Matrix multiplication

Matrix

An $m \times n$ matrix is a rectangular array of numbers

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

We use capital letters, such as A, B and C, to denote matrices.

The entries of a matrix are denoted by the corresponding lowercase letter with "double subscript", such as a_{ij}. And a_{ij} is called the (i,j)-entry.

Matrix

An $m \times n$ matrix is a rectangular array of numbers

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

We use capital letters, such as A, B and C, to denote matrices.

The entries of a matrix are denoted by the corresponding lowercase letter with "double subscript", such as a_{ij}. And a_{ij} is called the (i, j)-entry.

In this way, to describe a matrix, is to know what the entries a_{ij} are.

A special case

Definition If $A = [a_{1j}]$ is a $1 \times n$ matrix (aka, a row vector) and $B = [b_{i1}]$ is an $n \times 1$ matrix (aka, a colunm vector), the *matrix product* AB is a 1×1 matrix whose entry is given by

$$a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1}.$$

That is to say,

$$AB = [a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1}].$$

Example

Find the following product matrices.

(a) AB, where $A = [1 \quad 2]$ and $B^T = [3 \quad 4]$.

(b) BA^T, where $A = [1 \quad 0 \quad 3 \quad 2]$ and $B = [1 \quad -3 \quad 2 \quad 5]$.

(c) $C^T C$, where $C^T = [2 \quad 1 \quad -3]$.

Matrix multiplication

Let A be an $m \times n$ matrix and let B be an $n \times s$ matrix (i.e., the number of columns of A equals to the number of the rows of B).

Then the *product* $C = AB$ is an $m \times s$ matrix, defined by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

for $1 \leq i \leq m$ and $1 \leq j \leq s$.

Matrix multiplication

Let A be an $m \times n$ matrix and let B be an $n \times s$ matrix (i.e., the number of columns of A equals to the number of the rows of B).

Then the *product* $C = AB$ is an $m \times s$ matrix, defined by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

for $1 \leq i \leq m$ and $1 \leq j \leq s$.

Alternatively, the rows of the matrix A and the columns of B give vectors in $\mathbb{R}^n$.
Let $\vec{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ (the i-th row of A) and $\vec{b^j} = (b_{1j}, b_{2j}, \ldots, b_{nj})$ (the j-th column of B) be two vectors in $\mathbb{R}^n$. Then $c_{ij} = \vec{a}_i \cdot \vec{b^j}$.

Lecture Note Example 7.6.

Consider the matrices shown here:

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 2 & 5 \end{bmatrix} \quad B = \begin{bmatrix} -2 & 3 \\ 1 & 4 \end{bmatrix}$$

$$C = \begin{bmatrix} 5 \\ 0 \\ 3 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & -2 & 5 \\ 0 & 3 & 4 & -1 \\ 5 & -1 & 2 & 4 \end{bmatrix}$$

How many different matrix products of the form M_1, M_2 are defined, where each of M_1 and M_2 is either one of the given matrices or the transpose of one of the given matrices?

Given the pairs of matrices A and B, find the product AB and BA, if defined.

(1)

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

(2)

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 3 & 6 & 5 \\ 2 & -4 & 7 \end{bmatrix} \quad B = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

(3) Find CC^T, where $C^T = [2 \quad 3 \quad -1 \quad 4]$.

(4) Let

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

What is AI_n? How about I_mA?

(5) Let

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

Find AB and BA. Show that $AB \neq BA$.

Powers of matrices

Definition Let A be a square matrix of order n (i.e., an $n \times n$ matrix). Then

$$\begin{aligned} A^1 &= A \\ A^2 &= AA \\ A^3 &= AA^2 \\ &\vdots \\ A^k &= AA^{k-1} \quad (k \geq 2) \end{aligned}$$

Examples

(1) Let

$$A = \begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix}$$

Find matrices A^1, A^2, A^3 and A^4.

(2) Consider the following matrices (the first two are diagonal matrices)

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Find I^k, A^k and B^k for $k \geq 1$.

Theorem

Theorem Properties of Operations for Matrices

Let A, B and C be matrices. Let a and b be scalars. Assume that the dimensions of the matrices are such that each operation is defined.

1. $A + B = B + A$ (matrix addition is commutative)
2. $A + (B + C) = (A + B) + C$ (matrix addition is also associative)

Theorem

Theorem Properties of Operations for Matrices

Let A, B and C be matrices. Let a and b be scalars. Assume that the dimensions of the matrices are such that each operation is defined.
1. $A + B = B + A$ (matrix addition is commutative)

2. $A + (B + C) = (A + B) + C$ (matrix addition is also associative)

3. $A + \mathit{0} = A$, where *0* denotes the zero matrix with the same dimensions as A.

4. $A + (-A) = \mathit{0}$, where again *0* is the zero matrix with the same dimensions as A.

Theorem

Theorem Properties of Operations for Matrices

Let A, B and C be matrices. Let a and b be scalars. Assume that the dimensions of the matrices are such that each operation is defined.
1. $A + B = B + A$ (matrix addition is commutative)

2. $A + (B + C) = (A + B) + C$ (matrix addition is also associative)

3. $A + \mathit{0} = A$, where $\mathit{0}$ denotes the zero matrix with the same dimensions as A.

4. $A + (-A) = \mathit{0}$, where again $\mathit{0}$ is the zero matrix with the same dimensions as A.

5. $A(BC) = (AB)C$ (matrix multiplication is associative)

6. $AI_n = A$ and $I_mA = A$, where A is an $m \times n$ matrix

7. $A(B + C) = AB + AC$ and $(B + C)A = BA + CA$
(matrix multiplication is distributive over matrix addition)

8. $a(B + C) = aB + aC$
(scalar multiplication of matrices is distributive over matrix addition)

9. $(a + b)C = aC + bC$
(scalar multiplication of matrices is also distributive over scalar addition)

7. $A(B+C) = AB + AC$ and $(B+C)A = BA + CA$
(matrix multiplication is distributive over matrix addition)

8. $a(B+C) = aB + aC$
(scalar multiplication of matrices is distributive over matrix addition)

9. $(a+b)C = aC + bC$
(scalar multiplication of matrices is also distributive over scalar addition)

10. $(ab)C = a(bC)$

11. $1A = A$ Note: 1 is the scalar 1.

12. $A\mathit{0} = \mathit{0}$ and $\mathit{0}A = \mathit{0}$, where 0 denotes a zero matrix and the two zero matrices have appropriate dimensions.

13. $a\mathit{0} = \mathit{0}$

14. $a(AB) = (aA)B = A(aB)$

15. $(A+B)^T = A^T + B^T$
(matrix transposition is distributive over matrix addition)

16. $(AB)^T = B^T A^T$
(matrix transposition is distributive over matrix multiplication, but the order of multiplication is reversed)

Examples

Let

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

Find A^TA, AA^T, AB^T, $(AB^T)^T$, B^TA, BA^T, $A(BC)$, $(AB)C$, $A(B+C)$ $AB + AC$ and $A + C$ if defined.

Matrix equation and SLE

- Express an SLE in terms of a matrix equation

Consider m linear equations with n variables x_1, x_2 ..., x_n

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

Consider m linear equations with n variables x_1, x_2 ..., x_n

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

Recall that the coefficient matrix of an SLE is a matrix whose i-th row is given by the coefficients in front of variables at the i-th equation. So

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

is the coefficient matrix of the SLE above.

Because of the matrix multiplication, we have

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

Because of the matrix multiplication, we have

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

If we denote $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, then the matrix equation $A\vec{x} = \vec{b}$ exactly represents the SLE.

For instance, consider the SLE

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

The corresponding matrix equation is

For instance, consider the SLE

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

The corresponding matrix equation is

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

We can think of an augmented matrix as an “abbreviation” of a matrix equation.

Some remarks

• When we have a solution of an SLE, we write

$$(x, y, z) = (1, 2, 3)$$

instead of

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

• In a matrix equation $A\vec{x} = \vec{b}$ representing an SLE, A is the coefficient matrix, $\vec{x}$ is a column vector consisting of all variables and $\vec{b}$ is a column vector consisting of all constants.

Definition

Definition Any SLE involving m equations and n variables can be represented by the *matrix form* of the SLE $A\vec{x} = \vec{b}$, where A is the $m \times n$ coefficient matrix, $\vec{x}$ is the column vector of the unknowns and $\vec{b}$ is the column vector of right hand side values.

This means that solving an SLE is equivalent to find all $\vec{x}$ such that $A\vec{x} = \vec{b}$.

How to do this?

Definition

Definition Any SLE involving m equations and n variables can be represented by the *matrix form* of the SLE $A\vec{x} = \vec{b}$, where A is the $m \times n$ coefficient matrix, $\vec{x}$ is the column vector of the unknowns and $\vec{b}$ is the column vector of right hand side values.

This means that solving an SLE is equivalent to find all $\vec{x}$ such that $A\vec{x} = \vec{b}$.

How to do this? Think about a single equation $2x = 3$. We know that $(\frac{1}{2})(2) = 1$ so that $x = (\frac{1}{2})(2x) = (\frac{1}{2})(3) = \frac{3}{2}$.

Definition

Definition Any SLE involving m equations and n variables can be represented by the *matrix form* of the SLE $A\vec{x} = \vec{b}$, where A is the $m \times n$ coefficient matrix, $\vec{x}$ is the column vector of the unknowns and $\vec{b}$ is the column vector of right hand side values.

This means that solving an SLE is equivalent to find all $\vec{x}$ such that $A\vec{x} = \vec{b}$.

How to do this? Think about a single equation $2x = 3$. We know that $(\frac{1}{2})(2) = 1$ so that $x = (\frac{1}{2})(2x) = (\frac{1}{2})(3) = \frac{3}{2}$.

We would like to mimic this process in a case of matrix equations. That is, if A is a **square matrix**, find a matrix B such that $BA = I$. Thus $\vec{x} = BA\vec{x} = B\vec{b}$. If such B exists, it is called the *inverse* matrix of A.