MA1522: Linear Algebra for Computing

Chapter 2: Matrix Algebra

2.1 Definition and Special types of Matrices

Definition

Definition

A (real-valued) *matrix* is a rectangular array of (real) numbers

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

where $a_{ij} \in \mathbb{R}$ are real numbers. The *size* of the matrix is said to be $m \times n$ (read as m by n), where m is the number of rows and n is the number of columns.

The numbers in the array are called *entries*. The *(i,j)-entry*, a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, is the number in the i-th row j-th column.

Question

Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 & 5 & 10 \\ 3 & 5 & 13 & 23 & 41 \\ -7 & 2 & 0 & 0 & 11 \end{pmatrix}$$

1. What is the size of the matrix?
2. What is the $(2,3)$-entry?

Remarks

- The size of a matrix is read as m by n. One should not multiply the numbers. For example, $\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 & 5 & 10 \\ 3 & 5 & 13 & 23 & 41 \\ -7 & 2 & 0 & 0 & 11 \end{pmatrix}$ is a 3 by 5 matrix, not a size 15 matrix.

- There is a need to add a comma when labelling the (i, j)-entry, $a_{i,j}$, if there is ambiguity. For example, a_{123} may mean the $(1, 23)$-entry, in which case we label it is as $a_{1,23}$ instead, or the $(12, 3)$-entry, in which case we label it as $a_{12,3}$ instead.

Example

We may define a matrix by some formula on its entry. For example,

$$\mathbf{A} = (a_{ij})_{2\times 3}, \quad a_{ij} = i + j$$

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

Example

$\mathbf{B} = (b_{ij})_{3\times 2}, b_{ij} = (-1)^{i+j}$

$$\mathbf{B} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}$$

Special types of Matrices

Vectors

A $n \times 1$ matrix is called a (column) *vector*, and a $1 \times n$ matrix is called a (row) *vector*.

Remark

If it is not specified whether the vector is a column or a row vector, by default we will assume it is a column vector.

Zero matrices

All entries equal 0, denoted as $\mathbf{0}_{m\times n}$. Not necessarily a square matrix.

Example

$$\mathbf{0}_{2\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{0}_{3\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{0}_{4\times 1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{0}_{1\times 1} = \begin{pmatrix} 0 \end{pmatrix}.$$

Square matrices

Number of rows = number of columns

$$\mathbf{A} = (a_{ij})_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

Example

$$\begin{pmatrix} 2 & -3 \\ 7 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}.$$

Definition

i. A size $n \times n$ matrix is a square matrix of order n.

ii. The entries a_{ii}, $i = 1, 2, \ldots, n$, (explicitly, $a_{11}, a_{22}, \ldots, a_{nn}$) are called the diagonal entries of the (square) matrix.

Diagonal, Scalar, Identity matrices

1. Diagonal matrix $\mathbf{D} = (a_{ij})_n$, $a_{ij} = 0$ for $i \neq j$. Denote as $\mathbf{D} = diag(d_1, d_2, ..., d_n) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$.
2. Scalar matrix $\mathbf{C} = (a_{ij})$, $a_{ij} = \begin{cases} c & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$, $\mathbf{C} = diag(c, c, ..., c) = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}$.
3. Identity matrix $\mathbf{I} = (a_{ij})$, $a_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$, $\mathbf{I}_n = diag(1, 1, ..., 1) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$.

A scalar matrix can also be denoted as $\mathbf{C} = c\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. See later for definition of scalar multiplication.

Question

Which of the following options are true? The 3 by 3 zero matrix $\mathbf{0}_{3\times 3}$ is a

(a) diagonal matrix.

(b) scalar matrix.

(c) identity matrix.

Triangular matrices

Upper triangular $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for all $i > j$: $\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix}$

Strictly upper triangular $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for all $i \geq j$: $\begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$

Lower triangular $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for all $i < j$: $\begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}$

Strictly lower triangular $\mathbf{A} = (a_{ij})$, $a_{ij} = 0$ for all $i \leq j$: $\begin{pmatrix} 0 & 0 & \cdots & 0 \\ * & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 0 \end{pmatrix}$

Question

Which of the following statements are true?

(i) Every square matrix is row equivalent to an upper triangular matrix.

(ii) Every square matrix is row equivalent to a strictly upper triangular matrix.

Symmetric matrices

$\mathbf{A} = (a_{ij})_n, \ a_{ij} = a_{ji}.$

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

Example

$$\begin{pmatrix} 1 & 2 & b \\ a & 0 & 4 \\ 3 & c & d \end{pmatrix} \text{ is symmetric } \Leftrightarrow a = 2, b = 3, c = 4, d \in \mathbb{R}.$$

Question

A symmetric upper triangular matrix is a zero matrix. True or false?

2.2 Matrix Algebra

Equality

Two matrices are *equal* if they have the same size and their corresponding entries are equal;

$\mathbf{A} = (a_{ij})_{n \times m}$ and $\mathbf{B} = (b_{ij})_{k \times l}$ are equal if and only if $n = k$, $m = l$, and $a_{ij} = b_{ij}$ for all $i = 1, ..., n$, $j = 1, ..., m$.

Example

1. $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \neq \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ for any $a, b, c, d, e, f \in \mathbb{R}$ since the matrices do not have the same sizes.

2. $\begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Leftrightarrow a = b = 1, c = 3, d = 2.$

Matrix Addition and Scalar Multiplication

Definition

1. Scalar multiplication: $c \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}$.
2. Matrix addition:
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

Example

1. $5 \begin{pmatrix} 6 & 1 & -1 \\ 2 & -4 & 3 \\ 4 & 9 & -11 \end{pmatrix} = \begin{pmatrix} 30 & 5 & -5 \\ 10 & -20 & 15 \\ 20 & 45 & -55 \end{pmatrix}$
2. $\begin{pmatrix} 1 & 2 & -3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 2 & 1 & 5 \\ 1 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 2 \\ 5 & 2 & 8 \end{pmatrix}$

Remark

1. Matrix addition is only defined between matrices of the same size.

2. $-\mathbf{A} = (-1)\mathbf{A}$.

3. Matrix subtraction is defined to be the addition of a negative multiple of another matrix,

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B}.$$

Properties of Matrix Addition and Scalar Multiplication

Theorem

For matrices $\mathbf{A} = (a_{ij})_{m\times n}$, $\mathbf{B} = (b_{ij})_{m\times n}$, $\mathbf{C} = (c_{ij})_{m\times n}$, *and real numbers* $a, b \in \mathbb{R}$,

(i) *(Commutative)* $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$,

(ii) *(Associative)* $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$,

(iii) *(Additive identity)* $\mathbf{0}_{m\times n} + \mathbf{A} = \mathbf{A}$,

(iv) *(Additive inverse)* $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}_{m\times n}$,

(v) *(Distributive law)* $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$,

(vi) *(Scalar addition)* $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$,

(vii) *(Associative)* $(ab)\mathbf{A} = a(b\mathbf{A})$,

(viii) *If* $a\mathbf{A} = \mathbf{0}_{m\times n}$, *then either* $a = 0$ *or* $\mathbf{A} = \mathbf{0}$.

Remarks

- The proof of the theorem can be found in the appendix.
- However, intuitively, the properties follow from the properties of addition and multiplication of real numbers, since scalar multiplication and matrix addition is defined entries-wise.
- By the associativity of matrix addition, if $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_k$ are k matrices of the same size, we simply write the sum as

$$\mathbf{A}_1 + \mathbf{A}_2 + \cdots + \mathbf{A}_k,$$

without the parenthesis (brackets).

Matrix Multiplication

$$\mathbf{AB} = (a_{ij})_{m\times p}(b_{ij})_{p\times n} = (\textstyle\sum_{k=1}^{p} a_{ik}b_{kj})_{m\times n}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33} & a_{11}b_{14}+a_{12}b_{24}+a_{13}b_{34} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33} & a_{21}b_{14}+a_{22}b_{24}+a_{23}b_{34} \end{pmatrix}$$

That is, the (i, j)-entry of the product $\mathbf{AB}$ is the sum of the product of the entries in the i-th row of $\mathbf{A}$ with the j-th column of $\mathbf{B}$.

Example

$$\begin{aligned}\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}\begin{pmatrix}1 & 1\\ 2 & 3\\ -1 & -2\end{pmatrix} &= \begin{pmatrix}(1)(1)+(2)(2)+(3)(-1) & (1)(1)+(2)(3)+(3)(-2)\\ (4)(1)+(5)(2)+(6)(-1) & (4)(1)+(5)(3)+(6)(-2)\end{pmatrix}\\ &= \begin{pmatrix}2 & 1\\ 8 & 7\end{pmatrix}\end{aligned}$$

Question

- What is the $(2,3)$-entry of the product
$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ -1 & 2 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 3 & 1 & 0 & 1 \\ 2 & -5 & 9 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 0 & 0 \\ 3 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}?$$

- What is the $(2,3)$-entry of the product
$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 1 & 3 & -2 \end{pmatrix}?$$

Remark

1. For **AB** to be defined, the number of columns of **A** must agree with the number of rows of **B**. The resultant matrix has the same number of rows as **A**, and the same number of columns as **B**.

$$(m \times p)(p \times n) = (m \times n).$$

2. Matrix multiplication is not commutative,

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

3. The order of the product matter.
   i. If we multiply **A** to the left of **B**, we are pre-multiplying **A** to **B**.
   ii. If we multiply **A** to the right of **B**, we are post-multiplying **A** to **B**.

Properties of Matrix Multiplication

Theorem

(i) *(Associative)* $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

(ii) *(Left distributive law)* $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

(iii) *(Right distributive law)* $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

(iv) *(Commute with scalar multiplication)* $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B} = \mathbf{A}(c\mathbf{B})$.

(v) *(Multiplicative identity) For any* $m \times n$ *matrix* $\mathbf{A}$, $\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n$.

(vi) *(Nonzero Zero divisor) There exists* $\mathbf{A} \neq \mathbf{0}_{m\times p}$ *and* $\mathbf{B} \neq \mathbf{0}_{p\times n}$ *such that* $\mathbf{AB} = \mathbf{0}_{m\times n}$.

(vii) *(Zero matrix) For any* $m \times n$ *matrix* $\mathbf{A}$, $\mathbf{A}\mathbf{0}_{n\times p} = \mathbf{0}_{m\times p}$ *and* $\mathbf{0}_{p\times m}\mathbf{A} = \mathbf{0}_{p\times n}$

Remarks

- To not overcrowd the slide, we left out the sizes of the matrices in the theorem, assuming that the matrices have the appropriate sizes for the operations to be well-defined. See the appendix for the detail statements.

- The proof of the theorem can be found in the appendix.

- By the associativity of matrix multiplication, if $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_k$ are k matrices of the right sizes such that their product is well-defined, we write it as

$$\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_k,$$

  without the parenthesis (brackets).

Zero Divisors

Consider

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Powers of Square Matrices

Definition

Define the power of square matrices inductively as such.

(i) $\mathbf{A}^0 = \mathbf{I}$,

(ii) $\mathbf{A}^n = \mathbf{A}\mathbf{A}^{n-1}$, for $n \geq 1$.

That is, $\mathbf{A}^n$ is $\mathbf{A}$ multiplied to itself n times, for $n \geq 2$. It follows that $\mathbf{A}^n\mathbf{A}^m = \mathbf{A}^{n+m}$ for positive integers m, n.

Example

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}^3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}$$

Note that powers of a matrix can only be defined for square matrices.

Question

Is it true that for any square matrices $\mathbf{A}$ and $\mathbf{B}$ of order n, $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$?

Challenge

Show that for diagonal matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$.

Transpose

Let $\mathbf{A} = (a_{ij})$ be a $m \times n$ matrix. The *transpose* of $\mathbf{A}$, denoted as $\mathbf{A}^T$, is the $n \times m$ matrix whose (i, j)-entry is the (j, i)-entry of $\mathbf{A}$, $\mathbf{A}^T = (b_{ij})_{n\times m}$, $b_{ij} = a_{ji}$.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m\times n} \qquad \mathbf{A}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}_{n\times m}$$

Examples

1. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 5 \end{pmatrix}$

2. $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$

Properties of Transpose

Theorem

(i) $(\mathbf{A}^T)^T = \mathbf{A}$.

(ii) $(c\mathbf{A})^T = c\mathbf{A}^T$.

(iii) $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.

(iv) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

The proof is left as an exercise

The transpose provides an alternative definition of symmetric matrix. A square matrix $\mathbf{A}$ is symmetric if $\mathbf{A} = \mathbf{A}^T$.

Example

Here is an example to demonstrate the last property $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ of the previous theorem.

$$\left(\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}\begin{pmatrix}1 & 1\\ 2 & 3\\ -1 & -2\end{pmatrix}\right)^T = \begin{pmatrix}2 & 1\\ 8 & 7\end{pmatrix}^T = \begin{pmatrix}2 & 8\\ 1 & 7\end{pmatrix} = \begin{pmatrix}1 & 2 & -1\\ 1 & 3 & -2\end{pmatrix}\begin{pmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{pmatrix}$$
$$= \begin{pmatrix}1 & 1\\ 2 & 3\\ -1 & -2\end{pmatrix}^T\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}^T$$

In fact, if $\mathbf{A}$ is a $m \times p$ matrix and $\mathbf{B}$ a $p \times n$ matrix, then $\mathbf{A}^T$ is $p \times m$ and $\mathbf{B}^T$ is $n \times p$, and thus the product $\mathbf{A}^T\mathbf{B}^T$ is not well defined if $m \neq n$.

Question

Is it true that if $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices of the same size, then so is $\mathbf{A}+\mathbf{B}$?

Challenge

Is it true that if $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices of the same order, then so is $\mathbf{AB}$?

2.3 Linear System and Matrix Equation

Matrix Equation

A linear system in standard form

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

Can be expressed as a *matrix equation*

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \qquad \mathbf{A}\mathbf{x} = \mathbf{b}$$

Here $\mathbf{A} = (a_{ij})_{m \times n}$ is called the *coefficient matrix*, $\mathbf{x} = (x_i)_{n \times 1}$ the *variable vector*, and $\mathbf{b} = (b_i)_{m \times 1}$ the *constant vector*.

Vector Equation

The linear system

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

can also be expressed as a *vector equation*:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

Here $\mathbf{a}_i$ is called the coefficient vector for variable x_i, for $i = 1, \dots, n$.

Example

The linear system

$$\begin{cases} 3x + 2y - z = 1 \\ x + 2y + z = 3 \\ x + z = 2 \end{cases}$$

can be written as a matrix equation

$$\begin{pmatrix} 3 & 2 & -1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

or a vector equation

$$x \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.$$

Example

The (unique) solution to the system is $x = \frac{1}{2}$, $y = \frac{1}{2}$, $z = \frac{3}{2}$. We can express the solution as a 3 by 1 matrix, $\begin{pmatrix} 1/2 \\ 1/2 \\ 3/2 \end{pmatrix}$. Check that

$$\begin{pmatrix} 3 & 2 & -1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 3/2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

and

$$\frac{1}{2} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + \frac{3}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.$$

Properties of Homogeneous Linear System

Recall that a linear system is homogeneous if it has the following corresponding matrix equation

$$\mathbf{Ax} = \mathbf{0},$$

for some $m \times n$ matrix $\mathbf{A}$, a variable n-vector $\mathbf{x}$, and the $m \times 1$ zero matrix (or zero m-vector) $\mathbf{0} = \mathbf{0}_{m\times 1}$.

Theorem

A homogeneous linear system is always consistent.

Proof.

Write the homogeneous linear system as $\mathbf{Ax} = \mathbf{0}$ for some $m \times n$ matrix $\mathbf{A}$. Then

$$\mathbf{A0} = \mathbf{0},$$

that is, the $m \times 1$ zero matrix is a solution to the system. $\square$

Properties of Homogeneous Linear System

Definition

The zero solution is called the *trivial solution*. If $\mathbf{x} \neq \mathbf{0}$ is a nonzero solution to the homogeneous system, it is called a *nontrivial solution*.

Theorem

A homogeneous linear system has infinitely many solutions if and only if it has a nontrivial solution.

Proof.

($\Rightarrow$) If the system has infinitely many solutions, it must surely have a nontrivial solution.

($\Leftarrow$) Suppose now $\mathbf{u} \neq \mathbf{0}$ is a nontrivial solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$, that is, $\mathbf{Au} = \mathbf{0}$. Now for any real number $s \in \mathbb{R}$, since scalar multiplication commutes with matrix multiplication,

$$\mathbf{A}(s\mathbf{u}) = s(\mathbf{Au}) = s\mathbf{0} = \mathbf{0},$$

that is, $s\mathbf{u}$ is a solution to the homogeneous linear system too. Hence, the system has infinitely many solutions. $\square$

Example

Consider the following homogeneous linear system

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

See that $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is a nontrivial solution to the system,

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

Now for any real number $s \in \mathbb{R}$, $s \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} s \\ 0 \end{pmatrix}$ is a solution to the system too,

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Question

Show that if the trivial solution is a solution to the linear system, then it must be a homogeneous linear system.

Question

Which of the following statements is/are true?

1. If the homogeneous system has a unique solution, it must be the trivial solution.

2. If the homogeneous system has the trivial solution, it must be the unique solution.

Solutions to Homogeneous and Non-homogeneous Linear System

Lemma

Let $\mathbf{v}$ *be a particular solution* $\mathbf{Ax} = \mathbf{b}$, *and* $\mathbf{u}$ *be a particular solution to the homogeneous system* $\mathbf{Ax} = \mathbf{0}$ *with the same coefficient matrix* $\mathbf{A}$. *Then* $\mathbf{v} + \mathbf{u}$ *is also a solution to* $\mathbf{Ax} = \mathbf{b}$.

Proof.

By hypothesis, $\mathbf{Av} = \mathbf{b}$ and $\mathbf{Au} = \mathbf{0}$. Hence, by the distribution properties of matrix multiplication,

$$\mathbf{A}(\mathbf{v} + \mathbf{u}) = \mathbf{Av} + \mathbf{Au} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

□

Lemma

Suppose $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *are solutions to the linear system* $\mathbf{Ax} = \mathbf{b}$. *Then* $\mathbf{v}_1 - \mathbf{v}_2$ *is a solution to the homogeneous linear system* $\mathbf{Ax} = \mathbf{0}$ *with the same coefficient matrix.*

Proof.

By hypothesis, $\mathbf{Av}_1 = \mathbf{b}$, $\mathbf{Av}_2 = \mathbf{b}$. Hence,

$$\mathbf{A}(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{Av}_1 - \mathbf{Av}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

□

Challenge

Let $\mathbf{v}$ be a particular solution to a non-homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{b} \neq \mathbf{0}$.

Show that

$$\mathbf{v} + s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + \cdots + s_k\mathbf{u}_k, \quad s_1, s_2, \ldots, s_k \in \mathbb{R}$$

is the general solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ if and only if

$$s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + \cdots + s_k\mathbf{u}_k, \quad s_1, s_2, \ldots, s_k \in \mathbb{R}$$

is the general solution to the homogeneous linear system $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Introduction to Submatrices

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

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Rows of $\mathbf{A}$:

$$\begin{gathered} \mathbf{r}_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix} \\ \mathbf{r}_2 = \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{pmatrix} \\ \vdots \\ \mathbf{r}_m = \begin{pmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \end{gathered}$$

Columns of $\mathbf{A}$:

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix} \quad \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{pmatrix} \quad \ldots \quad \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix}$$

The rows and columns of a matrix are examples of submatrices of a matrix.

Submatrix

Definition

A $p \times q$ *submatrix* of an $m \times n$ matrix $\mathbf{A}$, $p \leq m$, $q \leq n$, is formed by taking a $p \times q$ block of the entries of the matrix $\mathbf{A}$.

Example

Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 1 & 2 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix}.$$

(i) Each row is a 1 by 5 submatrix, each column is a 3 by 1 submatrix.

(ii) $\begin{pmatrix} 2 & 4 & 6 \\ 1 & 2 & 2 \end{pmatrix}$ is a 2 by 3 submatrix obtained by taking rows 2 and 3, and columns 1 to 3.

(iii) $\begin{pmatrix} 5 & 7 \\ 6 & 1 \\ 2 & 1 \end{pmatrix}$ is a 3 by 2 submatrix obtained by taking rows 1 to 3, and columns 3 and 4.

Block Multiplication

Theorem

Let $\mathbf{A}$ be an $m \times p$ matrix and $\mathbf{B}$ a $p \times n$ matrix. Let $\mathbf{A}_1$ be a $(m_2 - m_1 + 1) \times p$ submatrix of $\mathbf{A}$ obtained by taking rows m_1 to m_2, and $\mathbf{b}_1$ a $p \times (n_2 - n_1 + 1)$ submatrix of $\mathbf{B}$ obtained by taking columns n_1 to n_2. Then the product $\mathbf{A}_1\mathbf{B}_1$ is a $(m_2 - m_1 + 1) \times (n_2 - n_1 + 1)$ submatrix of $\mathbf{AB}$ obtained by taking rows m_1 to m_2 and columns n_1 to n_2.

The proof is left as an exercise. We call this *block multiplication*.

Example

Let $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 3 & 2 & 1 \\ 4 & 4 & 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 1 & 2 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix}$. Then

$$\mathbf{AB} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 3 & 2 & 1 \\ 4 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 3 & 5 & 7 & 1 \\ 2 & 4 & 6 & 1 & 2 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 5 & 7 & 8 & 4 \\ 7 & 15 & 21 & 11 & 11 \\ 8 & 19 & 29 & 24 & 10 \\ 14 & 32 & 48 & 34 & 18 \end{pmatrix}.$$

(i) $\mathbf{A}_1 = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ is obtained from $\mathbf{A}$ by taking rows 2 and 3, and $\mathbf{B}_1 = \begin{pmatrix} 5 & 7 \\ 6 & 1 \\ 2 & 1 \end{pmatrix}$ is obtained from $\mathbf{B}$ by taking columns 3 and 4. Then $\mathbf{A}_1\mathbf{B}_1 = \begin{pmatrix} 21 & 11 \\ 29 & 24 \end{pmatrix}$ is a submatrix of $\mathbf{AB}$ by taking rows 2 and 3, columns 3 and 4.

(ii) $\mathbf{A}_1 = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 4 & 2 \end{pmatrix}$ is obtained from $\mathbf{A}$ by taking rows 3 and 4, and $\mathbf{B}_1 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$ is column 2 of $\mathbf{B}$. Then $\mathbf{A}_1\mathbf{B}_1 = \begin{pmatrix} 19 \\ 32 \end{pmatrix}$ is a submatrix of $\mathbf{AB}$ by taking rows 3 and 4, and column 2.

Question

Given that

$$\begin{pmatrix}1&0&1\\1&2&2\\3&2&1\\4&4&2\end{pmatrix}\begin{pmatrix}1&3&5&7&1\\2&4&6&1&2\\1&2&2&1&3\end{pmatrix}=\begin{pmatrix}2&5&7&8&4\\7&15&21&11&11\\8&19&29&24&10\\14&32&48&34&18\end{pmatrix}.$$

1. What is $\begin{pmatrix}1&0&1\\1&2&2\\3&2&1\\4&4&2\end{pmatrix}\begin{pmatrix}3&5\\4&6\\2&2\end{pmatrix}$?

2. What is $\begin{pmatrix}4&4&2\end{pmatrix}\begin{pmatrix}1&3&5&7&1\\2&4&6&1&2\\1&2&2&1&3\end{pmatrix}$?

Block Multiplication

In particular, let $\mathbf{b}_j$ be the j-th column of $\mathbf{B}$. Then

$$\mathbf{AB} = \mathbf{A}\begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_n \end{pmatrix}.$$

That is, the j-th column of the product $\mathbf{AB}$ is the product of $\mathbf{A}$ with the j-th column of $\mathbf{B}$.

Also, if $\mathbf{a}_i$ is the i-th row of $\mathbf{A}$, then

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} \mathbf{B} = \begin{pmatrix} \mathbf{a}_1\mathbf{B} \\ \mathbf{a}_2\mathbf{B} \\ \vdots \\ \mathbf{a}_m\mathbf{B} \end{pmatrix}.$$

That is, the i-th row of the product $\mathbf{AB}$ is the product of the i-th row of $\mathbf{A}$ with $\mathbf{B}$.

Solving Matrix Equations

Let $\mathbf{A} = \begin{pmatrix} 3 & 2 & -1 \\ 5 & -1 & 3 \\ 2 & 1 & -1 \end{pmatrix}$. Find a 3×3 matrix $\mathbf{X}$ such that

$$\mathbf{AX} = \begin{pmatrix} 3 & 2 & -1 \\ 5 & -1 & 3 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix}.$$

By block multiplication, we are solving for the 3 linear systems

$$\mathbf{A} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{A} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{A} \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Example

Solve the first linear system,

$$\left(\begin{array}{ccc|c} 3 & 2 & -1 & 1 \\ 5 & -1 & 3 & 2 \\ 2 & 1 & -1 & 3 \end{array}\right) \xrightarrow[R_3-\frac{2}{3}R_1]{R_2-\frac{5}{3}R_1} \left(\begin{array}{ccc|c} 3 & 2 & -1 & 1 \\ 0 & -13/3 & 14/3 & 1/3 \\ 0 & -1/3 & -1/3 & 7/3 \end{array}\right) \xrightarrow{R_3-\frac{1}{13}R_2} \xrightarrow[-\frac{3}{13}R_2]{-\frac{13}{9}R_3} \left(\begin{array}{ccc|c} 3 & 2 & -1 & 1 \\ 0 & 1 & -14/13 & -1/13 \\ 0 & 0 & 1 & -10/3 \end{array}\right)$$

$$\xrightarrow[R_1+R_3]{R_2+\frac{14}{13}R_3} \left(\begin{array}{ccc|c} 3 & 2 & 0 & -7/3 \\ 0 & 1 & 0 & -11/3 \\ 0 & 0 & 1 & -10/3 \end{array}\right) \xrightarrow{R_1-2R_2} \xrightarrow{\frac{1}{3}R_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 5/3 \\ 0 & 1 & 0 & -11/3 \\ 0 & 0 & 1 & -10/3 \end{array}\right)$$

So, $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 5 \\ -11 \\ 10 \end{pmatrix}$.

Example

Solve the second linear system,

$$\left(\begin{array}{ccc|c} 3 & 2 & -1 & 2 \\ 5 & -1 & 3 & 1 \\ 2 & 1 & -1 & 1 \end{array}\right) \xrightarrow[R_3-\frac{2}{3}R_1]{R_2-\frac{5}{3}R_1} \left(\begin{array}{ccc|c} 3 & 2 & -1 & 2 \\ 0 & -13/3 & 14/3 & -7/3 \\ 0 & -1/3 & -1/3 & -1/3 \end{array}\right) \xrightarrow{R_3-\frac{1}{13}R_2} \xrightarrow[-\frac{3}{13}R_2]{-\frac{13}{9}R_3} \left(\begin{array}{ccc|c} 3 & 2 & -1 & 2 \\ 0 & 1 & -14/13 & 7/13 \\ 0 & 0 & 1 & 2/9 \end{array}\right)$$

$$\xrightarrow[R_1+R_3]{R_2+\frac{14}{13}R_3} \left(\begin{array}{ccc|c} 3 & 2 & 0 & 20/9 \\ 0 & 1 & 0 & 7/9 \\ 0 & 0 & 1 & 2/9 \end{array}\right) \xrightarrow{R_1-2R_2} \xrightarrow{\frac{1}{3}R_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2/9 \\ 0 & 1 & 0 & 7/9 \\ 0 & 0 & 1 & 2/9 \end{array}\right)$$

So, $\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 2 \\ 7 \\ 2 \end{pmatrix}$. The row operations used here in the reduction process is exactly those in solving the first linear system. This is because they both have the same coefficient matrix.

We can conclude likewise that the exact same row operations could be used to solve the third linear system. Hence, we could instead solve all 3 linear systems simultaneously!

Combining Augmented Matrices

Solve the following 3 linear systems

$$\begin{array}{rcrcrcl} 3x & + & 2y & - & z & = & a \\ 5x & - & y & + & 3z & = & b \\ 2x & + & y & - & z & = & c \end{array}, \text{ for } \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Matrix equation:

$$\begin{pmatrix} 3 & 2 & -1 \\ 5 & -1 & 3 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix}$$

Augmented matrix:

$$\left(\begin{array}{ccc|c|c|c} 3 & 2 & -1 & 1 & 2 & 1 \\ 5 & -1 & 3 & 2 & 1 & 1 \\ 2 & 1 & -1 & 3 & 1 & 0 \end{array}\right)$$

Solve the 3 linear system simultaneously.

Example

$$\left(\begin{array}{ccc|c|c|c} 3 & 2 & -1 & 1 & 2 & 1 \\ 5 & -1 & 3 & 2 & 1 & 1 \\ 2 & 1 & -1 & 3 & 1 & 0 \end{array}\right) \xrightarrow[R_3-\frac{2}{3}R_1]{R_2-\frac{5}{3}R_1} \left(\begin{array}{ccc|ccc} 3 & 2 & -1 & 1 & 2 & 1 \\ 0 & -13/3 & 14/3 & 1/3 & -7/3 & -2/3 \\ 0 & -1/3 & -1/3 & 7/3 & -1/3 & -2/3 \end{array}\right) \xrightarrow{R_3-\frac{1}{13}R_2} \xrightarrow[-\frac{3}{13}R_2]{-\frac{13}{9}R_3}$$

$$\left(\begin{array}{ccc|ccc} 3 & 2 & -1 & 1 & 2 & 1 \\ 0 & 1 & -14/13 & -1/13 & 7/13 & 2/13 \\ 0 & 0 & 1 & -10/3 & 2/9 & 8/9 \end{array}\right) \xrightarrow[R_1+R_3]{R_2+\frac{14}{13}R_3} \left(\begin{array}{ccc|ccc} 3 & 2 & 0 & -7/3 & 20/9 & 17/9 \\ 0 & 1 & 0 & -11/3 & 7/9 & 10/9 \\ 0 & 0 & 1 & -10/3 & 2/9 & 8/9 \end{array}\right)$$

$$\xrightarrow{R_1-2R_2} \xrightarrow{\frac{1}{3}R_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 5/3 & 2/9 & -1/9 \\ 0 & 1 & 0 & -11/3 & 7/9 & 10/9 \\ 0 & 0 & 1 & -10/3 & 2/9 & 8/9 \end{array}\right)$$

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 5 \\ -11 \\ -10 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 2 \\ 7 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} -1 \\ 10 \\ 8 \end{pmatrix}.$$

That is,

$$\mathbf{X} = \begin{pmatrix} 5/3 & 2/9 & -1/9 \\ -11/3 & 7/9 & 10/9 \\ -10/3 & 2/9 & 8/9 \end{pmatrix}$$

Combining Augmented Matrices

In general: p linear systems with the same coefficient matrix $\mathbf{A} = (a_{ij})_{m\times n}$, for $k = 1, \dots, p$,

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_{1k} \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_{2k} \\ & & & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_{mk} \end{array}\right.$$

Combined augmented matrix:

$$\left(\begin{array}{cccc|c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} & b_{11} & b_{12} & & b_{1p} \\ a_{21} & a_{22} & \cdots & a_{2n} & b_{21} & b_{22} & & b_{2p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_{m1} & b_{m2} & & b_{mp} \end{array}\right)$$

2.4 Inverse of Matrices

Introduction

1. Suppose $ab = ac$ for some $a, b, c \in \mathbb{R}$ with $a \neq 0$, then $b = c$.

2. This is because if $a \neq 0$, $\frac{1}{a}$ exists, and $a\frac{1}{a} = 1$.

3. Hence, multiplying both sides of $ab = ac$ with $\frac{1}{a}$, we have $b = 1 \times b = \frac{1}{a}ab = \frac{1}{a}ac = 1 \times c = c$.

4. This is used to solve $ax = b$ for some $a, b \in \mathbb{R}$. We can conclude that $x = \frac{b}{a}$.

5. Ideally, we want to apply the same idea to solve a linear system,

$$\mathbf{Ax} = \mathbf{b} \quad \Rightarrow \quad \mathbf{x} = \frac{\mathbf{b}}{\mathbf{A}}.$$

6. Idea: Try to define $\frac{\mathbf{1}}{\mathbf{A}}$.

Introduction

1. To define $\frac{\mathbf{1}}{\mathbf{A}}$, we need the matrix equivalent of 1, since $\frac{1}{a}$ is defined such that $\frac{1}{a}a = 1$.

2. 1 is the multiplicative identity of the real numbers, $1 \times a = a$ for any $a \in \mathbb{R}$.

3. The matrix multiplicative identity is the identity matrix, $\mathbf{I}_n$,

$$\mathbf{IA} = \mathbf{A} = \mathbf{AI}$$

   for any matrix $\mathbf{A}$.

4. So, define $\frac{\mathbf{1}}{\mathbf{A}}$ as the matrix such that $\frac{\mathbf{1}}{\mathbf{A}}\mathbf{A} = \mathbf{I}$.

Problem 1

For matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ with the appropriate sizes, if $\mathbf{AB} = \mathbf{AC}$ and $\mathbf{A} \neq \mathbf{0}$ is not the zero matrix, can we conclude that $\mathbf{B} = \mathbf{C}$?

No. For example,

$$\begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix},$$

but $\begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix}$.

Here although $\mathbf{A}$ is not the zero matrix, we do not have the cancellation law.

Problem 2

Recall that matrix multiplication is not commutative, $\mathbf{AB} \neq \mathbf{BA}$. Suppose a matrix $\frac{\mathbf{1}}{\mathbf{A}}$ exists such that $\frac{\mathbf{1}}{\mathbf{A}}\mathbf{A} = \mathbf{I}$. Is it true that $\mathbf{A}\frac{\mathbf{1}}{\mathbf{A}} = \mathbf{I}$?

No. For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}, \quad \text{but} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbf{I}$$

Conclusion

- So, it seems like our dream of solving a linear system

$$\mathbf{Ax} = \mathbf{b} \quad \Rightarrow \quad \mathbf{x} = \frac{\mathbf{b}}{\mathbf{A}}$$

  is IMPOSSIBLE!

- Except, it is possible in some cases.

- Restrict our attention only to square matrices.

Inverse of Square Matrices

Definition

A $n \times n$ square matrix $\mathbf{A}$ is *invertible* if there exists a matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}.$$

A matrix is said to be *non-invertible* otherwise.

Example

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Remarks

1. By definition, for a square matrix $\mathbf{A}$ to be invertible, there must be a $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}_n$ AND $\mathbf{AB} = \mathbf{I}_n$ simultaneously.

2. This necessitates $\mathbf{B}$ to also be a $n \times n$ square matrix.

3. Only square matrices are invertible. If $\mathbf{A}$ is a $n \times m$ matrix with $n \neq m$, then for $\mathbf{AB} = \mathbf{I}_n$, $\mathbf{B}$ must be of size $m \times n$, and hence $\mathbf{BA} \neq \mathbf{I}_n$.

4. Hence, all non-square matrices are non-invertible. In fact, we will see later that if $\mathbf{A}$ is a non-square matrix such that $\mathbf{BA} = \mathbf{I}$ (or $\mathbf{AB} = \mathbf{I}$, respectively) for some matrix $\mathbf{B}$, then there exists no matrix $\mathbf{C}$ such that $\mathbf{AC} = \mathbf{I}$ (or $\mathbf{CA} = \mathbf{I}$, respectively).

Uniqueness of inverse

Theorem

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

Proof.

Suppose $\mathbf{B}$ and $\mathbf{C}$ are such that $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$ and $\mathbf{AC} = \mathbf{I} = \mathbf{CA}$. Then by associativity of matrix multiplication,

$$\mathbf{B} = \mathbf{BI} = \mathbf{B}(\mathbf{AC}) = (\mathbf{BA})\mathbf{C} = \mathbf{IC} = \mathbf{C}.$$

□

So since the inverse is unique, we can denote the inverse of an invertible matrix $\mathbf{A}$ by $\mathbf{A}^{-1}$ and call it the *inverse* of $\mathbf{A}$.

Question: Is the identity matrix $\mathbf{I}$ invertible? If it is, what is it's inverse?

Non-Invertible Square Matrix

Not all matrices are invertible. For example, consider the matrix $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.

It is not invertible since for any order 2 square matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ a+c & b+d \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Definition

A non-invertible square matrix is called a *singular* matrix.

Remark

Some textbook use the term singular interchangeably with non-invertible, that is, they do not insist that singular matrices are square matrices.

Inverse of 2 by 2 Square Matrices

Theorem

A 2×2 *square matrix* $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *is invertible if and only if* $ad - bc \neq 0$. *In this case, the inverse is given by*

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

We will proof the claim that $\mathbf{A}$ is invertible if and only if $ad - bc \neq 0$ later. The verification that it's inverse is $\mathbf{A}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ is left as an exercise for the readers. The formula above is known as the adjoint formula for inverse.

Cancellation law for matrices

Theorem

Let $\mathbf{A}$ be an invertible matrix of order n.

(i) *(Left cancellation) If $\mathbf{B}$ and $\mathbf{C}$ are $n \times m$ matrices with $\mathbf{AB} = \mathbf{AC}$, then $\mathbf{B} = \mathbf{C}$.*

(ii) *(Right cancellation) If $\mathbf{B}$ and $\mathbf{C}$ are $m \times n$ matrices with $\mathbf{BA} = \mathbf{CA}$, then $\mathbf{B} = \mathbf{C}$.*

Proof.

(i)

$$\mathbf{AB} = \mathbf{AC} \quad \Rightarrow \quad \mathbf{B} = \mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{AC} = \mathbf{C}.$$

(ii)

$$\mathbf{BA} = \mathbf{CA} \quad \Rightarrow \quad \mathbf{B} = \mathbf{BAA}^{-1} = \mathbf{CAA}^{-1} = \mathbf{C}.$$

$\square$

Caution

If $\mathbf{AB} = \mathbf{CA}$, we cannot conclude that $\mathbf{B} = \mathbf{C}$.

Question

Suppose $\mathbf{A}$ is invertible. Can we conclude that the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is consistent? If so, what is a solution, and is it unique?

Invertibility and Linear System

Theorem

Suppose $\mathbf{A}$ *is an* $n \times n$ *invertible square matrix. Then for any* $n \times 1$ *vector* $\mathbf{b}$, $\mathbf{Ax} = \mathbf{b}$ *has a unique solution.*

Proof.

There are two claims in the theorem; it claims that the system $\mathbf{Ax} = \mathbf{b}$ is consistent, and that the solution is unique. Firstly, we will check that $\mathbf{u} = \mathbf{A}^{-1}\mathbf{b}$ is a solution. Indeed,

$$\mathbf{Au} = \mathbf{A}(\mathbf{A}^{-1}\mathbf{b}) = \mathbf{b}.$$

Now suppose $\mathbf{v}$ is also a solution, that is, $\mathbf{Av} = \mathbf{b}$. Then

$$\mathbf{Av} = \mathbf{b} = \mathbf{Au}.$$

By the cancellation law, we have $\mathbf{v} = \mathbf{u} = \mathbf{A}^{-1}\mathbf{b}$. Hence, $\mathbf{A}^{-1}\mathbf{b}$ is the unique solution to the system. $\square$

Invertibility and Linear System

Corollary

Suppose $\mathbf{A}$ is invertible. Then the trivial solution is the only solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$.

Proof.

This follows immediately from the previous theorem by letting $\mathbf{b} = \mathbf{0}$. □

Example

Consider that linear system

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}.$$

Let $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. Readers should check that $\mathbf{A}^{-1} = \begin{pmatrix} 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 \end{pmatrix}$. Solving the system, we have

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 4 \\ 0 & 1 & 1 & 6 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 2 \end{array}\right).$$

The unique solution is $\begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}$, which is exactly equal to $\mathbf{A}^{-1}\mathbf{b}$,

$$\mathbf{A}^{-1} \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}.$$

Exercise

Verify that the homogeneous system

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has only the trivial solution.

Remark

1. The theorem and its corollary on invertibility and linear system are actually equivalent statements. That is,
   (i) a square matrix $\mathbf{A}$ is invertible if and only if $\mathbf{Ax} = \mathbf{b}$ has a unique solution for all $\mathbf{b}$, and

   (ii) a square matrix $\mathbf{A}$ is invertible if and only if the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

2. In fact, one can observe that $\mathbf{Ax} = \mathbf{b}$ has a unique solution if and only if the reduced row-echelon form of the augmented matrix $\left(\mathbf{A} \mid \mathbf{b}\right)$ is $\left(\mathbf{I} \mid \mathbf{A}^{-1}\mathbf{b}\right)$.

3. This also hint towards the fact that $\mathbf{A}$ is invertible if and only if the reduced row-echelon form of $\mathbf{A}$ is the identity matrix.

4. The proofs will be given in the next section.

Algorithm to Computing Inverse

Suppose $\mathbf{A}$ is an invertible $n \times n$ matrix. By uniqueness of the inverse, there must be a unique solution to

$$\mathbf{AX} = \mathbf{I}.$$

By block multiplication, we are solving the augmented matrix

$$\left(\mathbf{A} \mid \mathbf{I}\right) \xrightarrow{RREF} \left(\mathbf{I} \mid \mathbf{A}^{-1}\right).$$

Example

1. Find the inverse of $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$.

Solve $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

$$\left(\begin{array}{cc|c|c} 3 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \xrightarrow{R_1-3R_2} \xrightarrow{R_2+2R_1} \left(\begin{array}{cc|c|c} 0 & -1 & 1 & -3 \\ 1 & 0 & 2 & -5 \end{array}\right) \xrightarrow{-R_1} \xrightarrow{R_1\leftrightarrow R_2} \left(\begin{array}{cc|c|c} 1 & 0 & 2 & -5 \\ 0 & 1 & -1 & 3 \end{array}\right).$$

So, $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$. Indeed, this is exactly the answer obtained if we used

$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{(3)(2)-(5)(1)}\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}.$$

Example

2. Find the inverse of $\begin{pmatrix} 2 & 7 & 1 \\ 1 & 4 & -1 \\ 1 & 3 & 0 \end{pmatrix}$.

$$\left(\begin{array}{ccc|ccc} 2 & 7 & 1 & 1 & 0 & 0 \\ 1 & 4 & -1 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -3/2 & -3/2 & 11/2 \\ 0 & 1 & 0 & 1/2 & 1/2 & -3/2 \\ 0 & 0 & 1 & 1/2 & -1/2 & -1/2 \end{array}\right).$$

So, $\begin{pmatrix} 2 & 7 & 1 \\ 1 & 4 & -1 \\ 1 & 3 & 0 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} -3 & -3 & 11 \\ 1 & 1 & -3 \\ 1 & -1 & -1 \end{pmatrix}$.

Question

Consider the matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. Following the steps above, we have

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{array}\right).$$

Is $\mathbf{A}$ invertible? If it is, what is its inverse?

Remarks

1. In the algorithm to finding the inverse, we are solving to $\mathbf{AX} = \mathbf{I}$. Technically, we are solving for a *right inverse*.

2. Can we guarantee that the solution is also a *left inverse*? That is, if $\mathbf{AB} = \mathbf{I}$, can we be sure that $\mathbf{BA} = \mathbf{I}$ too?

Properties of Inverses

Theorem

Let $\mathbf{A}$ *be an invertible matrix of order* n*.*

(i) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

(ii) *For any nonzero real number* $a \in \mathbb{R}$, $(a\mathbf{A})$ *is invertible with inverse* $(a\mathbf{A})^{-1} = \frac{1}{a}\mathbf{A}^{-1}$.

(iii) $\mathbf{A}^T$ *is invertible with inverse* $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

(iv) *If* $\mathbf{B}$ *is an invertible matrix of order* n*, then* $(\mathbf{AB})$ *is invertible with inverse* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Partial Proof.

The proof of (i) to (iii) is left as a exercise. For (iv),

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

$\square$

Remark

1. The precise statement for (iii) is that as follows.

   A square matrix $\mathbf{A}$ is invertible if and only if its transpose $\mathbf{A}^T$ is. In this case, the inverse of its transpose is the transpose of its inverse, $(\mathbf{A}^{-1})^T$.

2. **Caution**: $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \neq \mathbf{A}^{-1}\mathbf{B}^{-1}$.

3. If $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_k$ are invertible, the product $\mathbf{A}_1\mathbf{A}_s \cdots \mathbf{A}_k$ is invertible with $(\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1} \cdots \mathbf{A}_2^{-1}\mathbf{A}_1^{-1}$ if $\mathbf{A}_i$ is an invertible matrix for $i = 1, ..., k$.

4. The *negative power* of an invertible matrix is defined to be

$$\mathbf{A}^{-n} = (\mathbf{A}^{-1})^n$$

   for any $n > 0$.

Question

Statement (iv) of the theorem on properties of inverse says that the product of invertible matrices is invertible. Is the converse true? That is, if the product of two square matrices $\mathbf{AB}$ is invertible, can we conclude that both $\mathbf{A}$ and $\mathbf{B}$ are invertible? Why?

Question

Is the inverse of an invertible symmetric matrix symmetric?

2.5 Elementary Matrices

Elementary Matrices

Definition

A square matrix $\mathbf{E}$ of order n is called an *elementary matrix* if it can be obtained from the identity matrix $\mathbf{I}_n$ by performing a single elementary row operation

$$\mathbf{I}_n \xrightarrow{r} \mathbf{E},$$

where r is an elementary row operation. The elementary row operation is said to be the row operation corresponding to the elementary matrix.

Example

1. $\mathbf{I}_4 \xrightarrow{R_2+3R_4} \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
2. $\mathbf{I}_4 \xrightarrow{R_1 \leftrightarrow R_3} \mathbf{E} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$
3. $\mathbf{I}_4 \xrightarrow{3R_2} \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$

Elementary Matrices and Elementary Row Operations

Let $\mathbf{A}$ be an $n \times m$ matrix and let $\mathbf{E}$ be the $n \times n$ elementary matrix corresponding to the elementary row operation r. Then the product $\mathbf{EA}$ is the resultant of performing the row operation r on $\mathbf{A}$,

$$\mathbf{A} \xrightarrow{r} \mathbf{EA}.$$

That is, performing elementary row operations is equivalent to premultiplying by the corresponding elementary matrix.

Example

1. $$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} \xrightarrow{R_2+2R_1} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 4 & 5 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 4 & 5 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix}$$

Example

2. $$\begin{pmatrix}1 & 1 & 1 & 0\\ 1 & 2 & 3 & -1\\ 2 & 1 & 4 & 2\end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix}1 & 1 & 1 & 0\\ 2 & 1 & 4 & 2\\ 1 & 2 & 3 & -1\end{pmatrix}$$
$$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 1 & 0\\ 1 & 2 & 3 & -1\\ 2 & 1 & 4 & 2\end{pmatrix} = \begin{pmatrix}1 & 1 & 1 & 0\\ 2 & 1 & 4 & 2\\ 1 & 2 & 3 & -1\end{pmatrix}$$

3. $$\begin{pmatrix}1 & 1 & 1 & 0\\ 1 & 2 & 3 & -1\\ 2 & 1 & 4 & 2\end{pmatrix} \xrightarrow{-2R_2} \begin{pmatrix}1 & 1 & 1 & 0\\ -2 & -4 & -6 & 2\\ 2 & 1 & 4 & 2\end{pmatrix}$$
$$\begin{pmatrix}1 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 1 & 1 & 0\\ 1 & 2 & 3 & -1\\ 2 & 1 & 4 & 2\end{pmatrix} = \begin{pmatrix}1 & 1 & 1 & 0\\ -2 & -4 & -6 & 2\\ 2 & 1 & 4 & 2\end{pmatrix}$$

Row Equivalent Matrices

Recall that matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be row equivalent if $\mathbf{B}$ can be obtained from $\mathbf{A}$ by performing a series of elementary row operations.

Suppose now $\mathbf{B}$ is row equivalent to $\mathbf{A}$,

$$\mathbf{A} \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} \mathbf{B}.$$

Let $\mathbf{E}_i$ be the elementary matrix corresponding to the row operation r_i, for $i = 1, 2, ..., k$. Then

$$\mathbf{B} = \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}.$$

This follows from the previous discussion,

$$\mathbf{A} \xrightarrow{r_1} \mathbf{E}_1\mathbf{A} \xrightarrow{r_2} \mathbf{E}_2\mathbf{E}_1\mathbf{A} \xrightarrow{r_3} \cdots \xrightarrow{r_k} \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{B}.$$

Example

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} \xrightarrow{R_2+2R_1} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 2 \\ 3 & 4 & 5 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 3 & 4 & 5 & -1 \\ 2 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 2 \\ 3 & 4 & 5 & -1 \end{pmatrix}$$

Question

Consider the following

$$\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_2 + 3R_1} \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{B}$$

Give a single row operation that reduces $\mathbf{B}$ to the identity matrix $\mathbf{I}_3$,

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{???} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Question

Consider the following

$$\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_2 + 3R_1} \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{B}$$

Give a single row operation that reduces $\mathbf{B}$ to the identity matrix $\mathbf{I}_3$,

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_2 - 3R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Inverse of Elementary Matrices

Recall that every elementary row operations has an inverse, and that performing a row operation is equivalent to premultiplying by the corresponding elementary matrix.

Consider now a row operation r, with its corresponding elementary matrix, $\mathbf{E}$. Let r' be the reverse of the row operation r, and $\mathbf{E}'$ its corresponding elementary matrix. By definition, $\mathbf{I} \xrightarrow{r} \mathbf{E}$. Now if we apply r' to $\mathbf{E}$, since it is the reverse of r, we should get back the identity matrix,

$$\mathbf{I} \xrightarrow{r} \mathbf{E} \xrightarrow{r'} \mathbf{I}.$$

Hence, we have

$$\mathbf{I} = \mathbf{E}'\mathbf{E}\mathbf{I} = \mathbf{E}'\mathbf{E}.$$

Similarly, $\mathbf{I} \xrightarrow{r'} \mathbf{E}' \xrightarrow{r} \mathbf{I}$ tells us that

$$\mathbf{I} = \mathbf{E}\mathbf{E}'\mathbf{I} = \mathbf{E}\mathbf{E}'.$$

This shows that $\mathbf{E}$ is invertible with inverse $\mathbf{E}'$, which is also an elementary matrix.

Inverse of Elementary Matrices

Theorem

Every elementary matrices $\mathbf{E}$ *are invertible. The inverse* $\mathbf{E}^{-1}$ *is the elementary matrix corresponding to the reverse of the row operation corresponding to* $\mathbf{E}$.

(i)

$$\mathbf{I}_n \xrightarrow{R_i+cR_j} \mathbf{E} \xrightarrow{R_i-cR_j} \mathbf{I}_n \quad \Rightarrow \quad \mathbf{E} : R_i + cR_j, \ \mathbf{E}^{-1} : R_i - cR_j.$$

(ii)

$$\mathbf{I}_n \xrightarrow{R_i \leftrightarrow R_j} \mathbf{E} \xrightarrow{R_i \leftrightarrow R_j} \mathbf{I}_n \quad \Rightarrow \quad \mathbf{E} : R_i \leftrightarrow R_j, \ \mathbf{E}^{-1} : R_i \leftrightarrow R_j.$$

(iii)

$$\mathbf{I}_n \xrightarrow{cR_i} \mathbf{E} \xrightarrow{\frac{1}{c}R_i} \mathbf{I}_n \quad \Rightarrow \quad \mathbf{E} : cR_i, \ \mathbf{E}^{-1} : \frac{1}{c}R_i.$$

Example

1. $$\begin{pmatrix}1&0&0&0\\0&1&0&3\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&-3\\0&0&1&0\\0&0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\Rightarrow\begin{pmatrix}1&0&0&0\\0&1&0&3\\0&0&1&0\\0&0&0&1\end{pmatrix}^{-1}=\begin{pmatrix}1&0&0&0\\0&1&0&-3\\0&0&1&0\\0&0&0&1\end{pmatrix}$$
$$\xrightarrow{R_2+3R_4}\qquad\xrightarrow{R_2-3R_4}$$

2. $$\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\Rightarrow\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}^{-1}=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$$
$$\xrightarrow{R_1\leftrightarrow R_3}\qquad\xrightarrow{R_1\leftrightarrow R_3}$$

3. $$\begin{pmatrix}1&0&0&0\\0&2&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1/2&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\Rightarrow\begin{pmatrix}1&0&0&0\\0&2&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}^{-1}=\begin{pmatrix}1&0&0&0\\0&1/2&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$
$$\xrightarrow{2R_2}\qquad\xrightarrow{\frac{1}{2}R_2}$$

Question

Find the inverse of this elementary matrix

$$\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2.6 Equivalent Statements for Invertibility

Introduction

This section will be mainly proving important equivalent statements of invertibility. That is, as long as any one of these statements holds, we know that all the other statements will hold true too. The equivalent statements of invertibility is like a junction, where knowing that any of the statements opens you up to all the other statements, of which one of them might be useful in solving the problem you have at hand.

Before we get lost in a sea of theorems and proofs, we will be illustrating the statements with some examples.

Example

Consider the matrix $\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$. Reducing, we have

$$\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \xrightarrow{R_1 - R_3} \xrightarrow{R_2 + R_1} \xrightarrow{R_3 + R_2} \xrightarrow{\frac{1}{2} R_3} \xrightarrow{R_2 - R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here we can conclude that the reduced row-echelon form of **A** is the identity matrix, and from the previous section,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By taking the inverse of the elementary matrices, we get

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Observe that **A** is a product of elementary matrices!

Example

Next, let us evaluate the product of the elementary matrices in the reduction of $\mathbf{A}$,

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = \begin{pmatrix}1&0&0\\0&1&-1\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1/2\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&-1\\0&1&0\\0&0&1\end{pmatrix}\mathbf{A} = \begin{pmatrix}1&0&-1\\1/2&1/2&-1\\1/2&1/2&0\end{pmatrix}\mathbf{A}.$$

Observe that if we multiply the matrix to the right of $\mathbf{A}$, we do get the identity matrix too,

$$\begin{pmatrix}1&-1&1\\-1&1&1\\0&-1&1\end{pmatrix}\begin{pmatrix}1&0&-1\\1/2&1/2&-1\\1/2&1/2&0\end{pmatrix} = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

This shows that $\mathbf{A}^{-1} = \begin{pmatrix}1&0&-1\\1/2&1/2&-1\\1/2&1/2&0\end{pmatrix}$.

Example

Now consider the linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, for some $b_1, b_2, b_3 \in \mathbb{R}$. Let's solve the equation by row reduction

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & b_1 \\ -1 & 1 & 1 & b_2 \\ 0 & -1 & 1 & b_3 \end{array}\right) \xrightarrow{R_1-R_3} \xrightarrow{R_2+R_1} \xrightarrow{R_3+R_2} \xrightarrow{\frac{1}{2}R_3} \xrightarrow{R_2-R_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & b_1 - b_3 \\ 0 & 1 & 0 & b_1/2 + b_2/2 - b_3 \\ 0 & 0 & 1 & b_1/2 + b_2/2 \end{array}\right).$$

This shows that the system is not only consistent, but have a unique solution for every $\mathbf{b}$. In fact,

$$\mathbf{A}^{-1}\mathbf{b} = \begin{pmatrix} 1 & 0 & -1 \\ 1/2 & 1/2 & -1 \\ 1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 - b_3 \\ b_1/2 + b_2/2 - b_3 \\ b_1/2 + b_2/2 \end{pmatrix}.$$

Example

Now consider $\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & -1 & 1 \end{pmatrix}$. Reducing, we have

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & -1 & 1 \end{pmatrix} \xrightarrow{R_3+R_1} \xrightarrow{R_2 \leftrightarrow R_1} \xrightarrow{R_2 - R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

The reduced row-echelon form of $\mathbf{A}$ has both a non-pivot column and a zero row. Now consider the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, for some $b_1, b_2, b_3 \in \mathbb{R}$. Let's solve the equation by row reduction.

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & b_1 \\ 1 & 0 & 0 & b_2 \\ -1 & -1 & 1 & b_3 \end{array}\right) \xrightarrow{R_3+R_1} \xrightarrow{R_2 \leftrightarrow R_1} \xrightarrow{R_2 - R_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & b_2 \\ 0 & 1 & -1 & b_1 - b_2 \\ 0 & 0 & 0 & b_1 + b_3 \end{array}\right).$$

If $b_1 + b_3 \neq 0$, then the system is inconsistent. If $b_1 + b_3 = 0$, then the system has infinitely many solutions. Can $\mathbf{A}$ be written as a product of elementary matrices? Can the reduced row-echelon form of $\mathbf{A}$ be the identity matrix? Is $\mathbf{A}$ invertible?

Elementary Matrices and Inverse

Theorem

If $\mathbf{A} = \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1$ is a product of elementary matrices, then $\mathbf{A}$ is invertible.

Proof.

This follows from the fact that the product of invertible matrices is invertible. Hence, if a square matrix $\mathbf{A}$ can be written as a product of elementary matrices, then since elementary matrices are invertible, $\mathbf{A}$ invertible. $\square$

Corollary

If the reduce row-echelon form of $\mathbf{A}$ is the identity matrix, then $\mathbf{A}$ is invertible.

Proof.

By the hypothesis, $\mathbf{A}$ is row equivalent to $\mathbf{I}$. This means that there are elementary matrices $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_k$ such that

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

By premultiplying the inverse of the elementary matrices, we have

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}\mathbf{I} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}.$$

However, since the inverse of elementary matrices are elementary matrices, this shows that $\mathbf{A}$ is a product of elementary matrices, and hence invertible. $\square$

Equivalent Statements for Invertibility

Theorem

A square matrix $\mathbf{A}$ is invertible if and only if the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

Proof.

($\Rightarrow$) We have shown that if $\mathbf{A}$ is invertible, then the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

($\Leftarrow$) Now suppose the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution. This necessitates the reduced row-echelon form of $\mathbf{A}$ to be the identity matrix. For otherwise, the reduced row-echelon form must have a non-pivot column, which then implies that the homogeneous system $\mathbf{Ax} = \mathbf{0}$ has infinitely many solutions, a contradiction. Hence, by the previous corollary, $\mathbf{A}$ is invertible. $\square$

The proof also shows that if $\mathbf{A}$ is invertible, then its reduced row-echelon form is the identity matrix.

Theorem

A square matrix $\mathbf{A}$ is invertible if and only if it's reduced row-echelon form is the identity matrix.

Theorem

A square matrix $\mathbf{A}$ is invertible if and only if it is a product of elementary matrices.

The proofs of theorems are left as exercises.

Left and Right Inverses

Definition

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

(i) A $m \times n$ matrix $\mathbf{B}$ is said to be a *left inverse* of $\mathbf{A}$ if $\mathbf{BA} = \mathbf{I}_m$, where $\mathbf{I}_m$ is the $m \times m$ identity matrix.

(ii) A $m \times n$ matrix $\mathbf{B}$ is said to be a *right inverse* of $\mathbf{A}$ if $\mathbf{AB} = \mathbf{I}_n$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix.

Observe that $\mathbf{B}$ is a left inverse of $\mathbf{A}$ if and only if $\mathbf{A}$ is a right inverse of $\mathbf{B}$.

Example

1. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ is a left inverse of $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$; $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ is a right inverse of $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

2. The inverse of a square matrix $\mathbf{A}$ is both a (the) left and right inverse, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$.

Equivalent Statements for Invertibility

Theorem

A square matrix $\mathbf{A}$ *is invertible if and only if it has a left inverse.*

Proof.

($\Rightarrow$) An inverse is a left inverse.

($\Leftarrow$) Suppose now $\mathbf{B}$ is a left inverse of $\mathbf{A}$, $\mathbf{BA} = \mathbf{I}$. Consider now the homogeneous system $\mathbf{Ax} = \mathbf{0}$. Let $\mathbf{u}$ be a solution, $\mathbf{Au} = \mathbf{0}$. Then premultiplying by the left inverse $\mathbf{B}$,

$$\mathbf{0} = \mathbf{B}(\mathbf{0}) = \mathbf{B}(\mathbf{Au}) = (\mathbf{BA})\mathbf{u} = \mathbf{Iu} = \mathbf{u}$$

tells us that $\mathbf{u} = \mathbf{0}$, that is, the homogeneous system has only the trivial solution. Hence, $\mathbf{A}$ is invertible. $\square$

Equivalent Statements for Invertibility

Theorem

A square matrix $\mathbf{A}$ *is invertible if and only if it has a right inverse.*

Proof.

Recall that $\mathbf{A}$ is invertible if and only if $\mathbf{A}^T$ is. Let $\mathbf{B}$ be a right inverse of $\mathbf{A}$, $\mathbf{AB} = \mathbf{I}$. Then by taking the transpose, and observing that the identity matrix is symmetric, we have $\mathbf{B}^T\mathbf{A}^T = \mathbf{I}$. This shows that $\mathbf{B}^T$ is a left inverse of $\mathbf{A}^T$. Therefore, by the previous theorem, $\mathbf{A}^T$ is invertible, which thus proves that $\mathbf{A}$ is invertible. $\square$

Equivalent Statement for Invertibility

Theorem

A square matrix $\mathbf{A}$ is invertible if and only if $\mathbf{Ax} = \mathbf{b}$ has a unique solution for all $\mathbf{b}$.

Proof.

($\Rightarrow$) We have shown that if $\mathbf{A}$ is invertible, then $\mathbf{Ax} = \mathbf{b}$ has a unique solution for all $\mathbf{b}$.

($\Leftarrow$) Now suppose $\mathbf{Ax} = \mathbf{b}$ is consistent for all $\mathbf{b}$. In particular, $\mathbf{Ax} = \mathbf{e}_i$ is consistent, where $\mathbf{e}_i$ is the i-th column of the identity matrix, for $i = 1, ..., n$, where n is the order of $\mathbf{A}$. Let $\mathbf{b}_i$ be a solution to $\mathbf{Ax} = \mathbf{e}_i$. Let $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{pmatrix}$. Then, by block multiplication

$$\mathbf{AB} = \mathbf{A}\begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_n \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{pmatrix} = \mathbf{I},$$

which shows that $\mathbf{B}$ is a right inverse of $\mathbf{A}$. Thus, by the previous theorem, $\mathbf{A}$ is invertible. $\square$

Algorithm for Finding Inverse

Thus, we will now formally introduce the algorithm to testing if a matrix is invertible, and finding its inverse it it is invertible.

Let $\mathbf{A}$ be a $n \times n$ matrix.

Step 1: Form the $n \times 2n$ (augmented) matrix $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I}_n \end{array} \right)$.

Step 2: Reduce the matrix $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array} \right) \longrightarrow \left(\begin{array}{c|c} \mathbf{R} & \mathbf{B} \end{array} \right)$ to its REF or RREF.

Step 3: If RREF $\mathbf{R} \neq \mathbf{I}$ or REF has a zero row, then $\mathbf{A}$ is not invertible. If RREF $\mathbf{R} = \mathbf{I}$ or REF has no zero row, $\mathbf{A}$ is invertible with inverse $\mathbf{A}^{-1} = \mathbf{B}$.

Question

Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times m$ matrices. Show that $\mathbf{A}$ and $\mathbf{B}$ are row equivalent if and only if $\mathbf{B} = \mathbf{PA}$ for some invertible $n \times n$ matrix $\mathbf{P}$.

Equivalent Statements of Invertibility

Theorem (Equivalent statements of invertibility)

Let $\mathbf{A}$ *be a square matrix of order* n. *The following statements are equivalent.*

(i) $\mathbf{A}$ *is invertible.*

(ii) $\mathbf{A}^T$ *is invertible.*

(iii) *(left inverse) There is a matrix* $\mathbf{B}$ *such that* $\mathbf{BA} = \mathbf{I}$.

(iv) *(right inverse) There is a matrix* $\mathbf{B}$ *such that* $\mathbf{AB} = \mathbf{I}$.

(v) *The reduced row-echelon form of* $\mathbf{A}$ *is the identity matrix.*

(vi) $\mathbf{A}$ *can be expressed as a product of elementary matrices.*

(vii) *The homogeneous system* $\mathbf{Ax} = \mathbf{0}$ *has only the trivial solution.*

(viii) *For any* $\mathbf{b}$, *the system* $\mathbf{Ax} = \mathbf{b}$ *has a unique solution.*

2.7 LU Factorization

Example

Consider the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 1 & 0 & 1 & -1 & 0 \\ 3 & 1 & 0 & 0 & 3 \end{pmatrix}$. Reducing to a row-echelon form,

$$\mathbf{A} \xrightarrow{R_2 - R_1} \xrightarrow{R_3 - 3R_1} \xrightarrow{R_3 - \frac{5}{2} R_2} \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 & 3 \end{pmatrix}.$$

This means that

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 & 3 \end{pmatrix},$$

and thus

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & 5/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 & 3 \end{pmatrix}.$$

Example

In the example above, we may write $\mathbf{A}$ as a product of a lower triangular matrix and a row-echelon of $\mathbf{A}$,

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & 5/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 & 3 \end{pmatrix}.$$

Observe furthermore that the diagonal entries of the lower triangular matrix are 1. Such matrices are known as *unit lower triangular* matrices. We will write it as $\mathbf{A} = \mathbf{LU}$, where $\mathbf{L}$ is a unit lower triangular matrix, and $\mathbf{U}$ is a row-echelon form of $\mathbf{A}$.

Example

Consider now the linear system $\mathbf{Ax} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. Replacing $\mathbf{A}$ with $\mathbf{LU}$, we have $\mathbf{LUx} = \mathbf{b}$. Let $\mathbf{Ux} = \mathbf{y}$, and we first solve for $\mathbf{Ly} = \mathbf{b}$. But since $\mathbf{L}$ is a unit lower triangular matrix, this is easy. From

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 2 \\ 3 & 5/2 & 1 & 1 \end{array}\right)$$

We can observe that $y_1 = 1$, $y_2 = 1$, $y_3 = -9/2$ is the unique solution. Next, we solve for $\mathbf{Ux} = \mathbf{y}$. Now since $\mathbf{U}$ is in row-echelon form, this is easy too,

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & -1 & 0 & 1 \\ 0 & -2 & 0 & 0 & 0 & 1 \\ 0 & 0 & -3 & 3 & 3 & -9/2 \end{array}\right) \xrightarrow[-\frac{1}{3}R_3]{-\frac{1}{2}R_2} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & -1/2 \\ 0 & 0 & 1 & -1 & -1 & 3/2 \end{array}\right)$$

tells us that $x_1 = \frac{1}{2} - t$, $x_2 = -\frac{1}{2}$, $x_3 = \frac{3}{2} + s + t$, $x_4 = s$, $x_5 = t$, $s, t, \in \mathbb{R}$ is the general solution.

LU Factorization

Definition

A square matrix $\mathbf{L}$ is a *unit lower triangular* matrix if $\mathbf{L}$ is a lower triangular matrix with 1 in the diagonal entries.

An *LU factorization* of a $m \times n$ matrix $\mathbf{A}$ is the decomposition

$$\mathbf{A} = \mathbf{LU},$$

where $\mathbf{L}$ is a unit lower triangular matrix, and $\mathbf{U}$ is a row-echelon form of $\mathbf{A}$.

If such LU factorization exits for $\mathbf{A}$, we say that $\mathbf{A}$ is *LU factorizable*.

Unit Lower Triangular Matrices

Lemma

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be unit lower triangular matrices of the same size. Then* $\mathbf{AB}$ *is a unit lower triangular matrix too.*

Proof.

Write $\mathbf{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ b_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & 1 \end{pmatrix}$. Then

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ b_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_{21} + b_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + a_{n2}b_{21} + \cdots + b_{n1} & a_{n2} + a_{n3}b_{b2} + \cdots + b_{n2} & \cdots & 1 \end{pmatrix} \end{aligned}$$

$\square$

Unit Lower Triangular Matrix

(i) Observe that the elementary matrix $\mathbf{E}$ corresponding to the operation $R_i + cR_j$ for $i > j$ for some real number c is a lower triangular matrix.

(ii) Also, since the inverse of an elementary matrix is an elementary matrix corresponding to an elementary row operation same type, $\mathbf{E}^{-1}$ is also a unit lower triangular matrix.

(iii) Hence, by the previous lemma, if $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_k$ are series of elementary matrices corresponding to row operations of the type $R_i + cR_j$ for $i > j$ for some c, then $\mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}$ is a unit lower triangular matrix.

Algorithm to LU Factorization

Suppose $\mathbf{A} \xrightarrow{r_1, r_2, \ldots, r_k} \mathbf{U}$, where each row operation r_l is of the form $R_i + cR_j$ for some $i > j$ and real number c, and $\mathbf{U}$ is an row-echelon form of $\mathbf{A}$. Let $\mathbf{E}_i$ be the elementary matrix corresponding for r_i, for $i = 1, 2, ..., k$. Then

$$\mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{U} \quad \Rightarrow \quad \mathbf{A} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1} \mathbf{U} = \mathbf{L}\mathbf{U},$$

where $\mathbf{L} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}$. Then

$$\mathbf{A} = \mathbf{L}\mathbf{U} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{pmatrix} \begin{pmatrix} * & & & & & \cdots & * \\ 0 & \cdots & 0 & * & & \cdots & * \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & & & \cdots & * \end{pmatrix}$$

is a LU factorization of $\mathbf{A}$.

In this case, we could obtain $\mathbf{L}$ quickly without computing $\mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_k^{-1}$. For each row operation $r_l = R_i + c_l R_j$ for some $i > j$ and real number c_l, we will put $-c_l$ in the (i, j)-entry of $\mathbf{L}$.

Question

1. What if we use other row operations that are not of the type $R_i + cR_j$ for some $i > j$ and real number c? Is $\mathbf{L}$ still a unit lower triangular matrix?

2. Is it possible to reduce any matrix $\mathbf{A}$ to a row-echelon form with only the type of row operations mentioned above?

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{pmatrix} \xrightarrow[\substack{R_3 - R_1 \\ R_4 + 3R_1}]{R_2+2R_1} \begin{pmatrix} 2 & 4 & 1 & 5 & -2 \\ 0 & 3 & 5 & 2 & -3 \\ 0 & -9 & -5 & -4 & 10 \\ 0 & 12 & 10 & 12 & -5 \end{pmatrix} \xrightarrow[R_4-4R_2]{R_3+3R_2} \begin{pmatrix} 2 & 4 & 1 & 5 & -2 \\ 0 & 3 & 5 & 2 & -3 \\ 0 & 0 & 10 & 2 & 1 \\ 0 & 0 & -10 & 4 & 7 \end{pmatrix}$$

$$\xrightarrow{R_4+R_3} \begin{pmatrix} 2 & 4 & 1 & 5 & -2 \\ 0 & 3 & 5 & 2 & -3 \\ 0 & 0 & 10 & 2 & 1 \\ 0 & 0 & 0 & 6 & 8 \end{pmatrix}$$

So, $\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & -1 & 1 \end{pmatrix}$ and thus

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 1 & 5 & -2 \\ 0 & 3 & 5 & 2 & -3 \\ 0 & 0 & 10 & 2 & 1 \\ 0 & 0 & 0 & 6 & 8 \end{pmatrix}.$$

Example

Solve the system

$$\begin{pmatrix} 2 & 4 & 1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ -4 \\ 7 \end{pmatrix}.$$

First solve $\mathbf{Ly} = \mathbf{b}$, $\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 6 \\ -2 & 1 & 0 & 0 & 1 \\ 1 & -3 & 1 & 0 & -4 \\ -3 & 4 & -1 & 1 & 7 \end{array}\right)$. $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 6 \\ 13 \\ 29 \\ 2 \end{pmatrix}$ is the unique solution. Now solve for $\mathbf{Ux} = \mathbf{y}$, $\left(\begin{array}{ccccc|c} 2 & 4 & 1 & 5 & -2 & 6 \\ 0 & 3 & 5 & 2 & -3 & 13 \\ 0 & 0 & 10 & 2 & 1 & 29 \\ 0 & 0 & 0 & 6 & 8 & 2 \end{array}\right)$. The general solution is $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \frac{1}{36} \begin{pmatrix} 71 + 37s \\ -22 + 58s \\ 102 + 6s \\ 12 - 48s \\ 36s \end{pmatrix}$, $s \in \mathbb{R}$.

Questions

Let $\mathbf{A} = \mathbf{LU}$ be a LU factorization of $\mathbf{A}$.

1. Show that the system $\mathbf{Ly} = \mathbf{b}$ has a unique solution for any $\mathbf{b}$.

2. Is every matrix LU factorizable? If not, provide a counter-example.

Remarks

1. The question on the characterization of whether a matrix is LU factorizable is beyond the scope of this course.

2. Any matrix admits a LU factorization with pivoting (LUP factorization), that is, any matrix $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{PLU}$, where $\mathbf{L}$ is a unit lower triangular matrix, $\mathbf{U}$ is a row-echelon form of $\mathbf{A}$, and $\mathbf{P}$ is a permutation matrix (see below for definition).

Definition

A $n \times n$ matrix $\mathbf{P}$ is a *permutation matrix* if every rows and columns has a 1 in only one entry, and 0 everywhere else.

Equivalently, $\mathbf{P}$ is a permutation matrix if and only if $\mathbf{P}$ is the product of elementary matrices corresponding to row swaps.

2.8 Determinant by Cofactor Expansion

Order 1 and 2 Square Matrices

We will define the *determinant* of $\mathbf{A}$ of order n, denoted as $\det(\mathbf{A})$, or $|\mathbf{A}|$, by induction.

1. For $n = 1$, $\mathbf{A} = (a)$, $\det(\mathbf{A}) = a$.
2. For $n = 2$, $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\det(\mathbf{A}) = ad - bc$.

Example

(i) $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $\det(\mathbf{A}) = (1)(1) - (2)(3) = -5.$

(ii) $\mathbf{A} = \begin{pmatrix} 3 & -1 \\ 5 & 1 \end{pmatrix}$, $\det(\mathbf{A}) = (3)(1) - (-1)(5) = 8.$

Inductive Step: Matrix Minor

Suppose we have defined the determinant of all square matrices of order $\leq n-1$. Let $\mathbf{A}$ be a square matrix of order n.

Define $\mathbf{M}_{ij}$, called the *(i,j) matrix minor* of $\mathbf{A}$, to be the matrix obtained from $\mathbf{A}$ be deleting the i-th row and j-th column.

Example

$$\mathbf{A} = \begin{pmatrix} 5 & 1 & 2 & -1 \\ -1 & -3 & 1 & 3 \\ 3 & 8 & 2 & 1 \\ 2 & 0 & 1 & 11 \end{pmatrix}$$

$$\mathbf{M}_{11} = \begin{pmatrix} -3 & 1 & 3 \\ 8 & 2 & 1 \\ 0 & 1 & 11 \end{pmatrix}, \quad \mathbf{M}_{12} = \begin{pmatrix} -1 & 1 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 11 \end{pmatrix}, \quad \mathbf{M}_{23} = \begin{pmatrix} 5 & 1 & -1 \\ 3 & 8 & 1 \\ 2 & 0 & 11 \end{pmatrix}, \quad \mathbf{M}_{43} = \begin{pmatrix} 5 & 1 & -1 \\ -1 & -3 & 3 \\ 3 & 8 & 1 \end{pmatrix}$$

Inductive Step: Cofactor

The (i, j)-cofactor of $\mathbf{A}$, denoted as A_{ij}, is the (real) number given by

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

Take note of the sign of the (i, j)-entry, $(-1)^{i+j}$. Here's a visualization of the sign of the entries of the matrix

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \ddots & \end{pmatrix}.$$

Example

$$\begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\begin{aligned} A_{11} &= \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} \\ &= (1)(1) - (2)(3) \\ &= -5 \end{aligned} \qquad \begin{aligned} A_{12} &= -\begin{vmatrix} -1 & 3 \\ 3 & 1 \end{vmatrix} \\ &= -(-1)(1) + (3)(3) \\ &= 10 \end{aligned} \qquad \begin{aligned} A_{13} &= \begin{vmatrix} -1 & 1 \\ 3 & 2 \end{vmatrix} \\ &= (-1)(2) - (1)(3) \\ &= -5 \end{aligned}$$

$$\begin{aligned} A_{21} &= -\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} \\ &= -(2)(1) + (-1)(2) \\ &= -4 \end{aligned} \qquad \begin{aligned} A_{22} &= \begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix} \\ &= (1)(1) - (-1)(3) \\ &= 4 \end{aligned} \qquad \begin{aligned} A_{23} &= -\begin{vmatrix} 1 & 2 \\ 3 & 2 \end{vmatrix} \\ &= -(1)(2) + (2)(3) \\ &= 4 \end{aligned}$$

$$\begin{aligned} A_{31} &= \begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix} \\ &= (2)(3) - (-1)(1) \\ &= 7 \end{aligned} \qquad \begin{aligned} A_{32} &= -\begin{vmatrix} 1 & -1 \\ -1 & 3 \end{vmatrix} \\ &= -(1)(3) + (-1)(-1) \\ &= -2 \end{aligned} \qquad \begin{aligned} A_{33} &= \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} \\ &= (1)(1) - (2)(-1) \\ &= 3 \end{aligned}$$

Determinant by Cofactor Expansion

The *determinant* of $\mathbf{A}$ is defined to be

$$\det(\mathbf{A}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \tag{1}$$

$$= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj} \tag{2}$$

This is called the *cofactor expansion* along $\left\{ \begin{array}{ccc} \text{row} & i & (3) \\ \text{column} & j & (4) \end{array} \right.$.

Remark

The above is both a theorem and a definition. The theorem states that evaluating the cofactor expansion along any row or column produces the same result. Hence, we may define the determinant to be the cofactor expansion along any rows or columns. Readers may refer to the appendix for details.

Property of Determinant

Theorem (Determinant is invariant under tanspose)

Let $\mathbf{A}$ *be a square matrix. Then the determinant of* $\mathbf{A}$ *is equal to the determinant of* $\mathbf{A}^T$*,*

$$\det(\mathbf{A}) = \det(\mathbf{A}^T).$$

Students can try to prove the theorem, the details can be found in the appendix. The idea is that taking the cofactor expansion of $\mathbf{A}$ along column 1 is equal to the cofactor expansion of $\mathbf{A}^T$ along row 1. But the first produces $\det(\mathbf{A})$, while the latter $\det(\mathbf{A}^T)$.

Order 3 Matrices

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

$$\left|\begin{array}{ccc|cc} \mathrm{a} & \mathrm{b} & \mathrm{c} & \mathrm{a} & \mathrm{b} \\ \mathrm{d} & \mathrm{e} & \mathrm{f} & \mathrm{d} & \mathrm{e} \\ \mathrm{g} & \mathrm{h} & \mathrm{i} & \mathrm{g} & \mathrm{h} \end{array}\right|$$

Example

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

Example

1. $\begin{vmatrix} 1 & 3 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 3 \end{vmatrix} = (1)(1)(3) + (3)(0)(1) + (1)(2)(2) - (1)(1)(1) - (3)(2)(3) - (1)(0)(2) = -12$

2. $\begin{vmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{vmatrix} = 2\begin{vmatrix} 0 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 3 \end{vmatrix} - \begin{vmatrix} 1 & 3 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 3 \end{vmatrix} = 2(-3) - (-12) = 6$ (cofactor expansion along column 1)

3. $\begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = (1)\begin{vmatrix} 2 & 6 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = (1)(2)\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = (1)(2)(1)(1) = 2.$ Observe that the determinant of a triangular matrix is the product of the diagonal entries.

4. $\begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 1 & 2 & 3 \\ 0 & 2 & 0 & 2 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 2 & 2 \end{vmatrix} = (2)(1)\begin{vmatrix} 2 & 1 \\ 2 & 2 \end{vmatrix} = 2(4-2) = 4$ (cofactor expansion along column 3, then along column 1)

Question

What is the determinant of the following matrix?

$$\begin{pmatrix} 1 & 1 & 3 & 0 & 5 & -2 \\ 2 & 2 & 0 & 0 & 1 & 0 \\ -1 & 7 & 2 & 0 & 0 & 9 \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 0 & 2 & 1 \\ 1 & -5 & 7 & 0 & 8 & 0 \end{pmatrix}$$

Determinant of Triangular Matrices

Theorem (Determinant of a triangular matrix is the product of diagonal entries)

If $\mathbf{A} = (a_{ij})_n$ is a triangular matrix, then

$$\det(\mathbf{A}) = a_{11}a_{22}\cdots a_{nn} = \prod_{k=1}^{n} a_{ii}.$$

Sketch of proof.

Upper triangular matrix , continuously cofactor expand along first column,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ 0 & a_{33} & a_{34} \\ 0 & 0 & a_{44} \end{vmatrix} = a_{11}a_{22} \begin{vmatrix} a_{33} & a_{34} \\ 0 & a_{44} \end{vmatrix} = a_{11}a_{22}a_{33}a_{44}.$$

Lower triangular matrix, continuously cofactor expand along the first row,

$$\begin{vmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & 0 & 0 \\ a_{32} & a_{33} & 0 \\ a_{42} & a_{43} & a_{44} \end{vmatrix} a_{11}a_{22} \begin{vmatrix} a_{33} & 0 \\ a_{43} & a_{44} \end{vmatrix} = a_{11}a_{22}a_{33}a_{44}.$$

$\square$

Example

$$\begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = (1)(2)(1)(1) = 2.$$

Upshot: Computation of determinant of triangular matrices is easy!

But is it useful? Majority of the square matrices are not triangular. But, every matrix is related to some triangular matrix!

Question

1. Fill in the blanks. Every square matrix is ____________________ a triangular matrix.

2. Suppose $\mathbf{A}$ and $\mathbf{B}$ are row equivalent. How are the determinants of $\mathbf{A}$ and $\mathbf{B}$ related?

2.9 Determinant by Reduction

Determinant and Elementary Row Operations

$R_i + aR_j$

$$\det(\mathbf{B}) = \begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2$$

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 3 & 15 & 4 & 8 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 - 3R_1} \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{B}$$

$$\det(\mathbf{A}) = \begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 3 & 15 & 4 & 8 \\ 0 & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 1 \\ 0 & 2 & 6 \\ 3 & 15 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 6 \\ 15 & 4 \end{vmatrix} + 3\begin{vmatrix} 5 & 1 \\ 2 & 6 \end{vmatrix} = 2.$$

$$\det(\mathbf{B}) = \det(\mathbf{A}).$$

Determinant and Elementary Row Operations

cR_i

$$\det(\mathbf{B}) = \begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2$$

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 4 & 12 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{B}$$

$$\det(\mathbf{A}) = \begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 4 & 12 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 4.$$

$$\det(\mathbf{B}) = \frac{1}{2}\det(\mathbf{A}).$$

Determinant and Elementary Row Operations

$R_i \leftrightarrow R_j$

$$\det(\mathbf{B}) = \begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2$$

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 2 & 6 & 3 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_4} \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{B}$$

$$\det(\mathbf{A}) = \begin{vmatrix} 1 & 5 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 2 & 6 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 2 & 6 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 2 & 6 \end{vmatrix} = -2.$$

$$\det(\mathbf{B}) = -\det(\mathbf{A}).$$

Determinant and Elementary Row Operations

Theorem

Let $\mathbf{A}$ *be a* $n \times n$ *square matrix. Suppose* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *via a single elementary row operation. Then the determinant of* $\mathbf{B}$ *is obtained as such.*

$\mathbf{A} \xrightarrow{R_i + aR_j} \mathbf{B}$	$\det(\mathbf{B}) = \det(\mathbf{A})$
$\mathbf{A} \xrightarrow{cR_i} \mathbf{B}$	$\det(\mathbf{B}) = c\det(\mathbf{A})$
$\mathbf{A} \xrightarrow{R_i \leftrightarrow R_j} \mathbf{B}$	$\det(\mathbf{B}) = -\det(\mathbf{A})$

The proof can be found in the appendix.

Corollary

The determinant of an elementary matrix $\mathbf{E}$ *is given as such.*

(i) *If* $\mathbf{E}$ *corresponds to* $R_i + aR_j$, *then* $\det(\mathbf{E}) = 1$.

(ii) *If* $\mathbf{E}$ *corresponds to* cR_j, *then* $\det(\mathbf{E}) = c$.

(iii) *If* $\mathbf{E}$ *corresponds to* $R_i \leftrightarrow R_j$, *then* $\det(\mathbf{E}) = -1$.

Proof.

This follows immediately from the previous theorem by letting $\mathbf{A} = \mathbf{I}$. □

Determinant of Row Equivalent Matrices

Theorem

Let $\mathbf{A}$ *and* $\mathbf{R}$ *be square matrices such that*

$$\mathbf{R} = \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}$$

for some elementary matrices $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_k$. *Then*

$$\det(\mathbf{R}) = \det(\mathbf{E}_k) \cdots \det(\mathbf{E}_2) \det(\mathbf{E}_1) \det(\mathbf{A}).$$

Proof.

This follows immediately from the previous theorem and corollary. $\square$

Remark

Note that the determinant of an elementary matrix is nonzero. This means that the determinant of $\mathbf{R}$ could be computed from the determinant of $\mathbf{A}$ and vice versa.

Determinant of Row Equivalent Matrices

Corollary

Let $\mathbf{A}$ *be a* $n \times n$ *square matrix. Suppose* $\mathbf{A} \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} \mathbf{R} = \begin{pmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$, *where* $\mathbf{R}$ *is the reduced row-echelon form of* $\mathbf{A}$. *Let* $\mathbf{E}_i$ *be the elementary matrix corresponding to the elementary row operation* r_i, *for* $i = 1, ..., k$. *Then*

$$\det(\mathbf{A}) = \frac{d_1 d_2 \cdots d_n}{\det(\mathbf{E}_k) \cdots \det(\mathbf{E}_2) \det(\mathbf{E}_1)}.$$

Proof.

Using the previous theorem and the fact that the determinant of a triangular matrix is the product its diagonal entries,

$$d_1 d_2 \cdots d_n = \det(\mathbf{E}_k) \cdots \det(\mathbf{E}_2) \det(\mathbf{E}_1) \det(\mathbf{A}).$$

Hence,

$$\det(\mathbf{A}) = \frac{d_1 d_2 \cdots d_n}{\det(\mathbf{E}_k) \cdots \det(\mathbf{E}_2) \det(\mathbf{E}_1)}.$$

$\square$

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{pmatrix} \xrightarrow[\mathbf{E}_2 : R_3 - 2R_4]{\mathbf{E}_1 : R_1 - 2R_2} \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & -3 & -6 \\ 0 & 1 & 2 & 3 \end{pmatrix} \xrightarrow[\mathbf{E}_4 : -\frac{1}{3} R_3]{\mathbf{E}_3 : R_4 - R_1}$$

$$\begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{\mathbf{E}_5 : R_4 - R_3} \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix} \xrightarrow{\mathbf{E}_6 : R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$$\begin{aligned} \Rightarrow \quad 2 &= \det(\mathbf{E}_6)\det(\mathbf{E}_5)\det(\mathbf{E}_4)\det(\mathbf{E}_3)\det(\mathbf{E}_2)\det(\mathbf{E}_1)\det(\mathbf{A}) \\ &= (-1)(1)\left(-\frac{1}{3}\right)(1)(1)(1)\det(\mathbf{A}) \\ \Rightarrow \quad \det(\mathbf{A}) &= 6 \end{aligned}$$

2.10 Properties of Determinant

Determinant and Invertiblity

Theorem

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

Proof.

Suppose $\mathbf{A}$ is invertible. Then we can write $\mathbf{A} = \mathbf{E}_1\mathbf{E}_2\cdots\mathbf{E}_k$ as a product of elementary matrices. Hence,

$$\det(\mathbf{A}) = \det(\mathbf{E}_1)\det(\mathbf{E}_2)\cdots\det(\mathbf{E}_k),$$

which is nonzero since determinant of elementary matrices are nonzero.

Conversely, suppose $\mathbf{A}$ is singular. Then the last row of the reduce row-echelon form $\mathbf{R}$ of $\mathbf{A}$ is a zero row. Thus $\det(\mathbf{R}) = 0$ by cofactor expanding along the last row. Write $\mathbf{R} = \mathbf{E}_k\cdots\mathbf{E}_2\mathbf{E}_1\mathbf{A}$. Then

$$\det(\mathbf{A}) = \frac{\det(\mathbf{R})}{\det(\mathbf{E}_k)\cdots\det(\mathbf{E}_2)\det(\mathbf{E}_1)} = 0.$$

$\square$

We will add this to the equivalent statement of invertibility.

Equivalent Statements for Invertibility

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

(i) $\mathbf{A}$ is invertible.

(ii) $\mathbf{A}^T$ is invertible.

(iii) (left inverse) There is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.

(iv) (right inverse) There is a matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}$.

(v) The reduced row-echelon form of $\mathbf{A}$ is the identity matrix.

(vi) $\mathbf{A}$ can be expressed as a product of elementary matrices.

(vii) The homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

(viii) For any $\mathbf{b}$, the system $\mathbf{Ax} = \mathbf{b}$ has a unique solution.

(ix) The determinant of $\mathbf{A}$ is nonzero, $\det(\mathbf{A}) \neq 0$.

Determinant of Product of Matrices

Theorem

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be square matrices of the same size. Then*

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

Students may try to prove the theorem by considering the cases where $\mathbf{A}$ is invertible or not. The proof can be found in the appendix.

By induction, we get

$$\det(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k) = \det(\mathbf{A}_1)\det(\mathbf{A}_2)\cdots\det(\mathbf{A}_k).$$

Question

Let $\mathbf{A}$ be a LU factorizable square matrix with LU factorization

$$\mathbf{A} = \mathbf{LU} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{pmatrix} \begin{pmatrix} u_{11} & * & \cdots & * \\ 0 & u_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}.$$

What is the determinant of $\mathbf{A}$?

Determinant of inverse

Theorem

If $\mathbf{A}$ is invertible, then

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

Proof.

Using the fact that $\det(\mathbf{I}) = 1$, where $\mathbf{I}$ is the identity matrix, and that the determinant of product is the product of determinant, we have

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

Since $\mathbf{A}$ is invertible, its determinant is nonzero. Hence, we can divide both sides of the equation above by $\det(\mathbf{A})$ to obtain the conclusion

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

$\square$

Determinant of Scalar Multiplication

Theorem

For any square matrix $\mathbf{A}$ *of order* n *and scalar* $c \in \mathbb{R}$,

$$\det(c\mathbf{A}) = c^n \det(\mathbf{A}).$$

Proof.

Observe that scalar multiplication is equivalent to matrix multiplication by scalar matrix,
$c\mathbf{A} = (c\mathbf{I})\mathbf{A} = \begin{pmatrix} c & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c \end{pmatrix} \mathbf{A}$. Hence,

$$\det(c\mathbf{A}) = \det((c\mathbf{I})\mathbf{A}) = \begin{vmatrix} c & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c \end{vmatrix} \det(\mathbf{A}) = c^n \det(\mathbf{A}).$$

$\square$

Example

Let $\mathbf{A} = \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

$$\mathbf{A} \xrightarrow{R_4 - R_3} \begin{pmatrix} 1 & 5 & 1 & 2 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

So $\det(\mathbf{A}) = -2$ and $\det(\mathbf{B}) = 3$.

1. $\det(3\mathbf{A}^T) = 3^4 \det(\mathbf{A}^T) = 3^4 \det(\mathbf{A}) = -162$
2. $\det(3\mathbf{A}\mathbf{B}^{-1}) = 3^4 \det(\mathbf{A}\mathbf{B}^{-1}) = 3^4 \det(\mathbf{A}) \det(\mathbf{B}^{-1}) = 3^4 \det(\mathbf{A}) \frac{1}{\det(\mathbf{B})} = -54$
3. $\det((3\mathbf{B})^{-1}) = \frac{1}{\det(3\mathbf{B})} = \frac{1}{3^4 \det(\mathbf{B})} = \frac{1}{3^5} = \frac{1}{243}$.

Adjoint

Definition

Let $\mathbf{A}$ be a $n \times n$ square matrix. The *adjoint* of $\mathbf{A}$, denoted as $\mathbf{adj}(\mathbf{A})$, is the $n \times n$ square matrix whose (i, j) entry is the (j, i)-cofactor of $\mathbf{A}$,

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

Example

1. Let $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then
$$\mathbf{adj}(\mathbf{A}) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2. Let $\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ (see the example of computing the cofactors). Then
$$\mathbf{adj}(\mathbf{A}) = \begin{pmatrix} -5 & -4 & 7 \\ 10 & 4 & -2 \\ -5 & 4 & 3 \end{pmatrix}.$$

Adjoint Formula

Theorem

Let $\mathbf{A}$ *be a square matrix and* $\mathbf{adj}(\mathbf{A})$ *it adjoint. Then*

$$\mathbf{A}(\mathbf{adj}(\mathbf{A})) = \det(\mathbf{A})\mathbf{I},$$

where $\mathbf{I}$ *is the identity matrix.*

The proof can be found in the appendix.

Corollary (Adjoint Formula for Inverse)

Let $\mathbf{A}$ *be an invertible matrix. Then the inverse of* $\mathbf{A}$ *is given by*

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

The corollary follows immediately from the previous theorem, and the fact that $\det(\mathbf{A}) \neq 0$. From the adjoint formula for inverse, we have the inverse formula for 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Examples

1. Let $\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ (see example in computing adjoint). Then

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{adj}(\mathbf{A}) = \frac{1}{20}\begin{pmatrix} -5 & -4 & 7 \\ 10 & 4 & -2 \\ -5 & 4 & 3 \end{pmatrix}.$$

Indeed, $\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & 1 & 3 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/4 & -1/5 & 7/20 \\ 0 & 1 & 0 & 1/2 & 1/5 & -1/10 \\ 0 & 0 & 1 & -1/4 & 1/5 & 3/20 \end{array}\right)$

2. Let $\mathbf{A} = \begin{pmatrix} -1 & 2 & 0 \\ -2 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Then $\mathbf{adj}(\mathbf{A}) = \begin{pmatrix} 4 & -2 & 2 \\ 2 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ (check it). Observe that $\mathbf{A}$ is singular. Then

$$\mathbf{A}(\mathbf{adj}(\mathbf{A})) = \begin{pmatrix} -1 & 2 & 0 \\ -2 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 4 & -2 & 2 \\ 2 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Question

1. Show that a square matrix $\mathbf{A}$ is singular if and only if $\mathbf{A}$ times its adjoint is the zero matrix, $\mathbf{A}(\mathbf{adj}(\mathbf{A})) = \mathbf{0}$.

2. Is it true that $\mathbf{A}$ is singular if and only if the adjoint of $\mathbf{A}$ times $\mathbf{A}$ is the zero matrix, $(\mathbf{adj}(\mathbf{A}))\mathbf{A} = \mathbf{0}$?

Appendix

Properties of Matrix Addition and Scalar Multiplication

Theorem

For matrices $\mathbf{A} = (a_{ij})_{m\times n}$, $\mathbf{B} = (b_{ij})_{m\times n}$, $\mathbf{C} = (c_{ij})_{m\times n}$, *and real numbers* $a, b \in \mathbb{R}$,

(i) *(Commutative)* $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$,

(ii) *(Associative)* $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$,

(iii) *(Additive identity)* $\mathbf{0}_{m\times n} + \mathbf{A} = \mathbf{A}$,

(iv) *(Additive inverse)* $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}_{m\times n}$,

(v) *(Distributive law)* $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$,

(vi) *(Scalar addition)* $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$,

(vii) *(Associative)* $(ab)\mathbf{A} = a(b\mathbf{A})$,

(viii) *If* $a\mathbf{A} = \mathbf{0}_{m\times n}$, *then either* $a = 0$ *or* $\mathbf{A} = \mathbf{0}$.

Proof of the Properties of Scalar Multiplication and Matrix Addition

Proof.

(i) $\mathbf{A}+\mathbf{B} = (a_{ij}+b_{ij})_{m\times n} = (b_{ij}+a_{ij})_{m\times n} = \mathbf{B}+\mathbf{A}$, since addition of real numbers is commutative.

(ii) $\mathbf{A}+(\mathbf{B}+\mathbf{C}) = (a_{ij}+(b_{ij}+c_{ij}))_{m\times n} = ((a_{ij}+b_{ij})+c_{ij})_{m\times n} = (\mathbf{A}+\mathbf{B})+\mathbf{C}$, since addition of real numbers is associative.

(iii) $\mathbf{A}+\mathbf{0} = (a_{ij}+0)_{m\times n} = (a_{ij})_{m\times n} = \mathbf{A}$.

(iv) $\mathbf{A}+(-\mathbf{A}) = (a_{ij}+(-a_{ij}))_{m\times m} = (a_{ij}-a_{ij})_{m\times n} = (0)_{m\times n} = \mathbf{0}_{m\times n}$.

(v) $a(\mathbf{A}+\mathbf{B}) = (a(a_{ij}+b_{ij}))_{m\times n} = (aa_{ij}+ab_{ij})_{m\times n} = (aa_{ij})_{m\times n}+(ab_{ij})_{m\times n} = a(a_{ij})_{m\times n}+a(b_{ij})_{m\times n} = a\mathbf{A}+a\mathbf{B}$, where the 2^{nd} equality follows from the distributive property of real number addition and multiplication, and the 4^{th} equality follows from the definition of scalar multiplication.

(vi) $(a+b)\mathbf{A} = ((a+b)a_{ij})_{m\times n} = (aa_{ij}+ba_{ij})_{m\times n} = (aa_{ij})_{m\times n}+(ba_{ij})_{m\times n} = a(a_{ij})_{m\times n}+b(a_{ij})_{m\times n} = a\mathbf{A}+b\mathbf{A}$.

(vii) $(ab)\mathbf{A} = ((ab)a_{ij})_{m\times n} = (aba_{ij})_{m\times n} = (a(ba_{ij}))_{m\times n} = a(ba_{ij})_{m\times n} = a(b\mathbf{A})$.

(viii) Suppose $a\mathbf{A} = (aa_{ij})_{m\times n} = (0)_{m\times n}$. This means that $aa_{ij}=0$ for all i,j. So, if $a\neq 0$, then necessarily $a_{ij}=0$ for all i,j, which means that $\mathbf{A}=\mathbf{0}$.

$\square$

Properties of Matrix Multiplication

Theorem

(i) *(Associative) For matrices* $\mathbf{A} = (a_{ij})_{m\times p}$, $\mathbf{B} = (b_{ij})_{p\times q}$, *and* $\mathbf{C} = (c_{ij})_{q\times n}$ $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

(ii) *(Left distributive law) For matrices* $\mathbf{A} = (a_{ij})_{m\times p}$, $\mathbf{B} = (b_{ij})_{p\times n}$, *and* $\mathbf{C} = (c_{ij})_{p\times n}$, $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

(iii) *(Right distributive law) For matrices* $\mathbf{A} = (a_{ij})_{m\times p}$, $\mathbf{B} = (b_{ij})_{m\times p}$, *and* $\mathbf{C} = (c_{ij})_{p\times n}$, $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

(iv) *(Commute with scalar multiplication) For any real number* $c \in \mathbb{R}$, *and matrices* $\mathbf{A} = (a_{ij})_{m\times p}$, $\mathbf{B} = (b_{ij})_{p\times n}$, $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B} = \mathbf{A}(c\mathbf{B})$.

(v) *(Multiplicative identity) For any* $m \times n$ *matrix* $\mathbf{A}$, $\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n$.

(vi) *(Zero divisor) There exists* $\mathbf{A} \neq \mathbf{0}_{m\times p}$ *and* $\mathbf{B} \neq \mathbf{0}_{p\times n}$ *such that* $\mathbf{AB} = \mathbf{0}_{m\times n}$.

(vii) *(Zero matrix) For any* $m \times n$ *matrix* $\mathbf{A}$, $\mathbf{A}\mathbf{0}_{n\times p} = \mathbf{0}_{m\times p}$ *and* $\mathbf{0}_{p\times m}\mathbf{A} = \mathbf{0}_{p\times n}$.

Proof of the Properties of Matrix Multiplication

Proof.

We will check that the corresponding entries on each side agrees. The check for the size of matrices agree is trivial and is left to the reader.

(i) The (i, j)-entry of $(\mathbf{AB})\mathbf{C}$ is

$$\sum_{l=1}^{q}(\sum_{k=1}^{p} a_{ik} b_{kl}) c_{lj} = \sum_{l=1}^{q}\sum_{k=1}^{p} a_{ik} b_{kl} c_{lj}.$$

The (i, j)-entry of $\mathbf{A}(\mathbf{BC})$ is

$$\sum_{k=1}^{p} a_{ik}(\sum_{l=1}^{q} b_{kl} c_{lj}) = \sum_{k=1}^{p}\sum_{l=1}^{q} a_{ik} b_{kl} c_{lj}.$$

Since both sums has finitely many terms, the sums commute and thus the (i, j)-entry of $(\mathbf{AB})\mathbf{C}$ is equal to the (i, j)-entry of $\mathbf{A}(\mathbf{BC})$.

(ii) The (i, j)-entry of $\mathbf{A}(\mathbf{B} + \mathbf{C})$ is $\sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{p}(a_{ik} b_{kj} + a_{ik} c_{kj}) = \sum_{k=1}^{p} a_{ik} b_{kj} + \sum_{k=1}^{p} a_{ik} c_{kj}$, which is the (i, j)-entry of $\mathbf{AB} + \mathbf{AC}$.

(iii) The proof is analogous to left distributive law.

Proof of the Properties of Matrix Multiplication

Continue.

(iv) The i, j entry of $c\mathbf{AB}$ is $c(\sum_{k=1}^{p} a_{ik} b_{kj}) = \sum_{k=1}^{p}(ca_{ik})b_{kj} = \sum_{k=1}^{p} a_{ik}(cb_{kj})$.

(v) Note that $\mathbf{I} = (\delta_{ij})$, where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$. So the (i, j)-entry of $\mathbf{I}_m\mathbf{A}$ is

$$\delta_{i1}a_{1j} + \cdots + \delta_{ii}a_{ij} + \cdots + \delta_{im}a_{mj} = 0a_{1j} + \cdots + 1a_{ij} + \cdots + 0a_{mj} = a_{ij}.$$

The proof for $\mathbf{A} = \mathbf{AI}_n$ is analogous.

(vi) Consider for example $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

(vii) Left to reader, if you have read till this far, surely this proof is trivial to you.

$\square$

Row Equivalent Augmented Matrices have the Same Solutions

Theorem

Let $\mathbf{Ax} = \mathbf{b}$ *and* $\mathbf{Bx} = \mathbf{c}$ *be two linear system such that their augmented matrices* $(\mathbf{A} \mid \mathbf{b})$ *and* $(\mathbf{B} \mid \mathbf{c})$ *are row equivalent. Then* $\mathbf{u}$ *is a solution to* $\mathbf{Ax} = \mathbf{b}$ *if and only if it is a solution to* $\mathbf{Bx} = \mathbf{c}$*. That is, row equivalent augmented matrices have the same set of solutions.*

Proof.

By the hypothesis, there exists an invertible matrix $\mathbf{P}$ such that

$$(\mathbf{PA} \mid \mathbf{Pb}) = \mathbf{P}(\mathbf{A} \mid \mathbf{b}).$$

This means that $\mathbf{PA} = \mathbf{B}$ and $\mathbf{Pb} = \mathbf{c}$. Now suppuse $\mathbf{u}$ is a solution to $\mathbf{Ax} = \mathbf{b}$, that is, $\mathbf{Au} = \mathbf{b}$. Then premultiplying the equation by $\mathbf{P}$, we have

$$\mathbf{Bu} = \mathbf{PAu} = \mathbf{Pb} = \mathbf{c},$$

which shows that $\mathbf{u}$ is a solution to $\mathbf{Bx} = \mathbf{c}$ too. Conversely, suppose $\mathbf{u}$ is a solution to $\mathbf{Bx} = \mathbf{c}$, $\mathbf{Bu} = \mathbf{c}$. Premultiplying both sides of the equation by $\mathbf{P}^{-1}$, we have

$$\mathbf{Au} = \mathbf{P}^{-1}\mathbf{Bu} = \mathbf{P}^{-1}\mathbf{c} = \mathbf{b},$$

which shows that $\mathbf{u}$ is a solution to $\mathbf{Ax} = \mathbf{b}$ too. $\square$

Permutation Groups

Definition

The group of permutations of n objects, denoted as S_n, is called the *n-permutation group*. An element in the permutation group is called a *permutation*. An *transposition* is a permutation that only exchanges 2 objects.

Example

$$\begin{aligned} S_1 &= \{(1)\} \\ S_2 &= \{(1,2),(2,1)\} \\ S_3 &= \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\} \end{aligned}$$

We may also interpret a member of the permutation group as a function. For example $\sigma = (2,1,3)$ as the function $\sigma(1) = 2$, $\sigma(2) = 1$, $\sigma(3) = 3$. $(2,1)$, $(1,3,2)$, and $(3,2,1)$ are examples of inversions.

Permutation Groups

Theorem

Every permutation is a composition of transposition.

Example

1. The permutation $(2,3,1)$ is obtained by exchanging 1 and 2, then 3 and 2.
2. The permutation $(3,1,4,2)$ is obtained by exchanging 1 and 2, 1 and 3, and 3 and 4.

Definition

Let $\sigma \in S_n$ be a permutation. The *sign* of σ, denoted by $sgn(\sigma)$, is define to be

$$sgn(\sigma) = (-1)^{\text{number of transposition in the decomposition of } \sigma}.$$

Leibniz Formula for Determinant

Definition

Let $\mathbf{A} = (a_{ij})_{n\times n}$ be a $n \times n$ square matrix. Let S_n denote the n-permutation group. The *determinant* of $\mathbf{A}$ is defined to be

$$\det(\mathbf{A}) = \sum_{\sigma\in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Example

1. For $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\det(\mathbf{A}) = sgn((1,2))(a_{11}a_{22}) + sgn((2,1))(a_{12}a_{21}) = a_{11}a_{22} - a_{12}a_{21}$.
2. For $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$,

$$\begin{aligned} \det(\mathbf{A}) &= sgn((1,2,3))a_{11}a_{22}a_{33} + sgn((1,3,2))a_{11}a_{23}a_{32} + sgn((2,1,3))a_{12}a_{21}a_{33} \\ &\quad + sgn((2,3,1))a_{12}a_{23}a_{31} + sgn((3,1,2))a_{13}a_{21}a_{32} + sgn((3,2,1))a_{13}a_{22}a_{31} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned}$$

Equivalence in the Definition of Determinant

Theorem

Let $\det(\mathbf{A})$ *denote the determinant of a* $n \times n$ *square matrix* $\mathbf{A}$ *computed using the Leibniz formula. Then*

$$\begin{aligned} \det(\mathbf{A}) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj} \end{aligned}$$

That is, the determinant is equal to the cofactor expansion along any row or column.

This proves the claim that the evaluation of the cofactor expansion along any row or column produces the same result, and that the result is the determinant of a square matrix.

Proof that Determinant is Invariant under Transpose

Theorem (Determinant is invariant under transpose)

Let $\mathbf{A}$ *be a square matrix. Then the determinant of* $\mathbf{A}$ *is equal to the determinant of* $\mathbf{A}^T$*,*

$$\det(\mathbf{A}) = \det(\mathbf{A}^T).$$

Proof.

We will prove by induction on the order of $\mathbf{A}$. The theorem holds trivially for a 1×1 matrix since the transpose of a real number is itself.

Now suppose $\det(\mathbf{A}) = \det(\mathbf{A}^T)$ for all matrices of order k. Let $\mathbf{A}$ be a $k+1 \times k+1$ matrix. Write $\mathbf{A} = (a_{ij})_{k+1\times k+1}$, for all $i, j = 1, ..., n$. Cofactor expand along the first row of $\mathbf{A}$, we have $\det(\mathbf{A}) = \prod_{i=1}^{k+1} a_{1i}(-1)^{i+1} \det \mathbf{M}_{1i})$. Cofactor expand along the first column of $\mathbf{A}^T$, and noting that the (i,j)-matrix minor of $\mathbf{A}^T$ is the transpose of the (j,i)-matrix minor of $\mathbf{A}$ we have

$$\det(\mathbf{A}^T) = \prod_{i=1}^{k+1} b_{i1}(-1)^{i+1} \det \mathbf{M}_{1i}^T) = \prod_{i=1}^{k+1} a_{1i}(-1)^{i+1} \det \mathbf{M}_{1i}^T) = \prod_{i=1}^{k+1} a_{1i}(-1)^{i+1} \det \mathbf{M}_{1i}) = \det(\mathbf{A}),$$

where for the second equality we use the fact that $b_{ij} = a_{ji}$, and the induction hypothesis in the third equality since the matrix minor $\mathbf{M}_{ij}$ is a $k \times k$ matrix. $\square$

Determinant of Elementary Matrices

Theorem

Let $\mathbf{A}$ *be a* $n \times n$ *square matrix, and* $\mathbf{B}$ *the matrix obtained from* $\mathbf{A}$ *via exchanging the* k*-th row and the* l*-th row,* $\mathbf{A} \xrightarrow{R_k \leftrightarrow R_l} \mathbf{B}$*. Then*

$$\det(\mathbf{B}) = -\det(\mathbf{A}).$$

Proof.

We will proof by induction on the order n of $\mathbf{A}$ for $n \geq 2$. The case for $n = 2$ is clear. Suppose now the statement is true for all matrices of size $n \times n$. Write $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$. Fix $m \neq k, l$. Observe that for each $j = 1, ..., n_1$, the (m, j) matrix minor of $\mathbf{B}$ can be obtained from the (m, j) matrix minor of $\mathbf{A}$ by exchanging the k-th row and the l-th row. Hence, if we let $\mathbf{M}_{mj}$ be the (m, j) matrix minor of $\mathbf{A}$, then by the induction hypothesis, the determinant of (m, j) matrix minor of $\mathbf{B}$ is $-\det(\mathbf{M}_{mj})$. Observe also that $b_{mj} = a_{mj}$ for all $j = 1, ..., n+1$. Hence, cofactor expand along the m-th row, we have

$$\det(\mathbf{B}) = \sum_{i=1}^{n+1} b_{mj}(-1)^{m+j}(-\det(\mathbf{M}_{mj})) = -\sum_{i=1}^{n+1} a_{mj}(-1)^{m+j}\det(\mathbf{M}_{mj}) = -\det(\mathbf{A}).$$

$\square$

Determinant of Elementary Matrices

Theorem

Let $\mathbf{A}$ *be a square matrix, and* $\mathbf{B}$ *the matrix obtained from* $\mathbf{A}$ *via multiplying row* m *by* $c \neq 0$, $\mathbf{A} \xrightarrow{cR_m} \mathbf{B}$. *Then*

$$\det(\mathbf{B}) = c\det(\mathbf{A}).$$

Proof.

Write $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$. Then $b_{ij} = a_{ij}$ for all $i \neq m$ and all $j = 1, ..., n$, and $b_{mj} = ca_{mj}$ for all $j = 1, ..., n$. Also observe that the (m, j) matrix minor of $\mathbf{B}$ is equal to the (m, j) matrix minor of $\mathbf{A}$ for any $j = 1, ..., n$. We will denote the common (m, j) matrix minor as M_{mj}. Now cofactor expand along row m of $\mathbf{B}$,

$$\det(\mathbf{B}) = \sum_{j=1}^{n} b_{mj}\det(M_{mj}) = \sum_{j=1}^{n} ca_{mj}\det(M_{mj}) = c\sum_{j=1}^{n} a_{mj}\det(M_{mj}) = c\det(\mathbf{A}).$$

$\square$

Determinant of Elementary Matrices

Lemma

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

Proof.

We will proof by induction on the order n of $\mathbf{A}$, for $n \geq 2$. The case for $n = 2$ is clear. Suppose now the statement is true for all matrices of size $n \times n$. Let $\mathbf{A}$ be a $n+1 \times n+1$ matrix such that the k-th row is equal to the l-th row. Now compute the determinant of $\mathbf{A}$ by cofactor expansion along the m-th row, where $m \neq k, l$,

$$\det(\mathbf{A}) = \sum_{j=1}^{n+1} a_{mj}(-1)^{m+j} \det(\mathbf{M}_{mj}).$$

Since $\mathbf{M}_{mj}$ is obtained from $\mathbf{A}$ by deleting the $m \neq k, l$-th row and j-th column, $\mathbf{M}_{mj}$ is a $n \times n$ square matrix with 2 identical rows. Thus, by the induction hypothesis, $\det(\mathbf{M}_{mj}) = 0$ for all $j = 1, \dots, n+1$, and hence, $\det(\mathbf{A}) = 0$. $\square$

Lemma

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

Proof.

Follows immediately from the previous theorem, and the fact that determinant is invariant under transpose. $\square$

Determinant of Elementary Matrices

Theorem

Let $\mathbf{A}$ *be a square matrix, and* $\mathbf{B}$ *the matrix obtained from* $\mathbf{A}$ *by adding* a *times of row* l *to row* m*, for some real number* a, $\mathbf{A} \xrightarrow{R_m + aR_l} \mathbf{B}$*. Then*

$$\det(\mathbf{B}) = \det(\mathbf{A}).$$

Proof.

Write $\mathbf{A} = (a_{ij})$. Then the entries of $\mathbf{B}$ are those of $\mathbf{A}$ except for the m-th row, where the (m, j)-entry of $\mathbf{B}$ is $a_{mj} + aa_{lj}$. Now cofactor expanding along the m-th row of $\mathbf{B}$,

$$\det(\mathbf{B}) \quad = \quad \sum_{j=1}^{n}(a_{mj} + aa_{lj})A_{mj} = \sum_{j=1}^{n} a_{mj}A_{mj} + \sum_{j=1}^{n} aa_{lj}A_{mj} = \det(\mathbf{A}) + a\left(\sum_{j=1}^{n} a_{lj}A_{mj}\right),$$

where in the last equality, we note that $\sum_{j=1}^{n} a_{mj}A_{mj}$ is the determinant of $\mathbf{A}$ computed by cofactor expanding along the m-th row.

Determinant of Elementary Matrices

Comtinue.

Now consider the matrix $\mathbf{C}$ obtained from $\mathbf{A}$ by replacing the m-th row of $\mathbf{A}$ with the l-th row, that is, the m-th and l-th row of $\mathbf{C}$ are the l-th row of $\mathbf{A}$. Since $\mathbf{C}$ has 2 identical rows, $\det(\mathbf{C}) = 0$. Now since all the other rows of $\mathbf{C}$ are identical to $\mathbf{A}$ except the m-th row, the (m, j)-cofactor of $\mathbf{C}$ is A_{mj}, the (m, j)-cofactor of $\mathbf{A}$. Hence, if we cofactor expand along the m-th row of $\mathbf{C}$, remember that the m-th row is the l-th row of $\mathbf{A}$), we have

$$0 = \det(\mathbf{C}) = \sum_{j=1}^{n} a_{lj} A_{mj}.$$

Thus,

$$\det(\mathbf{B}) = \det(\mathbf{A}) + a\left(\sum_{j=1}^{n} a_{lj} A_{mj}\right) = \det(\mathbf{A}) + a0 = \det(\mathbf{A}).$$

$\square$

Determinant of Product of Matrices

Theorem

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be square matrices of the same size. Then*

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

Proof.

Suppose $\mathbf{A}$ is singular. Then the product is singular too. Hence,

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

Suppose $\mathbf{A}$ is invertible. Write $\mathbf{A} = \mathbf{E}_1\mathbf{E}_2\cdots\mathbf{E}_k$ as a product of elementary matrices. Then $\mathbf{AB} = \mathbf{E}_1\mathbf{E}_2\cdots\mathbf{E}_k\mathbf{B}$ and thus

$$\det(\mathbf{AB}) = \det(\mathbf{E}_1)\det(\mathbf{E}_2)\cdots\det(\mathbf{E}_k)\det(\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B}),$$

where in the last equality, we used the fact that $\det(\mathbf{A}) = \det(\mathbf{E}_1)\det(\mathbf{E}_2)\cdots\det(\mathbf{E}_k)$ $\square$

Proof of the Adjoint Formula

Theorem

Let $\mathbf{A}$ *be a square matrix and* $\mathbf{adj}(\mathbf{A})$ *it adjoint. Then*

$$\mathbf{A}(\mathbf{adj}(\mathbf{A})) = \det(\mathbf{A})\mathbf{I},$$

where $\mathbf{I}$ *is the identity matrix.*

Proof.

Let $(\mathbf{A}(\mathbf{adj}(\mathbf{A})))[i,j]$ denote the (i,j) entry of the product $\mathbf{A}(\mathbf{adj}(\mathbf{A}))$. Suffice to show that

$$\mathbf{A}(\mathbf{adj}(\mathbf{A}))[i,j] = \begin{cases} \det(\mathbf{A}) & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Now write $\mathbf{A} = (a_{ij})$, then

$$\mathbf{A}(\mathbf{adj}(\mathbf{A}))[i,j] = \sum_{k=1}^{n} a_{ik} A_{jk}.$$

If $i = j$, then the cofactor expansion of $\mathbf{A}$ along the i-th row.

Proof of the Adjoint Formula

Comtinue.

For $i \neq j$, consider the matrix $\mathbf{C}$ obtained from $\mathbf{A}$ by replacing the j-th row of $\mathbf{A}$ with the i-th row, that is the i-th row and the j-th row of $\mathbf{C}$ is the i-th row of $\mathbf{A}$. Since $\mathbf{C}$ has 2 identical rows, $\det(\mathbf{C}) = 0$. Also, since all the other rows of $\mathbf{C}$ are identical to $\mathbf{A}$ except the j-th row, the (j, k)-cofactor of $\mathbf{C}$ is A_{jk}, the (j, k)-cofactor of $\mathbf{A}$. Hence, cofactor expanding along the j-th row of $\mathbf{C}$, noting the the j-th row is the i-th row of $\mathbf{A}$,

$$0 = \det(\mathbf{C}) = \sum_{k=1}^{n} a_{ik} A_{jk}.$$

This completes the proof of the theorem. $\square$

Cramer's Rule

Definition

Let $\mathbf{A}$ be a $n \times n$ square matrix and $\mathbf{b}$ a $n \times 1$ vector. Construct a new matrix $\mathbf{A}_i(\mathbf{b})$ be replacing the i-th column of $\mathbf{A}$ with $\mathbf{b}$, for $i = 1, ..., n$.

Theorem (Cramer's Rule)

Let $\mathbf{A}$ be an invertible $n \times n$ matrix. For any $n \times 1$ vector $\mathbf{b}$, the unique solution to the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} \det(\mathbf{A}_1(\mathbf{b})) \\ \det(\mathbf{A}_2(\mathbf{b})) \\ \vdots \\ \det(\mathbf{A}_n(\mathbf{b})) \end{pmatrix}.$$

Cramer's Rule

Proof.

Since $\mathbf{A}$ is invertible, by the adjoint formula for inverse, the unique solution is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{\det(\mathbf{A})}\mathbf{adj}(\mathbf{A})\mathbf{b}.$$

Now the i-th entry of $\mathbf{adj}(\mathbf{A})\mathbf{b}$ is

$$\sum_{k=1}^{n}(\mathbf{adj}(\mathbf{A})[i,k])b_k = \sum_{k=1}^{n} b_k A_{ki}$$

On the other hand, by cofactor expansion along the i-th column of $\mathbf{A}_i(\mathbf{b})$, we have

$$\det(\mathbf{A}_i(\mathbf{b})) = \sum_{k=1}^{n} b_k A_{ki}$$

which is exactly the i-th entry of $\mathbf{adj}(\mathbf{A})\mathbf{b}$. Hence, the i-th entry of $\mathbf{A}^{-1}\mathbf{b}$ is $\frac{\det(\mathbf{A}_i(\mathbf{b}))}{\det(\mathbf{A})}$ as required. $\square$

Uniqueness of Reduced Row-Echelon Form

Theorem (Uniqueness of RREF)

Suppose $\mathbf{R}$ *and* $\mathbf{S}$ *are two reduced row-echelon forms of a* $m \times n$ *matrix* $\mathbf{A}$*. Then* $\mathbf{R} = \mathbf{S}$*.*

Proof.

First note that there exists an invertible matrix $\mathbf{P}$ such that

$$\mathbf{PR} = \mathbf{S}. \tag{3}$$

This is because $\mathbf{A}$ is row equivalent to $\mathbf{R}$ and $\mathbf{S}$, and so there are invertible matrices $\mathbf{P}_1, \mathbf{P}_2$ such that $\mathbf{A} = \mathbf{P}_1\mathbf{R}$ and $\mathbf{A} = \mathbf{P}_2\mathbf{S}$. Let $\mathbf{P} = \mathbf{P}_2^{-1}\mathbf{P}_1$. We will prove by induction on the numbers of rows n of $\mathbf{R}$ and $\mathbf{S}$.
Suppose $n = 1$. Then $\mathbf{R}$, $\mathbf{S}$ are row matrices and $\mathbf{P}$ is a nonzero real number. Since the leading entries of $\mathbf{R}$ and $\mathbf{S}$ must be 1, by the equation (3), $\mathbf{P} = 1$. So $\mathbf{R} = \mathbf{S}$.
Now suppose $n > 1$. Write $\mathbf{R} = \begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \cdots & \mathbf{r}_n \end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_n \end{pmatrix}$, where $\mathbf{r}_j$ and $\mathbf{s}_j$ is the j-th column of $\mathbf{R}$ and $\mathbf{S}$, respectively. By equation (3), we have

$$\mathbf{P}\mathbf{r}_j = \mathbf{s}_j, \tag{4}$$

for $j = 1, ..., n$. Since $\mathbf{P}$ is invertible, $\mathbf{R}$ and $\mathbf{S}$ must have the same zero columns. By deleting the zero columns and forming a new matrix, we may assume that $\mathbf{R}$ and $\mathbf{S}$ has no zero columns.

Uniqueness of Reduced Row-Echelon Form

Continue.
With this assumption, and the fact that **R** and **S** are in RREF, necessarily the first column of both **R** and **S** must have 1 in the first entry and 0 everywhere else. By the equation (3), the first column of **P** also have 1 in the first entry and zero everywhere else. So we write **R**, **S**, and **P** in is submatrices,

$$\mathbf{P} = \left(\begin{array}{cc} 1 & \mathbf{p}' \\ 0 & \\ \vdots & \mathbf{P}' \\ 0 & \end{array}\right), \ \mathbf{R} = \left(\begin{array}{cc} 1 & \mathbf{r}' \\ 0 & \\ \vdots & \mathbf{R}' \\ 0 & \end{array}\right), \text{ and } \mathbf{S} = \left(\begin{array}{cc} 1 & \mathbf{s}' \\ 0 & \\ \vdots & \mathbf{S}' \\ 0 & \end{array}\right),$$

where $\mathbf{p}', \mathbf{r}', \mathbf{s}'$ are row matrices. By the equation (3) and block multiplication, we have $\mathbf{P}'\mathbf{R}' = \mathbf{S}'$. Note that $\mathbf{P}'$ is invertible. Since **R** and **S** are in RREF, $\mathbf{R}'$ and $\mathbf{S}'$ are in RREF too. Hence, by the induction hypothesis, $\mathbf{R}' = \mathbf{S}'$. We are left to show that $\mathbf{r}' = \mathbf{s}'$. Since $\mathbf{R}' = \mathbf{S}'$, and both **R** and **S** are in RREF, **R** and **S** must have the same pivot columns, say columns $i_1, i_2, ..., i_r$. In these columns, the entries of $\mathbf{r}'$ and $\mathbf{s}'$ must be zero. For the nonzero entries, by equation (4), and the fact that the entries of the columns agree from second row onward, the entries in the first row of each column agrees too, that is $\mathbf{r}' = \mathbf{s}'$ too. Thus the inductive step in complete, and the statement is proven. $\square$

Row Equivalent Matrices has the same Reduced Row-Echelon Form

Theorem

Two matrices are row equivalent if and only if they have the same reduced row-echelon form.

Proof.

Suppose $\mathbf{A}$ and $\mathbf{B}$ has the same RREF $\mathbf{R}$. Then there are invertible matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{PA} = \mathbf{R}$ and $\mathbf{QB} = \mathbf{R}$. Then

$$\mathbf{Q}^{-1}\mathbf{PA} = \mathbf{Q}^{-1}\mathbf{R} = \mathbf{B}.$$

Since $\mathbf{Q}^{-1}\mathbf{P}$ is invertible, it can be written as a product of elementary matrices, and so $\mathbf{A}$ is row equivalent to $\mathbf{B}$. Suppose now $\mathbf{A}$ is row equivalent to $\mathbf{B}$. Let $\mathbf{P}$ be an invertible matrix such that $\mathbf{PA} = \mathbf{B}$. Let $\mathbf{R}$ be the RREF of $\mathbf{A}$ and $\mathbf{S}$ be the RREF of $\mathbf{B}$. Then $\mathbf{R} = \mathbf{UA}$ and $\mathbf{S} = \mathbf{VB}$ for some invertible matrices $\mathbf{U}$ and $\mathbf{V}$. Then

$$\mathbf{VPU}^{-1}\mathbf{R} = \mathbf{VPA} = \mathbf{VB} = \mathbf{S},$$

which shows that $\mathbf{R}$ is row equivalent to $\mathbf{S}$. By the uniqueness of RREF, $\mathbf{R} = \mathbf{S}$. $\square$