show that if $\boldsymbol{A}$ is invertible, then $ad - bc \neq 0$. Read our textbook for the detailed proof.)

Invertible matrices (Theorem 2.4.12)

Let A and B be a square matrix of the same size.

If $AB = I$, then

(i) A is invertible,

(ii) B is invertible,

(iii) $A^{-1} = B$,

(iv) $B^{-1} = A$,

(v) $BA = I$.

Invertible matrices (Theorem 2.4.12)

Proof: Consider the linear system $\boldsymbol{Bx} = \mathbf{0}$:

$$\boldsymbol{Bx} = \mathbf{0} \;\Rightarrow\; \boldsymbol{ABx} = \boldsymbol{A}\mathbf{0} \;\Rightarrow\; \boldsymbol{Ix} = \mathbf{0} \;\Rightarrow\; \boldsymbol{x} = \mathbf{0}.$$

As the linear system $\boldsymbol{Bx} = \mathbf{0}$ has only the trivial solution, (by Theorem 2.4.7) $\boldsymbol{B}$ is invertible.

Since $\boldsymbol{B}$ is invertible,

$$\boldsymbol{AB} = \boldsymbol{I} \;\Rightarrow\; \boldsymbol{ABB}^{-1} = \boldsymbol{IB}^{-1} \;\Rightarrow\; \boldsymbol{AI} = \boldsymbol{B}^{-1} \;\Rightarrow\; \boldsymbol{A} = \boldsymbol{B}^{-1}.$$

Then (by Theorem 2.3.9.3) $\boldsymbol{A}$ is invertible

and $$\boldsymbol{A}^{-1} = (\boldsymbol{B}^{-1})^{-1} = \boldsymbol{B}.$$

Finally, $\boldsymbol{BA} = \boldsymbol{BB}^{-1} = \boldsymbol{I}$.

An example (Example 2.4.13)

Suppose $\boldsymbol{A}$ is a square matrix such that

$$\boldsymbol{A}^2 - 3\boldsymbol{A} - 6\boldsymbol{I} = \boldsymbol{0}.$$

Then $\boldsymbol{A}^2 - 3\boldsymbol{A} = 6\boldsymbol{I}$

$$\begin{aligned} &\Rightarrow \quad \boldsymbol{A}\boldsymbol{A} - \boldsymbol{A}(3\boldsymbol{I}) = 6\boldsymbol{I} \\ &\Rightarrow \quad \boldsymbol{A}(\boldsymbol{A} - 3\boldsymbol{I}) = 6\boldsymbol{I} \\ &\Rightarrow \quad \frac{1}{6}\left[\boldsymbol{A}(\boldsymbol{A} - 3\boldsymbol{I})\right] = \boldsymbol{I} \\ &\Rightarrow \quad \boldsymbol{A}\left[\frac{1}{6}(\boldsymbol{A} - 3\boldsymbol{I})\right] = \boldsymbol{I}. \end{aligned}$$

So $\boldsymbol{A}$ is invertible and $\boldsymbol{A}^{-1} = \frac{1}{6}(\boldsymbol{A} - 3\boldsymbol{I})$.

Singular matrices (Theorem 2.4.14)

Let A and B be square matrices of the same size.

If A is singular, then both AB and BA are singular.

Elementary column operations (Discussion 2.4.15)

Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{E}$ an $n \times n$ elementary matrix.

Then $\boldsymbol{A} \xrightarrow[\text{row operation}]{\text{an elementary}} \boldsymbol{EA}$.

How is $\boldsymbol{A}$ related to $\boldsymbol{AE}$?

Answer: $\boldsymbol{AE}$ can be obtained from $\boldsymbol{A}$ by doing an "elementary column operation".

(Read our textbook for more details.)

Section 2.5

Determinants

Invertible matrices (Discussion 2.5.1)

We know that a 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if and only if $ad - bc \neq 0$ (see Example 2.4.11.2).

We have a similar formula to determine whether a square matrix of higher order is invertible.

The formula involves a quantity called "determinant".

Determinants (Definition 2.5.2 & Notation 2.5.3)

Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix.

Let $\boldsymbol{M_{ij}}$ be an $(n-1) \times (n-1)$ matrix obtained from $\boldsymbol{A}$ by deleting the i^{th} row and the j^{th} column.

Then the determinant of $\boldsymbol{A}$ is defined to be

$$\det(\boldsymbol{A}) = \begin{cases} a_{11} & \text{if } n = 1 \\ a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & \text{if } n > 1 \end{cases}$$

where $A_{ij} = (-1)^{i+j} \det(\boldsymbol{M_{ij}})$ which is called the (i, j)-cofactor of $\boldsymbol{A}$.

Sometimes, we also write $\det(\boldsymbol{A})$ as $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$.

2×2 matrices (Example 2.5.4.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Then $\boldsymbol{M}_{11} = \begin{bmatrix} d \end{bmatrix}$,

$$A_{11} = (-1)^{1+1} \det(\boldsymbol{M}_{11}) = d$$

2×2 matrices (Example 2.5.4.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Then $\boldsymbol{M_{11}} = \begin{bmatrix} d \end{bmatrix}$, $\quad \boldsymbol{M_{12}} = \begin{bmatrix} c \end{bmatrix}$,

$$A_{11} = (-1)^{1+1}\det(\boldsymbol{M_{11}}) = d$$

and $\quad A_{12} = (-1)^{1+2}\det(\boldsymbol{M_{12}}) = -c.$

2×2 matrices (Example 2.5.4.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

$-$ $+$

Then $\boldsymbol{M}_{11} = [\,d\,]$, $\boldsymbol{M}_{12} = [\,c\,]$,

$$A_{11} = (-1)^{1+1}\det(\boldsymbol{M}_{11}) = d$$

and $A_{12} = (-1)^{1+2}\det(\boldsymbol{M}_{12}) = -c$.

So $\det(\boldsymbol{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} = ad - bc$.

Examples (Example 2.5.4.2)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = (-3)\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} - (-2)\begin{vmatrix} 4 & 1 \\ 0 & 4 \end{vmatrix} + 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix}$

Examples (Example 2.5.4.2)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = (-3)\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} - (-2)\begin{vmatrix} 4 & 1 \\ 0 & 4 \end{vmatrix} + 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix}$

Examples (Example 2.5.4.2)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = (-3)\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} - (-2)\begin{vmatrix} 4 & 1 \\ 0 & 4 \end{vmatrix} + 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix}$

Examples (Example 2.5.4.2)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = (-3)\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} - (-2)\begin{vmatrix} 4 & 1 \\ 0 & 4 \end{vmatrix} + 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix}$

$$= -3(3 \cdot 4 - 1 \cdot 2) + 2(4 \cdot 4 - 1 \cdot 0) + 4(4 \cdot 2 - 3 \cdot 0)$$

$$= 34.$$

Examples (Example 2.5.4.3)

Let $\boldsymbol{C} = \begin{bmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$.

Then $\det(\boldsymbol{C}) = 0 \begin{vmatrix} -3 & 3 & -2 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix} - (-1) \begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}$

$$+ 2 \begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix} - 0 \begin{vmatrix} 2 & -3 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{vmatrix}$$

Examples (Example 2.5.4.3)

Let $\boldsymbol{C} = \begin{bmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$.

Then $\det(\boldsymbol{C}) = 0 \begin{vmatrix} -3 & 3 & -2 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix} - (-1) \begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}$

$$+ 2 \begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix} - 0 \begin{vmatrix} 2 & -3 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{vmatrix}$$

Examples (Example 2.5.4.3)

Let $\boldsymbol{C} = \begin{bmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$.

Then $\det(\boldsymbol{C}) = 0 \begin{vmatrix} -3 & 3 & -2 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix} - (-1) \begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}$

$$+ 2 \begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix} - 0 \begin{vmatrix} 2 & -3 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{vmatrix}$$

Examples (Example 2.5.4.3)

Let $\boldsymbol{C} = \begin{bmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$.

Then $\det(\boldsymbol{C}) = 0\begin{vmatrix} -3 & 3 & -2 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix} - (-1)\begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}$

$$+ 2\begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix} - 0\begin{vmatrix} 2 & -3 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{vmatrix}$$

Examples (Example 2.5.4.3)

Let $\boldsymbol{C} = \begin{bmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$.

Then $\det(\boldsymbol{C}) = 0\begin{vmatrix} -3 & 3 & -2 \\ 2 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix} - (-1)\begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix}$

$$+\,2\begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix} - 0\begin{vmatrix} 2 & -3 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{vmatrix}$$

Examples (Example 2.5.4.3)

$$\det(\boldsymbol{C}) = \begin{vmatrix} 2 & 3 & -2 \\ 0 & 4 & 0 \\ 0 & 2 & -1 \end{vmatrix} + 2\begin{vmatrix} 2 & -3 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{vmatrix}$$

$$= \left(2\begin{vmatrix} 4 & 0 \\ 2 & -1 \end{vmatrix} - 3\begin{vmatrix} 0 & 0 \\ 0 & -1 \end{vmatrix} + (-2)\begin{vmatrix} 0 & 4 \\ 0 & 2 \end{vmatrix} \right)$$

$$+ 2\left(2\begin{vmatrix} 2 & 0 \\ 0 & -1 \end{vmatrix} - (-3)\begin{vmatrix} 0 & 0 \\ 0 & -1 \end{vmatrix} + (-2)\begin{vmatrix} 0 & 2 \\ 0 & 0 \end{vmatrix} \right)$$

$$= [2\cdot(-4) - 3\cdot 0 - 2\cdot 0] + 2[2\cdot(-2) + 3\cdot 0 - 2\cdot 0]$$

$$= -16.$$

3×3 matrices (Remark 2.5.5)

Let $\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

Then $\det(\boldsymbol{A}) = aei + bfg + cdh - ceg - afh - bdi$.

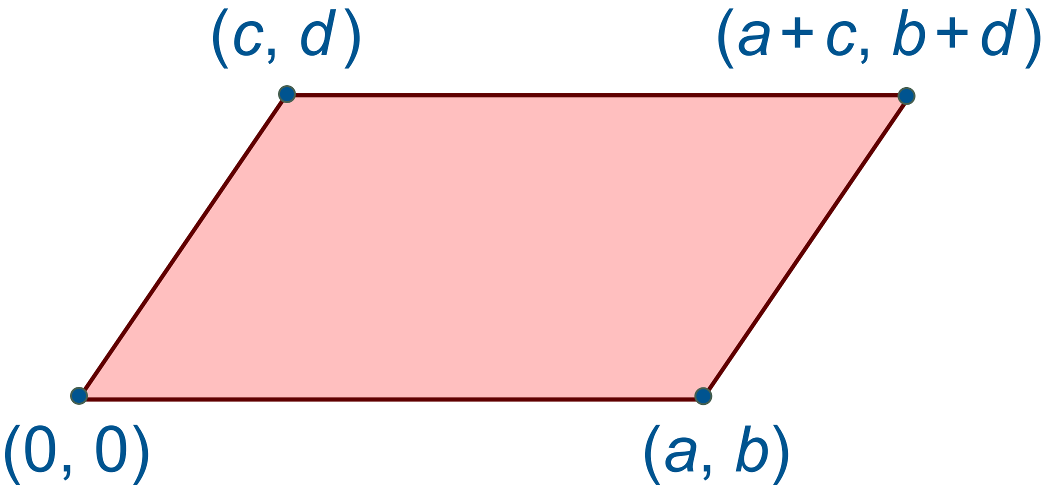
The formula in can be easily remembered by using diagram on the right:

$$\begin{array}{ccccc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}$$

$$- \quad - \quad - \qquad + \quad + \quad +$$

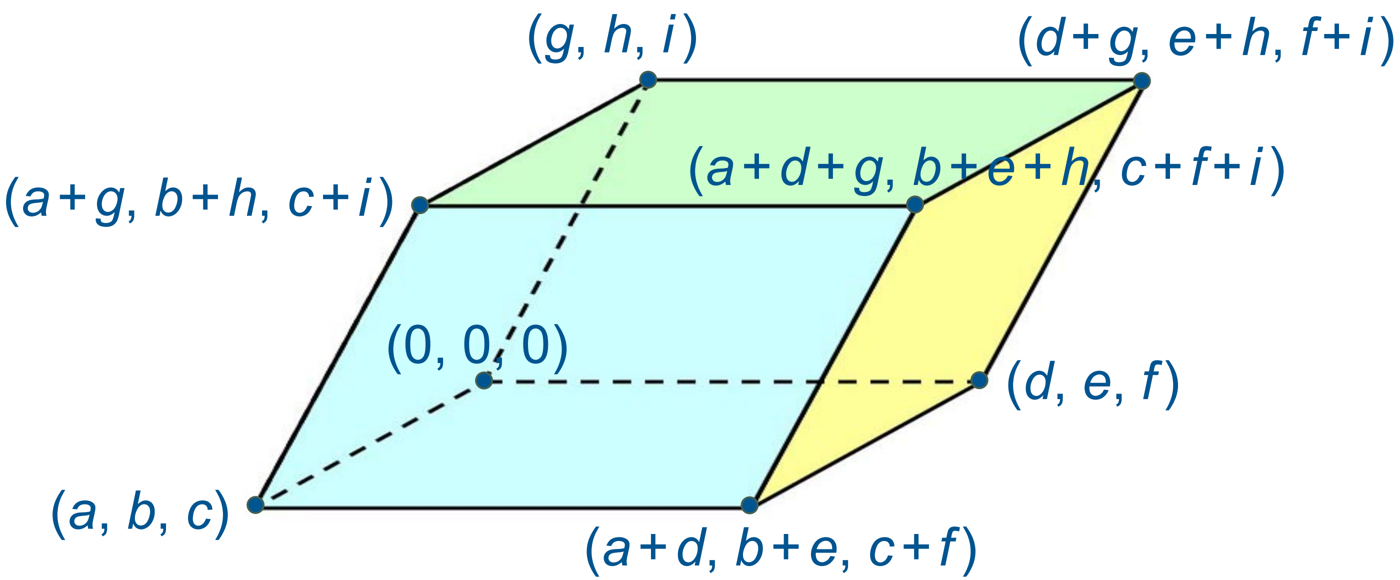
Warning: The method shown here cannot be generalized to higher order.

Geometrical interpretation



The area of the parallelogram is $ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Geometrical interpretation



The volume of the parallelepiped is

$$aei + bfg + cdh - ceg - afh - bdi = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}.$$

Cofactor expansions (Theorem 2.5.6)

Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix.

Recall that $\boldsymbol{M}_{ij}$ be an $(n-1) \times (n-1)$ matrix obtained from $\boldsymbol{A}$ by deleting the i^{th} row and the j^{th} column

and $A_{ij} = (-1)^{i+j} \det(\boldsymbol{M}_{ij})$.

Then for any $i = 1, 2, \ldots, n$ and $j = 1,2, \ldots, n$,

$$\begin{aligned}\det(\boldsymbol{A}) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}.\end{aligned}$$

cofactor expansion along the i^{th} row

cofactor expansion along the j^{th} column

(Since the proof involves some deeper knowledge of determinants, we use this result without proving it.)

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

$$= 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} + 4\begin{vmatrix} -3 & -2 \\ 4 & 3 \end{vmatrix} = 34.$$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4 \begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3 \begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1 \begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

$$= 4 \begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix} - 1 \begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} + 4 \begin{vmatrix} -3 & -2 \\ 4 & 3 \end{vmatrix} = 34.$$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

$$= 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} + 4\begin{vmatrix} -3 & -2 \\ 4 & 3 \end{vmatrix} = 34.$$

An example (Example 2.5.7)

Let $\boldsymbol{B} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$.

Then $\det(\boldsymbol{B}) = -4\begin{vmatrix} -2 & 4 \\ 2 & 4 \end{vmatrix} + 3\begin{vmatrix} -3 & 4 \\ 0 & 4 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} = 34$

$$= 4\begin{vmatrix} 4 & 3 \\ 0 & 2 \end{vmatrix} - 1\begin{vmatrix} -3 & -2 \\ 0 & 2 \end{vmatrix} + 4\begin{vmatrix} -3 & -2 \\ 4 & 3 \end{vmatrix} = 34.$$

Triangular matrices (Theorem 2.5.8)

If $\boldsymbol{A}$ is an $n \times n$ triangular matrix, i.e.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

then $\det(\boldsymbol{A}) = a_{11}a_{22} \cdots a_{nn}$.

Triangular matrices (Theorem 2.5.8)

Proof:

Let $$\boldsymbol{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

A proper proof using mathematics induction can be found in our textbook.

Then $$\det(\boldsymbol{A}) = a_{11}A_{11} + 0A_{12} + \cdots + 0A_{1n}$$

$$= a_{11} \begin{vmatrix} a_{22} & 0 & \cdots & 0 \\ a_{22} & a_{33} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n2} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

The result follows by repeating this process.

Examples (Example 2.5.9)

$$\det(\boldsymbol{I}) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} = 1 \cdot 1 \cdots 1 = 1.$$

$$\begin{vmatrix} -2 & 3.5 & 2 \\ 0 & 5 & -26 \\ 0 & 0 & 2 \end{vmatrix} = (-2) \cdot 5 \cdot 2 = -10.$$

$$\begin{vmatrix} -2 & 0 & 0 \\ 99 & 0 & 0 \\ 10 & -4.5 & 2 \end{vmatrix} = (-2) \cdot 0 \cdot 2 = 0.$$

Transposes (Theorem 2.5.10 & Example 2.5.11)

If $\boldsymbol{A}$ is a square matrix, then $\det(\boldsymbol{A}^T) = \det(\boldsymbol{A})$.

Let $\boldsymbol{C} = \begin{bmatrix} 0 & -1 & 2 & 0 \\ 2 & -3 & 3 & -2 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$. Then $\boldsymbol{C}^T = \begin{bmatrix} 0 & 2 & 0 & 0 \\ -1 & -3 & 2 & 0 \\ 2 & 3 & 4 & 2 \\ 0 & -2 & 0 & -1 \end{bmatrix}$.

Then $\det(\boldsymbol{C}^T) = -2 \begin{vmatrix} -1 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 0 & -1 \end{vmatrix}$

$$= -2[(-1)\cdot 4\cdot(-1) + 2\cdot 2\cdot 0 + 0\cdot 2\cdot 0 - 0\cdot 4\cdot 0 - (-1)\cdot 2\cdot 0 - 2\cdot 2\cdot(-1)]$$

$= -16 = \det(\boldsymbol{C})$ (see Example 2.5.4.3).

Identical rows or columns (Theorem 2.5.12 & Example 2.5.13)

1. The determinant of a square matrix with two identical rows is zero.
2. The determinant of a square matrix with two identical columns is zero.

The following matrices have zero determinant:

$$\begin{bmatrix} 4 & -2 \\ 4 & -2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 4 \\ -1 & 10 & 4 \\ 1 & 2 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & -3 & -3 & 9 \\ 2 & 4 & 4 & 0 \\ 0 & -2 & -2 & -1 \end{bmatrix}$$

Elementary row operations (Discussion 2.5.14.1)

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \xrightarrow{kR_2} \boldsymbol{B} = \begin{bmatrix} a & b & c \\ kd & ke & kf \\ g & h & i \end{bmatrix}$$

$$\begin{aligned} \det(\boldsymbol{B}) &= a(ke)i + b(kf)g + c(kd)h - c(ke)g - a(kf)h - b(kd)i \\ &= k(aei + bfg + cdh - ceg - afh - bdi) \\ &= k\det(\boldsymbol{A}). \end{aligned}$$

Elementary row operations (Discussion 2.5.14.1)

Let $\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Then $\boldsymbol{B} = \boldsymbol{EA}$.

Since $\det(\boldsymbol{E}) = 1 \cdot k \cdot 1 = k$,

$$\det(\boldsymbol{E}) \det(\boldsymbol{A}) = k \det(\boldsymbol{A}) = \det(\boldsymbol{B}) = \det(\boldsymbol{EA}).$$

Elementary row operations (Discussion 2.5.14.2)

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \boldsymbol{B} = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}$$

$$\begin{aligned} \det(\boldsymbol{B}) &= ahf + bid + cge - chd - aie - bgf \\ &= -(aei + bfg + cdh - ceg - afh - bdi) \\ &= -\det(\boldsymbol{A}). \end{aligned}$$

Elementary row operations (Discussion 2.5.14.2)

Let $\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Then $\boldsymbol{B} = \boldsymbol{EA}$.

Since $\det(\boldsymbol{E}) = 1 \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 0 \cdot 0 - 1 \cdot 1 = -1$,

$$\det(\boldsymbol{E}) \det(\boldsymbol{A}) = -\det(\boldsymbol{A}) = \det(\boldsymbol{B}) = \det(\boldsymbol{EA}).$$

Elementary row operations (Discussion 2.5.14.3)

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \xrightarrow{R_3 + kR_1} \boldsymbol{B} = \begin{bmatrix} a & b & c \\ d & e & f \\ g+ka & h+kb & i+kc \end{bmatrix}$$

$$\begin{aligned} \det(\boldsymbol{B}) &= ae(i+kc) + bf(g+ka) + cd(h+kb) \\ &\qquad - ce(g+ka) - af(h+kb) - bd(i+kc) \\ &= aei + bfg + cdh - ceg - afh - bdi \\ &\qquad + k(aec + bfa + cdb - cea - afb - bdc) \\ &= aei + bfg + cdh - ceg - afh - bdi \\ &= \det(\boldsymbol{A}). \end{aligned}$$

Elementary row operations (Discussion 2.5.14.3)

Let $\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$.

Then $\boldsymbol{B} = \boldsymbol{EA}$.

Since $\det(\boldsymbol{E}) = 1 \cdot 1 \cdot 1 = 1$,

$$\det(\boldsymbol{E}) \det(\boldsymbol{A}) = \det(\boldsymbol{A}) = \det(\boldsymbol{B}) = \det(\boldsymbol{EA}).$$

Elementary row operations (Theorem 2.5.15)

Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix.

$$\boldsymbol{A} \xrightarrow{kR_i} \boldsymbol{B_1} \qquad \det(\boldsymbol{B_1}) = k\det(\boldsymbol{A})$$

$$\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B_2} \qquad \det(\boldsymbol{B_2}) = -\det(\boldsymbol{A})$$

$$\boldsymbol{A} \xrightarrow{R_j + kR_i} \boldsymbol{B_3} \qquad \det(\boldsymbol{B_3}) = \det(\boldsymbol{A})$$

Furthermore, if $\boldsymbol{E}$ is an elementary matrix of the same size as $\boldsymbol{A}$, then $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$.

Proof of $\det(B_3) = \det(A)$ (Theorem 2.5.15)

To prove $\det(\boldsymbol{B_3}) = \det(\boldsymbol{A})$:

Since $\boldsymbol{B_3}$ is obtained from $\boldsymbol{A}$ by adding k times of the i^{th} row of $\boldsymbol{A}$ to j^{th} row,

$$\boldsymbol{B_3} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{j1}+ka_{i1} & a_{j2}+ka_{i2} & \cdots & a_{jn}+ka_{in} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ \cancel{a_{j1}} & \cancel{a_{j2}} & \cancel{\cdots} & \cancel{a_{jn}} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \boldsymbol{B_3} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ \cancel{a_{j1}+ka_{i1}} & \cancel{a_{j2}+ka_{i2}} & \cancel{\cdots} & \cancel{a_{jn}+ka_{in}} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

After deleting the j^{th} row, the matrices on both sides look exactly the same.

$$\text{The } (j, t)\text{-cofactor of } \boldsymbol{B_3} = (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1t-1} & a_{1t+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & \\ a_{j-1,1} & \cdots & a_{j-1,t-1} & a_{j-1,t+1} & \cdots & a_{j-1,n} \\ a_{j+1,1} & \cdots & a_{j+1,t-1} & a_{j+1,t+1} & \cdots & a_{j+1,n} \\ \vdots & & \vdots & \vdots & \vdots & \\ a_{m1} & \cdots & a_{mt-1} & a_{mt+1} & \cdots & a_{mn} \end{vmatrix}$$

$$= \text{ the } (j, t)\text{-cofactor of } \boldsymbol{A} = A_{jt}.$$

Proof of $\det(B_3) = \det(A)$ (Theorem 2.5.15)

$$\det(\boldsymbol{B_3}) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{j1}+ka_{i1} & a_{j2}+ka_{i2} & \cdots & a_{jn}+ka_{in} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}$$

cofactor expansion along the j^{th} row

$$= (a_{j1}+ka_{i1})A_{j1} + (a_{j2}+ka_{i2})A_{j2} + \cdots + (a_{jn}+ka_{in})A_{jn}$$

$$= a_{j1}A_{j1} + a_{j2}A_{j2} + \cdots + a_{jn}A_{jn} + k(a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn})$$

$$= \det(\boldsymbol{A}) + k(a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}).$$

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = ?$$

cofactor expansion along the j^{th} row

Compare with $\det(\boldsymbol{A}) = a_{j1}A_{j1} + a_{j2}A_{j2} + \cdots + a_{jn}A_{jn}$.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \boldsymbol{C} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The (j, t)-cofactor of $\boldsymbol{C}$ $= (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1t-1} & a_{1t+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{j-1,1} & \cdots & a_{j-1,t-1} & a_{j-1,t+1} & \cdots & a_{j-1,n} \\ a_{j+1,1} & \cdots & a_{j+1,t-1} & a_{j+1,t+1} & \cdots & a_{j+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mt-1} & a_{mt+1} & \cdots & a_{mn} \end{vmatrix}$

$=$ the (j, t)-cofactor of $\boldsymbol{A}$ $= A_{jt}$.

Proof of $\det(B_3) = \det(A)$ (Theorem 2.5.15)

Now, we compute $\det(C)$ in two different ways:

There exists two identical rows.

$$0 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}.$$

cofactor expansion along the j^{th} row

So $\det(B_3) = \det(A) + k(a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn})$
$= \det(A)$.

Examples (Remark 2.5.16 & Example 2.5.17.1)

Given a square matrix $\boldsymbol{A}$, we can use elementary row operations to transform a square matrix to a triangular matrix (e.g. using the Gaussian Elimination) and then compute the determinant accordingly.

$$\begin{vmatrix} 3 & -1 & 1 & 1 \\ 3 & -1 & 2 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix} \overset{R_2 - R_1}{=} \begin{vmatrix} 3 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix} \overset{R_2 \leftrightarrow R_3}{=} - \begin{vmatrix} 3 & -1 & 1 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 \end{vmatrix}$$

$$\overset{R_4 - 2R_3}{=} - \begin{vmatrix} 3 & -1 & 1 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} = -3 \cdot 2 \cdot 1 \cdot (-1) = 6.$$

Examples (Remark 2.5.17.2)

$$\boldsymbol{A} \xrightarrow{R_3 + \frac{2}{9}R_1} \xrightarrow{R_2 \leftrightarrow R_3} \xrightarrow{4R_2} \boldsymbol{B} = \begin{bmatrix} 5 & 0 & 8 & -1 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Find $\det(\boldsymbol{A})$.

$$5 \cdot (-2) \cdot 1 \cdot \frac{1}{3} = \det(\boldsymbol{B}) = 1 \cdot (-1) \cdot 4 \det(\boldsymbol{A})$$

Thus $-\frac{10}{3} = -4 \det(\boldsymbol{A})$ and hence $\det(\boldsymbol{A}) = \frac{5}{6}$.

Elementary column operations (Remark 2.5.18)

Instead of performing elementary operations on rows of a matrix, we can also perform these operations on columns.

Similar to the previous theorem (Theorem 2.5.15), one can use elementary column operations to compute the determinants of square matrices.

(Please read our textbook for the details.)

Invertible matrices (Theorem 2.5.19)

A square matrix $\boldsymbol{A}$ is invertible if and only if $\det(\boldsymbol{A}) \neq 0$.

Proof: Let $\boldsymbol{B}$ be the reduced row-echelon form of $\boldsymbol{A}$, i.e.

$$\boldsymbol{A} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \boldsymbol{B} = \boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}.$$

for some elementary matrices $\boldsymbol{E}_1, \ \boldsymbol{E}_2, \ \ldots, \ \boldsymbol{E}_k$.

Then
$$\begin{aligned}\det(\boldsymbol{B}) &= \det(\boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}) \\ &= \det(\boldsymbol{E}_k) \cdots \det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1)\det(\boldsymbol{A}).\end{aligned}$$

by Theorem 2.5.15.4

If $\boldsymbol{A}$ is invertible, then (by Theorem 2.4.7) $\boldsymbol{B} = \boldsymbol{I}$

and $1 = \det(\boldsymbol{I}) = \det(\boldsymbol{B}) = \det(\boldsymbol{E}_k) \cdots \det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1)\det(\boldsymbol{A})$

and hence $\det(\boldsymbol{A}) \neq 0$.

Invertible matrices (Theorem 2.5.19)

Suppose $\boldsymbol{A}$ is singular.

Then (by Remark 2.4.10) $\boldsymbol{B}$ has at least one zero row, i.e.

$$\boldsymbol{B} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}. \quad \longleftarrow \text{zero row}$$

So $0 = \det(\boldsymbol{B}) = \det(\boldsymbol{E}_k) \cdots \det(\boldsymbol{E}_2) \det(\boldsymbol{E}_1) \det(\boldsymbol{A})$.

Since $\det(\boldsymbol{E}_i) \neq 0$ for all i, $\det(\boldsymbol{A}) = 0$.

Scalar multiplication (Theorem 2.5.22.1)

If $\boldsymbol{A}$ is a square matrix of order n and c a scalar, then $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$.

Proof:

$$\boldsymbol{A} \xrightarrow{cR_1} \xrightarrow{cR_2} \cdots \xrightarrow{cR_n} c\boldsymbol{A}.$$

Thus $\det(c\boldsymbol{A}) = \underbrace{c \cdot c \cdots c}_{n \text{ times}} \det(\boldsymbol{A}) = c^n \det(\boldsymbol{A}).$

Matrix multiplication (Theorem 2.5.22.2)

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices of the same size, then

$$\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B}).$$

Proof: We have to consider the cases that $\boldsymbol{A}$ is singular and $\boldsymbol{A}$ is invertible separately.

Case 1: $\boldsymbol{A}$ is singular.

We learnt that (by Theorem 2.4.14) $\boldsymbol{AB}$ is also singular.

Then

$$\det(\boldsymbol{A})\det(\boldsymbol{B}) = 0 \cdot \det(\boldsymbol{B}) = 0 = \det(\boldsymbol{AB}).$$

Matrix multiplication (Theorem 2.5.22.2)

Case 2: $\boldsymbol{A}$ is invertible.

Then (by Theorem 2.4.7) $\boldsymbol{A} = \boldsymbol{E}_1\boldsymbol{E}_2 \cdots \boldsymbol{E}_k$ for some elementary matrices $\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_k$.

Thus

$$\begin{aligned}\det(\boldsymbol{AB}) &= \det(\boldsymbol{E}_1\boldsymbol{E}_2 \cdots \boldsymbol{E}_k\boldsymbol{B}) \\ &= \det(\boldsymbol{E}_1)\det(\boldsymbol{E}_2) \cdots \det(\boldsymbol{E}_k)\det(\boldsymbol{B}) \\ &= \det(\boldsymbol{E}_1\boldsymbol{E}_2 \cdots \boldsymbol{E}_k)\det(\boldsymbol{B}) \\ &= \det(\boldsymbol{A})\det(\boldsymbol{B}).\end{aligned}$$

Invertible matrices (Theorem 2.5.22.3)

If $\boldsymbol{A}$ is an invertible matrix, then

$$\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}.$$

Proof: Since $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$,

$$\det(\boldsymbol{A})\det(\boldsymbol{A}^{-1}) = \det(\boldsymbol{A}\boldsymbol{A}^{-1}) = \det(\boldsymbol{I}) = 1.$$

So $\det(\boldsymbol{A}^{-1}) = \dfrac{1}{\det(\boldsymbol{A})}$.

Examples (Example 2.5.23)

Let $\boldsymbol{A} = \begin{bmatrix} -3 & -2 & 4 \\ 4 & 3 & 1 \\ 0 & 2 & 4 \end{bmatrix}$. Note that $\det(\boldsymbol{A}) = 34$.

1. $\det(2\boldsymbol{A}) = 2^3 \det(\boldsymbol{A}) = 2^3 \cdot 34 = 272$.
2. Let $\boldsymbol{B} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$. Note that $\det(\boldsymbol{B}) = -1$.

   Then $\det(\boldsymbol{AB}) = \det(\boldsymbol{A}) \det(\boldsymbol{B}) = 34 \cdot (-1) = -34$.
3. $\det(\boldsymbol{A}^{-1}) = \dfrac{1}{\det(\boldsymbol{A})} = \dfrac{1}{34}$.

Adjoints (Definition 2.5.24)

Let $\boldsymbol{A}$ be a square matrix of order n.

The (classical) adjoint of $\boldsymbol{A}$ is the $n \times n$ matrix

$$\mathbf{adj}(\boldsymbol{A}) = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

where A_{ij} which is the (i, j)-cofactor of $\boldsymbol{A}$.

Adjoints (Theorem 2.5.25)

If $\boldsymbol{A}$ is an invertible matrix, then $\boldsymbol{A}^{-1} = \dfrac{1}{\det(\boldsymbol{A})}\,\mathbf{adj}(\boldsymbol{A})$.

Proof: It suffices to show $\boldsymbol{A}\left(\dfrac{1}{\det(\boldsymbol{A})}\,\mathbf{adj}(\boldsymbol{A})\right) = \boldsymbol{I}$.

Then (by Theorem 2.4.12) $\boldsymbol{A}^{-1} = \dfrac{1}{\det(\boldsymbol{A})}\,\mathbf{adj}(\boldsymbol{A})$.

Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ and $\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = (b_{ij})_{n\times n}$.

As

$$\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix},$$

$$b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}.$$

Adjoints (Theorem 2.5.25)

$\boldsymbol{A}\,[\,\mathbf{adj}(\boldsymbol{A})\,] = (b_{ij})_{n\times n}$
where $b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$.

Note that $b_{ii} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \det(\boldsymbol{A})$.

cofactor expansion along the i^{th} row of $\boldsymbol{A}$

For $i \neq j$,

$$0 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}.$$

$$= b_{ij}$$

There exists two identical rows.

cofactor expansion along the j^{th} row

Adjoints (Theorem 2.5.25)

$\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = (b_{ij})_{n \times n}$

where $b_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$.

Thus $\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = (b_{ij})_{n \times n}$

$$= \begin{bmatrix} \det(\boldsymbol{A}) & 0 & \cdots & 0 \\ 0 & \det(\boldsymbol{A}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \det(\boldsymbol{A}) \end{bmatrix}$$

$$= \det(\boldsymbol{A})\,\boldsymbol{I}.$$

So $\dfrac{1}{\det(\boldsymbol{A})}\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = \boldsymbol{I} \quad \Rightarrow \quad \boldsymbol{A}\left(\dfrac{1}{\det(\boldsymbol{A})}\mathbf{adj}(\boldsymbol{A})\right) = \boldsymbol{I}.$

Examples (Example 2.5.26.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc \neq 0$.

Then $\boldsymbol{M}_{11} = \begin{bmatrix} d \end{bmatrix}$,

$A_{11} = (-1)^{1+1}\,\det(\boldsymbol{M}_{11}) = d$,

Examples (Example 2.5.26.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc \neq 0$.

Then $\boldsymbol{M}_{11} = \begin{bmatrix} d \end{bmatrix}$, $\boldsymbol{M}_{12} = \begin{bmatrix} c \end{bmatrix}$,

$A_{11} = (-1)^{1+1}\det(\boldsymbol{M}_{11}) = d$, $A_{12} = (-1)^{1+2}\det(\boldsymbol{M}_{12}) = -c$,

Examples (Example 2.5.26.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc \neq 0$.

Then $\boldsymbol{M}_{11} = [\, d \,]$, $\boldsymbol{M}_{12} = [\, c \,]$, $\boldsymbol{M}_{21} = [\, b \,]$,

$A_{11} = (-1)^{1+1} \det(\boldsymbol{M}_{11}) = d$, $A_{12} = (-1)^{1+2} \det(\boldsymbol{M}_{12}) = -c$,

$A_{21} = (-1)^{2+1} \det(\boldsymbol{M}_{21}) = -b$,

Examples (Example 2.5.26.1)

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc \neq 0$.

Then $\boldsymbol{M}_{11} = \begin{bmatrix} d \end{bmatrix}$, $\boldsymbol{M}_{12} = \begin{bmatrix} c \end{bmatrix}$, $\boldsymbol{M}_{21} = \begin{bmatrix} b \end{bmatrix}$, $\boldsymbol{M}_{22} = \begin{bmatrix} a \end{bmatrix}$,

$A_{11} = (-1)^{1+1}\det(\boldsymbol{M}_{11}) = d$, $\quad A_{12} = (-1)^{1+2}\det(\boldsymbol{M}_{12}) = -c$,

$A_{21} = (-1)^{2+1}\det(\boldsymbol{M}_{21}) = -b$, $\quad A_{22} = (-1)^{2+2}\det(\boldsymbol{M}_{22}) = a$,

So $\mathbf{adj}(\boldsymbol{A}) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

and $\boldsymbol{A}^{-1} = \dfrac{1}{\det(\boldsymbol{A})}\mathbf{adj}(\boldsymbol{A}) = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \boxed{\begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix}} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \boxed{\begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}.$

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \boxed{\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix}} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^{T} = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.

$$\mathbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Examples (Example 2.5.26.2)

Let $\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$. $\qquad$ $\boldsymbol{B}^{-1} = \dfrac{1}{-2}\begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}$.

$$\textbf{adj}(\boldsymbol{B}) = \begin{bmatrix} \begin{vmatrix} -1 & 0 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \\ \begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -3 & 3 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & -1 \end{bmatrix}.$$

Cramer's Rule (Theorem 2.5.27)

Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ is a linear system

where $\boldsymbol{A} = (a_{ij})_{n \times n}$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

Let $\boldsymbol{A_i}$ be the $n \times n$ matrix obtained from $\boldsymbol{A}$ by replacing the i^{th} column of $\boldsymbol{A}$ by $\boldsymbol{b}$,

i.e. $\boldsymbol{A_i} = \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{bmatrix}$.

Cramer's Rule (Theorem 2.5.27)

If $\boldsymbol{A}$ is invertible, then the system has only one solution

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\begin{bmatrix} \det(\boldsymbol{A}_1) \\ \det(\boldsymbol{A}_2) \\ \vdots \\ \det(\boldsymbol{A}_n) \end{bmatrix}.$$

Proof: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \frac{1}{\det(\boldsymbol{A})}\,\mathbf{adj}(\boldsymbol{A})\,\boldsymbol{b}.$

$$\mathbf{adj}(\boldsymbol{A})\,\boldsymbol{b} = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1A_{11} + b_2A_{21} + \cdots + b_nA_{n1} \\ b_1A_{12} + b_2A_{22} + \cdots + b_nA_{n2} \\ \vdots \\ b_1A_{1n} + b_2A_{2n} + \cdots + b_nA_{nn} \end{bmatrix}$$

Cramer's Rule (Theorem 2.5.27)

So the solution to the system is

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\begin{bmatrix} b_1A_{11} + b_2A_{21} + \cdots + b_nA_{n1} \\ b_1A_{12} + b_2A_{22} + \cdots + b_nA_{n2} \\ \vdots \\ b_1A_{1n} + b_2A_{2n} + \cdots + b_nA_{nn} \end{bmatrix}$$

where for $i = 1, 2, ..., n$,

$$\det(\boldsymbol{A}_i) = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}$$

cofactor expansion along the i^{th} column

$$= b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni}.$$

An example (Example 2.5.28)

Consider the linear system

$$\begin{cases} x + y + 3z = 0 \\ 2x - 2y + 2z = 4 \\ 3x + 9y = 3. \end{cases}$$

First, we rewrite the system as

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix}.$$

An example (Example 2.5.28)

Then by Cramer's rule,

$$x = \frac{\begin{vmatrix} 0 & 1 & 3 \\ 4 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}} = 2.2.$$

$$y = \frac{\begin{vmatrix} 1 & 0 & 3 \\ 2 & 4 & 2 \\ 3 & 3 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}} = -0.4.$$

$$z = \frac{\begin{vmatrix} 1 & 1 & 0 \\ 2 & -2 & 4 \\ 3 & 9 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 3 \\ 2 & -2 & 2 \\ 3 & 9 & 0 \end{vmatrix}} = -0.6.$$