Matrix Algebra and Applications

MATH 2111 Lecture Notes

LI, Hongrui

November 3, 2024

Contents

1 Systems of Linear Equations

2 Matrix Algebra

2.1 Matrix Operations

Sums and Scalar Multiplication

Example 2.1. Matrix Addition

1. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$, then $A + B = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$, then $A - B = \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix}$

Given two matrices A and B of the same size, the sum of A and B is the matrix obtained by adding corresponding elements of A and B. The difference of A and B is the matrix obtained by subtracting corresponding elements of A and B.

Example 2.2. Scalar Multiplication

1. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $k = 2$, then $kA = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $k = -1$, then $kA = \begin{bmatrix} -1 & -2 \\ -3 & -4 \end{bmatrix}$

Given a matrix A and a scalar k, the product of k and A is the matrix whose colums are r times the corresponding columns of A.

Proposition 2.1. Properties of Matrix Addition and Scalar Multiplication

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $A + (-A) = 0$
5. $k(A + B) = kA + kB$
6. $(k + l)A = kA + lA$
7. $k(lA) = (kl)A$
8. $1A = A$

Proof. To prove Prop. 2.1, just consider the definition of matrix addition and scalar multiplication, and apply the Principles of numeral Addition and Multiplication.

Matrix Multiplication

Given two matrices A and B, for any vector $\mathbf{x}$, we want to make sure that $(AB)\mathbf{x} = A(B\mathbf{x})$. Starting from this point of view, we can define the matrix multiplication. Justification of this defination remains to be solved.

Definition 2.1. Matrix Multiplication
Let A be an $m \times n$ matrix and B be an $n \times p$ matrix, we now define that

$$AB = C = (A\mathbf{b_1}, \cdots, A\mathbf{b_p}).$$

Remark: The product AB is defined if and only if the number of columns of A is equal to the number of rows of B, and the size of the product is $m \times p$.

Example 2.3. Find the product of the AB, where

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

Solution: To avoid making mistaks or losing points, follow strictly this defination and write in steps, or write down the calculation details of each elements.

Note that this defination actually uses the defination of matrix-vector multiplication, which is defined as the linear combination of the columns of the matrix. In many textbooks, the defination of matrix multiplication is defined in the following way, which is equivalent to the defination above and we regard it as Row-Column Rule for Matrix Multiplication in the context of MATH2111. In general, we have the following properties of matrix multiplication.

Proposition 2.2. Row-Column Rule for Matrix Multiplication
Let A be an $m \times n$ matrix and B be an $n \times p$ matrix, we now define that the product AB is the $m \times p$ matrix whose (i, j)-entry is the dot product of the i-th row of A and the j-th column of B, that is,

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

where $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, B = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix}$.

Proposition 2.3. Properties of Matrix Multiplication

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $k(AB) = (kA)B = A(kB)$
5. $I_m A = A = AI_n$

Warning:
(1) In general, $AB \neq BA$, that is, the multiplication of matrices is not communtive. Moreover, the product of two matrices may not be defined, that is, AB may not be defined even if BA is defined.
(2) In general, $AB = AC$ does not imply $B = C$.
(3) In general, $AB = 0$ does not imply $A = 0$ or $B = 0$.

Powers and Transpose of a Matrix

Definition 2.2. Powers of a Matrix
Let A be a $n \times n$ matrix, then we define the power of A as follows:

$$\begin{aligned} A^0 &= I_n \\ A^1 &= A \\ A^2 &= A \cdot A \\ A^3 &= A \cdot A \cdot A \\ &\cdots \end{aligned}$$

Definition 2.3. Transpose of a Matrix
Let A be a $m \times n$ matrix, then the transpose of A, denoted by A^T, is the $n \times m$ matrix whose (i, j)-entry is the (j, i)-entry of A, that is,

$$(A^T)_{ij} = A_{ji}.$$

Example 2.4. Find the transpose of the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.

Solution: We have $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Proposition 2.4. Properties of Transpose of a Matrix Let A and B be matrices whose sizes are appropriate for the following operations, and let k be a scalar, then we have

1. $(A^T)^T = A$
2. $(kA)^T = kA^T$
3. $(A + B)^T = A^T + B^T$
4. $(AB)^T = B^T A^T$
5. $(A^k)^T = (A^T)^k$

Proof. For Prop. 2.4.4, we have $(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$, then

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} b_{ki} a_{jk} = (B^T A^T)_{ij}.$$

Note that in Prop. 2.4.4, $A^T B^T$ may not even be defined if A and B are not square matrices.

2.2 Inverses of Matrices and Elementary Matrices

Inverses of Matrices

In many situations, we need to solve the equation $A\mathbf{x} = B$, where A is a matrix and $\mathbf{x}$ is a vector. For numeral calculation, it's natural to find the solution by multiplying both sides by A^{-1} to get $\mathbf{x} = A^{-1}B$. In matrices, in order to solve this equation, we introduce the concept of the inverse of a matrix.

Definition 2.4. Identity Matrices
The $n \times n$ identity matrix, denoted by I_n, is the matrix whose (i,j)-entry is 1 if $i = j$ and 0 otherwise, that is,

$$(I_n)_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Definition 2.5. Inverses of Matrices
Let A be a square matrix. If there exists a matrix B such that

$$AB = BA = I_n$$

, then B is called the inverse of A, denoted by A^{-1}.
Otherwise, if A does not have an inverse, then A is called singular.

Note that the inverse of a matrix is unique if it exists, which can be proved by contradiction. Assume that B and C are both inverses of A, then we have

$$B = BI_n = B(AC) = (BA)C = I_nC = C,$$

which shows that $B = C$. To find the unique inverse of a matrix, we can use the following method.

Example 2.5. Find the inverse of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Solution: We have $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, where $a = 1$, $b = 2$, $c = 3$, $d = 4$. Thus, $A^{-1} = \frac{1}{1\cdot 4 - 2\cdot 3}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix}$.

Theorem. From this example, we can derive a general formula for the inverse of a 2×2 matrix. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then we have

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

If $ad - bc = 0$, then A is singular and does not have an inverse. In general, we can use the Gauss-Jordan method to find the inverse of a matrix, which will be explained later. No matter whether a matrix is 2×2 or not, however, we have the following properties covered in the lectures.

Proposition 2.5. Properties of Inverses of Matrices Let A and B be invertible matrices of the same size, and let k be a scalar, then we have

1. A^{-1} is invertible, and $(A^{-1})^{-1} = A$
2. $(kA)^{-1} = \frac{1}{k}A^{-1}$
3. $(A^T)^{-1} = (A^{-1})^T$
4. $(AB)^{-1} = B^{-1}A^{-1}$

Proof. For Prop. 2.5.3, just consider using the Properties of Transpose. For Prop. 2.5.4, we have

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

In the end, let's recall the problem of solving the equation $A\mathbf{x} = B$. With the concept of invertible matrices, we can now give a general solution. If A is invertible, then we can multiply both sides by A^{-1} to get $\mathbf{x} = A^{-1}B$. If $A\sigma$ is not invertible, then we can use the Gauss-Jordan method to solve the equation (which will be presented right after this part).

Proposition 2.6. Solution of the Equation $A\mathbf{x} = \mathbf{b}$
If A is an invertible matrix, then the equation $A\mathbf{x} = \mathbf{b}$ has **unique** solution $\mathbf{x} = A^{-1}\mathbf{b}$

Elementry Matrices

Recall the elementary row operations in Chapter 1, which are the following three operations: (1) Interchange two rows, (2) Multiply a row by a nonzero scalar, (3) Add a multiple of one row to another row. If an elementary row operation is performed on the $m \times n$ matrix A, the resulting matrix can be written as EA, where E is the $m \times m$ elementary matrix created by performing the elementary row operation on the $m \times m$ identity matrix I_m.

Each elementary matrix E is invertible, and its inverse is also an elementary matrix. The inverse of E is the elementary matrix of the same type that transforms E back to I_m.

Proposition 2.7. An $n \times n$ matrix A is invertible if and only if A is row equivalent to the $n \times n$ identity matrix I_n. In this case, any sequence of elementary row operations that reduces A to I_n will also transform I_n into A^{-1}.

Proof. All the following statements are equivalent:

$$\begin{aligned}
&A \text{ is invertible}\\
\Leftrightarrow &\text{The equation } A\mathbf{x} = \mathbf{b} \text{ has a unique solution for all vectors } \mathbf{b}\\
\Leftrightarrow &\text{The equation } A\mathbf{x} = \mathbf{0} \text{ has only the trivial solution}\\
\Leftrightarrow &\text{The reduced row echelon form of } A \text{ is } I_n \quad (A \sim I_n).
\end{aligned}$$

Conversly, if $A \sim I_n$, each row operation corresponds to multiplying A by an elementary matrix, so $E_p \cdots E_2 E_1 A = I_n$, where $E_1, E_2, \cdots, E_p$ are elementary matrices. Thus we have

$$A^{-1} = E_p \cdots E_2 E_1.$$

Proposition 2.8. An algorithm for finding the inverse of a matrix (Gauss-Jordan method)
Row operation reduce the matrix $\begin{pmatrix} A & I_n \end{pmatrix}$. If A is row equivalent to I_n, then $\begin{pmatrix} A & I_n \end{pmatrix}$ is row equivalent to $\begin{pmatrix} I_n & A^{-1} \end{pmatrix}$. Otherwise, A is not invertible.

Example 2.6. Find the inverse of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.
Solution: We can use the following method to find the inverse of a matrix.
Step 1: Write down the augmented matrix $\begin{pmatrix} A & I_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix}$.
Step 2: Use the row operations to transform the left side of the augmented matrix into the identity matrix: $\begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{pmatrix}$.
Step 3: The right side of the augmented matrix is the inverse of the matrix A.

Proof. $E_p \cdots E_1 \begin{pmatrix} A & I_n \end{pmatrix} = \begin{pmatrix} E_p \cdots E_1 A & E_p \cdots E_1 I_n \end{pmatrix} = \begin{pmatrix} I_n & A^{-1} \end{pmatrix}$

2.3 Characterization of Invertible Matrices

Proposition 2.9. Let A be an $n \times n$ matrix. The following statements are equivalent:

1. A is invertible.
2. A is row equivalent to I_n.
3. A has n pivot positions.
4. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
5. The equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for all vectors $\mathbf{b}$.
6. The columns of A form a linearly independent set.
7. The columns of A span $\mathbb{R}^n$.
8. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
9. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.
10. There exists an $n \times n$ matrix C such that $CA = I_n$.
11. There exists an $n \times n$ matrix D such that $AD = I_n$.
12. A^T is invertible.

Proof. The equivalence of the statements can be proved by the following chain of implications:

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (6) \Rightarrow (7) \Rightarrow (8) \Rightarrow (9) \Rightarrow (10) \Rightarrow (11) \Rightarrow (12) \Rightarrow (1)$$
$$(1) \Rightarrow (12) \Rightarrow (11) \Rightarrow (10) \Rightarrow (9) \Rightarrow (8) \Rightarrow (7) \Rightarrow (6) \Rightarrow (5) \Rightarrow (4) \Rightarrow (3) \Rightarrow (2) \Rightarrow (1)$$

Invertible Linear transformations

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is invertible if and only if there exists a map $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $T \circ S(\mathbf{x}) = S \circ T(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. In this case, S is called the inverse of T, denoted by T^{-1}.

Let A be the standard matrix of the linear transformation T, then T is invertible if and only if A is invertible. In this case, the standard matrix of T^{-1} is A^{-1}.

To bring this chapter to an end, we show the following proposition given by *Github Copilot*.

Proposition 2.10. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. The following statements are equivalent:

1. T is invertible.
2. T is one-to-one.
3. T is onto.

3 Determinants

3.1 Introduction to Determinants

Definition 3.1. Determinant of a Matrix
Let A be a 2×2 matrix, then the determinant of A, denoted by $\det(A)$ or $|A|$, is defined as the following

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

In general, the determinant of an $n \times n$ matrix A, and A_{ij} is the $(n-1) \times (n-1)$ matrix obtained by deleting the i-th row and j-th column of A, is defined as $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$.
Cofactor
The (i, j)-cofactor of A is defined as $C_{ij} = (-1)^{i+j} \det(A_{ij})$.

The defination of determinant is a recursive defination, which is based on the defination of determinant of 2×2 matrix. The determinant of a matrix can be calculated by the following formula:

Proposition 3.1. Calculation of Determinant
Let A be an $n \times n$ matrix, then the determinant of A can be computed by a cofactor expansion across any row or down any column, that is,

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} a_{ij} C_{ij} \text{ (expansion across the } i\text{-th row)} \\ &= \sum_{i=1}^{n} a_{ij} C_{ij} \text{ (expansion down the } j\text{-th column).} \end{aligned}$$

Thm for triangular matrix
If A is a triangular matrix, then $\det(A)$ is the product of the main diagonal entries of A.

$$\begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix} \text{ or } \begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ * & * & 0 & \cdots & 0 \\ * & * & * & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & * \end{pmatrix}$$

Warning: Mind the sign of the cofactor expansion - indexes changed after each recursive call !!

3.2 Properties of Determinants

Proposition 3.2. Row Operations
Let A be an $n \times n$ matrix.

- If a multiple of one row of A is added to another row to produce a matrix B, then $\det(B) = \det(A)$.
- If two rows of A are interchanged to produce a matrix B, then $\det(B) = -\det(A)$.
- If one row of A is multiplied by k to produce a matrix B, then $\det(B) = k \det(A)$.