

MATH 2101 Linear Algebra I–System of Linear Equations II

Row Echelon form

We first consider some examples.

Example

$$x_1 + 2x_3 = 1$$

$$x_2 + 3x_3 = 2$$

We can set t as a variable. Set $x_1 = 1 - 2t$, $x_2 = 2 - 3t$. The solution set is given by:

$$\{(1 - 2t, 2 - 3t, t) : t \in \mathbb{R}\}.$$

The above system of linear equations is *quite simple* since one can almost read off the solution. Our goal is to 'transform' the system of linear equations to such simple forms.

Row echelon form

We first introduce a notion involving in such form:

Definition

A matrix is said to be in a **reduced row echelon form** if the following conditions are satisfied:

- The **first nonzero entry in each row** is the only non-zero entry in its column.
- The rows with all zero entries are in the bottom of the matrices (if any).
- The first nonzero entry in each row is **1** and it occurs in a column to the right of the first nonzero entry in the preceding row.

The **first non-zero entry** in each row is called a **leading entry**. Those 1 in the leading entry is called a **leading one**.

Summary: The first bullet defines leading entries of a matrix and describes the column with a leading entry. The second and third bullets describe the distribution of leading entries.

Examples

Example

The matrices $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ are in reduced row echelon forms.

Remark

We rephrase some main features of a matrix in reduced row echelon form:

- ① The leading one is the first nonzero entry in a row.
- ② The leading ones come from left to right columns when going the rows from top to bottom.
- ③ The leading ones are in different rows.
- ④ For a column with a leading one, the leading entry is the only non-zero entry in the column.
- ⑤ Rows without a leading entry is the row with all entries to be zero.

Examples

Example

(Non-examples)

- $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ is not in a reduced row echelon form since the row of all zeros is in the second one, but the third row is not of all zeros.
- $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is not a reduced row echelon form since the first non-zero entries in the first and second rows are in the same column.
- $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is not a reduced row echelon form since the leading entries do not come from left to right when going from the first row to the second row.

2×2 matrices

Example

2×2 matrices in reduced row echelon form:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example

The following matrices are not in reduced row echelon forms:

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

Exercise

Write down all the reduced row echelon forms for 3×2 matrices.

Reduced row echelon form for an invertible matrix

Theorem

Let A be an $n \times n$ matrix in a reduced row echelon form. Then the following statements are equivalent:

- ① A has n leading ones;
- ② A is invertible;
- ③ $A = I_n$.

Proof.

(3) \Rightarrow (1) and (3) \Leftarrow (2) are clear. For (2) \Rightarrow (1), $\det(A) \neq 0$ and so A has no zero rows and so there are n leading ones.

It suffices to show that (1) \Rightarrow (3).

Suppose (1) holds. Since leading ones appear in different rows, the **n leading entries** have to be in each of the **n rows**. Since leading entries come from left to right, it must have to be in the diagonals. Moreover, in each column, the leading entry is the only non-zero entry. We must then have that $A = I_n$. This shows (3). □

Jordan-Gaussian eliminations

- The Jordan-Gaussian elimination is the process that transforms a $n \times m$ matrix to a reduced row echelon form by a sequence of elementary row operations.
- More precisely, given a system of linear equations $Ax = b$, we represent the system of linear equations by the augmented matrix $(A|b)$. If we carry out the elementary row operation on $(A|b)$, we obtain $E(A|b) = (EA|Eb)$ for some elementary matrix E . Then the corresponding system of linear equation is $EAx = Eb$. Since E is invertible, $Ax = b$ and $EAx = Eb$ has the same solution set.

Exercise

Show that $E(A|b) = (EA|Eb)$. (Here $E(A|b)$, EA and Eb are the matrix multiplications.)

Jordan-Gaussian eliminations

Suppose we have to solve the system of linear equations:

$$\begin{aligned}x_3 &= 2 \\2x_1 + 2x_2 + 4x_3 &= 6 \\x_1 + x_3 &= 1\end{aligned}$$

We write the augmented matrix:

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

Jordan-Gaussian elimination

- (1) Find the **leftmost non-zero column**. By interchanging rows (Type I), make the first row to contain a **non-zero entry** in that column e.g.

$$\left(\begin{array}{ccc|c} 0 & 0 & 1 & 2 \\ \textcolor{red}{2} & 2 & 4 & 6 \\ 1 & 0 & 1 & 1 \end{array} \right) \xrightarrow{2 \leftrightarrow 1} \left(\begin{array}{ccc|c} \textcolor{blue}{2} & 2 & 4 & 6 \\ 0 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{array} \right)$$

- (2) Change **that entry** to **1** by a scalar multiplication on the first row (Type II) e.g.

$$\left(\begin{array}{ccc|c} \textcolor{blue}{2} & 2 & 4 & 6 \\ 0 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{array} \right) \xrightarrow{1 \times (\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 \end{array} \right)$$

- (3) Add the scalar multiple of the first row to other rows to make the **other entries in that column to be zero** (Type III).

$$\left(\begin{array}{ccc|c} \textcolor{red}{1} & 1 & 2 & 3 \\ \textcolor{red}{0} & 0 & 1 & 2 \\ \textcolor{red}{1} & 0 & 1 & 1 \end{array} \right) \xrightarrow{1 \times (-1) + 3} \left(\begin{array}{ccc|c} \textcolor{red}{1} & 1 & 2 & 3 \\ \textcolor{red}{0} & 0 & 1 & 2 \\ \textcolor{red}{0} & -1 & -1 & -2 \end{array} \right)$$

Jordan-Gaussian eliminations

- (4) Repeat a similar process for the second row until all the entries in remaining rows are zero.

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & -1 & -1 & -2 \end{array} \right) \xrightarrow[2\leftrightarrow 3]{2\leftrightarrow 3} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & 2 \end{array} \right) \xrightarrow[2\times(-1)]{2\times(-1)} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

- (5) Start from the last row with non-zero entries, we add a scalar multiple to the row above to creates zero above that 1.

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right) \xrightarrow[3\times(-1)+2]{3\times(-1)+2} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array} \right) \xrightarrow[3\times(-2)+1]{3\times(-2)+1} \left(\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

- (6) Repeat the process for preceding rows.

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array} \right) \xrightarrow[2\times(-1)+1]{2\times(-1)+1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

- (7) The solution is $x_1 = -1, x_2 = 0, x_3 = 2$.

Leading variables and free variables

Definition

Given a system of linear equations $Ax = b$, we obtain an augmented matrix $(A|b)$. By the Gaussian elimination, we obtain a **reduced row echelon form**:

$$(A'|c').$$

A variable corresponding to a leading entry (of the reduced row echelon form) is called a **leading variable**. A variable which is not a leading variable is called a **free variable**. The solution (if any) for $Ax = b$ is given by setting free variables to be some arbitrary variables and then solve for the leading variables.

Example

For example, if the reduced row echelon form takes:

$$\left(\begin{array}{ccccc|c} 0 & 1 & 2 & 0 & 3 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{array} \right).$$

Then, x_2, x_4 are leading variables and x_1, x_3, x_5 are free variables. Set $x_1 = s, x_3 = t, x_5 = u$. We have $x_2 = 2 - 2t - 3u$ and $x_4 = 3 - 2u$.

More example on Jordan-Gaussian eliminations

Example

Solve the following system of linear equations:

$$x_1 + 2x_2 - x_3 + 4x_4 = -1$$

$$2x_1 + 4x_2 - x_3 + 8x_4 = -3$$

$$x_2 + x_4 = 2$$

The augmented matrix takes the form:

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & -1 \\ 2 & 4 & -1 & 8 & -3 \\ 0 & 1 & 0 & 1 & 2 \end{array} \right)$$

Example (con'd)

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & -1 \\ 2 & 4 & -1 & 8 & -3 \\ 0 & 1 & 0 & 1 & 2 \end{array} \right) \xrightarrow[1 \times (-2) + 2]{\rightarrow} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 2 \end{array} \right)$$
$$\xrightarrow[2 \leftrightarrow 3]{\rightarrow} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 4 & -1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 \end{array} \right)$$
$$\xrightarrow[3 \times (-1) + 1]{\rightarrow} \left(\begin{array}{cccc|c} 1 & 2 & 0 & 4 & -2 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 \end{array} \right) \xrightarrow[2 \times (-2) + 1]{\rightarrow} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 2 & -6 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 \end{array} \right)$$

Hence, x_4 is a free variable and we set $x_4 = t$. Then, the solution set is

$$\{(-2t - 6, -t + 2, -1) : t \in \mathbb{R}\}.$$

Another viewpoint on Jordan-Gaussian eliminations

To summarize again, the Jordan-Gaussian elimination is a process of taking a matrix to a reduced row echelon form. Another way to see the process is to reduce the redundancy of the system as well as the complexity of the system. For example, one may have some trivial '**redundant equations**' like:

$$x_1 + x_2 = 1 \quad (1)$$

$$2x_1 + 2x_2 = 2 \quad (2)$$

$$3x_1 + 3x_2 = 3 \quad (3)$$

The second and third equations are just essentially differed from the first one by a scalar. The elimination process **reduces three equations into solving the first one**. Of course, the elimination process can handle much more complicated situation than the above one. We now have a closer look on some special situations: systems of no solutions, systems of infinite solutions.

System of linear equations that has no solutions

Example

The equation $0x_1 + 0x_2 = 1$ has no solution!

Example

The system of linear equations:

$$2x_1 + x_2 = 1$$

$$4x_1 + 2x_2 = 3$$

also has no solution!

System of linear equations that has no solutions

Definition

A system of linear equations is said to be **consistent** if the solution set to that system is non-empty. Otherwise, we say that the system of linear equations is **inconsistent**.

Definition

A matrix B is said to be a **reduced row echelon form of a matrix A** if B is obtained by the Jordan-Gaussian elimination from A .

Theorem

*A system of linear equations $Ax = b$ is **inconsistent** if and only if the reduced row echelon form of the augmented matrix $(A|b)$ has a **row which has only one non-zero entry in the last column**.*

Example

If the reduced row echelon form of $(A|b)$ takes the form

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right),$$

then the system $Ax = b$ is inconsistent. (Actually, if one writes the system of linear equations for the reduced row echelon form, it takes

$$x_1 + 0x_2 + x_3 = 0$$

$$0x_1 + 0x_2 + 0x_3 = 1$$

More example

Example

Solve the following system of linear equations:

$$x_1 - 2x_2 - x_3 = -2$$

$$2x_1 - 3x_2 + x_3 = 0$$

$$3x_1 - 4x_2 - 2x_3 = -6$$

$$x_1 + 5x_3 = 4$$

We first obtain an augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 2 & -3 & 1 & 0 \\ 3 & -4 & -2 & -6 \\ 1 & 0 & 5 & 4 \end{array} \right)$$

Example (con'd)

This gives that

$$\begin{array}{ccc|c} \left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 2 & -3 & 1 & 0 \\ 3 & -4 & -2 & -6 \\ 1 & 0 & 5 & 4 \end{array} \right) & \xrightarrow{1 \times (-2) + 2} & \left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & 4 \\ 3 & -4 & -2 & -3 \\ 1 & 0 & 5 & 4 \end{array} \right) \\ \xrightarrow{1 \times (-3) + 3} & & \left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & 4 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 5 & 4 \end{array} \right) & \xrightarrow{1 \times (-1) + 4} \left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & 4 \\ 0 & 2 & 1 & 3 \\ 0 & 2 & 6 & 6 \end{array} \right) \\ \xrightarrow{2 \times (-2) + 3} & & \left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & -5 & -5 \\ 0 & 2 & 6 & 6 \end{array} \right) & \xrightarrow{2 \times (-2) + 4} \left(\begin{array}{ccc|c} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & -2 \end{array} \right) \end{array}$$

From this, we know the reduced row echelon form has a row with only non-zero in the last column. Hence, the system is inconsistent.

Infinitely many solutions

Definition

Let A be an $n \times m$ matrix. A system of linear equations $Ax = b$ is called **homogeneous** if $b = 0$. A homogeneous solution always has a solution, that is

$$x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Such solution is called the **trivial** solution.

Infinitely many solutions

Theorem

Let $Ax = 0$ be a homogeneous system of n linear equations and m variables.
If $m > n$, then there is an infinite number of solutions.

Proof.

We consider the augmented matrix of the form $(A \mid 0_{n \times 1})$. By carrying out the Jordan-Gaussian elimination, we obtain a reduced row echelon form, say $(A' \mid 0_{n \times 1})$. Since $m > n$, we must have some columns which do not contain a leading entry. In other words, the **reduced row echelon form has some free variables**. Those free variables constitute an infinite number of solutions. □

Example

The equation $2x_1 + 3x_2 - x_3 = 0$ has the solution set:

$$\left\{ \left(-\frac{3}{2}s + \frac{1}{2}t, s, t \right)^T : s, t \in \mathbb{R} \right\}.$$

In general, we have the following:

Theorem

Let $Ax = b$ be a system of n linear equations with m variables. Suppose $m > n$.

- The system has a solution, then there is an infinite number of solutions for the system.
- Moreover, let $x_p \in \mathbb{R}^m$ satisfy $Ax_p = b$. The solution set for the system $Ax = b$ is

$$\{x_p + y \in \mathbb{R}^m : Ay = 0\}.$$

Example

Let A be an $n \times 4$ matrix. Suppose the solution set of the system of linear equations $Ax = 0$ is given by $\{(2t, s, 3t + 2s, t - s)^T : t, s \in \mathbb{R}\}$. If $(0, 1, 0, 0)^T$ is a solution to $Ax = b$, find the solution set of $Ax = b$.

Solution: The solution set is

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2t \\ s \\ 3t + 2s \\ t - s \end{pmatrix} : t, s \in \mathbb{R} \right\}.$$

Proof of the theorem

Theorem

Let $Ax = b$ be a system of n linear equations with m variables. Suppose $m > n$.

- The system has a solution, then there is an infinite number of solutions for the system.
- Moreover, let x_p satisfy $Ax_p = b$. The solution set for the system $Ax = b$ is

$$\{x_p + y \in \mathbb{R}^m : Ay = 0\}.$$

Proof.

The proof for the first bullet is similar to that of the previous theorem and we leave as an exercise.

We only consider the second bullet. If $x' = x_p + y$ for some $Ay = 0$, we have that $Ax' = A(x_p + y) = Ax_p + Ay = b$. Hence, $x_p + y$ is a solution for $Ax = b$. Conversely, if $Ax' = b$, then $A(x' - x_p) = b - b = 0$. Hence,

$$x' = x_p + (x' - x_p) \in \{x_p + y \in \mathbb{R}^m : Ay = 0\}.$$



Invertibility and solution set

Theorem

Let A be an $n \times n$ matrix. Then the system $Ax = 0$ has a non-trivial solution if and only if A is not invertible.

Proof.

Suppose A is invertible. By taking the inverse, $x = 0$ is the only solution.
If A is not invertible, then the reduced row echelon form of A cannot be I_n . Thus the reduced row echelon form of $(A|0)$ has less leading entries than the number of columns. In other words, the system has free variables. Hence, the solution has non-trivial solution. □

Indeed, the proof of the previous theorem also gives:

Theorem

Let A be an $n \times n$ matrix. Then the following statements are equivalent:

- (a) *the system $Ax = 0$ has a non-trivial solution;*
- (b) *the system $Ax = 0$ has infinitely many solutions;*
- (c) *A is not invertible.*

Example

Example

Determine if the following system of linear equations has infinitely many solutions:

$$3x_1 - 4x_2 + 5x_3 = 0$$

$$2x_1 - x_2 + 4x_3 = 0$$

$$-x_1 - x_2 + 2x_3 = 0$$

MATH 2101 Linear Algebra I–Vector Spaces II (Linear combinations and vector subspaces)

Some motivating questions

Now the question is how to *produce* some vector subspaces from a vector space.

Example

If we start a *fixed* vector $v \in \mathbb{R}^n$, we can form a vector subspace of \mathbb{R}^n by considering the set:

$$\{cv : c \in \mathbb{R}\}.$$

For example, let $v = (3, 4, 5)$ and then $\{(3c, 4c, 5c) : c \in \mathbb{R}\}$ is a vector subspace.

Example

How about if we add one more vector $(2, -1, 3)$? How to form a vector subspace of \mathbb{R}^3 on that.

Linear combinations

Definition

Let S be a non-empty subset of V . A vector v in V is called a **linear combination** of vectors in S if there exists a finite number of vectors u_1, \dots, u_n in S and scalars a_1, \dots, a_n in \mathbb{R} such that

$$v = a_1u_1 + \dots + a_nu_n$$

We also say that v is a linear combination of vectors u_1, \dots, u_n .

Example

Example

Express the vector $(1, 3, 2)$ in linear combinations of $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$.

Solution: $(1, 3, 2) = (1, 0, 0) + 3(0, 1, 0) + 2(0, 0, 1)$.

Example

Example

Express the vector $(1, 3)$ in linear combinations of $(2, 3)$ and $(3, 5)$.

Solution: We have to solve $x_1, x_2 \in \mathbb{R}$ for $(1, 3) = x_1(2, 3) + x_2(3, 5)$ i.e.

$$\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Let $A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$. Then $A^{-1} = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}$. We then have:

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \end{pmatrix}.$$

Determining linear combinations=solving a system of linear equations

Remark

Determining questions about linear combinations of vectors is usually done by transferring to a problem of solving system of linear equations and then solving the system.

More examples

Example

Let $u_1 = (1, 1, 0, 0)$, $u_2 = (0, 1, 1, 0)$, $u_3 = (0, 0, 1, 1)$. Show that $(1, 2, 3, 4)$ cannot be written as a linear combination of u_1 , u_2 and u_3 .

Solution: We have to show the system of linear equations: $x_1u_1 + x_2u_2 + x_3u_3 = (1, 2, 3, 4)$ does not have a solution. We write the augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{array} \right)$$

By the Jordan-Gaussian eliminations, the row echelon form takes the form:

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

The last row implies there is no solution.

More examples

Example

Express $(0, -1, 4)$ as a linear combination of $(1, 2, -1)$ and $(2, 3, 2)$.

Solution: We have to solve x_1, x_2 such that

$$x_1(1, 2, -1) + x_2(2, 3, 2) = (0, -1, 4).$$

Then, we have to solve the system of linear equations with augmented matrix

$$\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 3 & -1 \\ -1 & 2 & 4 \end{array} \right).$$

By the Jordan-Gaussian eliminations (exercise: fill out the details),

$$\left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right).$$

Hence, $x_1 = -2$ and $x_2 = 1$. Then

$$(0, -1, 4) = -2(1, 2, -1) + (2, 3, 2).$$

Spanning sets

We now use the concept of linear combinations to define some vector subspaces:

Definition

Let S be a non-empty subset of a vector space V . The *span* of S , denoted by $\text{span}(S)$, is the set consisting of all linear combinations of the vectors in S . For convenience, we define $\text{span}(\emptyset) = \{0\}$.

Examples of spanning sets

Example

Let $S = \{(1, 0, 0), (0, 1, 0)\}$ be a subset in \mathbb{R}^3 . Then

$$\text{span}(S) = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}.$$

Example

Let v be a non-zero vector in \mathbb{R}^n and let $S = \{v\}$. Then

$$\text{span}(S) = \{av : a \in \mathbb{R}\}$$

Geometrically, $\text{span}(S)$ represents a line in \mathbb{R}^n .

Theorem

The span of any subset S of V is a vector subspace of V .

Proof.

When $S = \emptyset$, it follows from that $\text{span}(S)$ is a zero subspace.

Now assume $S \neq \emptyset$. For any $v, u \in \text{span}(S)$, there exists $a_1, \dots, a_r \in \mathbb{R}$ and $v_1, \dots, v_r \in S$ such that

$$v = a_1v_1 + \dots + a_rv_r,$$

and similarly there exists $b_1, \dots, b_s \in \mathbb{R}$ and $u_1, \dots, u_s \in S$ such that

$$u = b_1u_1 + \dots + b_su_s.$$

Hence,

$$v + u = a_1v_1 + \dots + a_rv_r + b_1u_1 + \dots + b_su_s$$

and so is a linear combination of vectors in S . Hence, $v + u \in \text{span}(S)$.

For $c \in \mathbb{R}$,

$$cv = c(a_1v_1) + \dots + c(a_rv_r) = (ca_1)v_1 + \dots + (ca_r)v_r.$$

Then, cv is also a linear combination of vectors in S and so is also in $\text{span}(S)$.



Exercise

Let A be an $n \times m$ matrix and let v_1, \dots, v_m be the columns of A . Show that

$$\text{span}(\{v_1, \dots, v_m\}) = \{Ax : x \in \mathbb{R}^m\}.$$

The former space is called **the column space** of A .

MATH 2101 Linear Algebra I–Vector Spaces III (Linear independence and basis)

Some ideas on linear independence

We have discussed the concept of a spanning set. In general, one may get a '**large' spanning set**' for a vector space V . For example, one may simply take $S = V$ (and so S is an infinite set), but one can ask if it is possible to find a smaller subset of V to span V . This will lead to a question on **whether a vector is in a linear combination of some others**. For example, suppose a set $\{v_1, v_2, v_1 + v_2\}$ spans a vector space V . Then, $v_1 + v_2$ is a linear combination of v_1 and v_2 , and so $\{v_1, v_2\}$ also spans V .

We shall introduce a concept of linearly independence to systematically address this.

Definition of linear dependence

Definition

A subset S of a vector space V is called *linearly dependent* if there exists a finite number of distinct vectors v_1, \dots, v_n in S and a_1, \dots, a_n in \mathbb{R} **not all zero** such that

$$a_1v_1 + \dots + a_nv_n = 0.$$

Roughly speaking, the linear independence on a set S of vectors in V is to *ask if one can do the addition and the scalar multiplication on some vectors in S to obtain some other vectors in S .*

Example

Let $u_1 = (1, 0, -1)$, $u_2 = (0, 1, -1)$ and $u_3 = (1, -1, 0)$. Then u_1, u_2, u_3 are linearly dependent since $u_1 + (-1)u_2 + (-1)u_3 = 0$.

Example

Let u, v be two non-zero vectors in \mathbb{R}^n . Then u and v are linearly dependent if and only if $u = av$ for some $a \in \mathbb{R}$.

Definition of linear independence

Definition

A subset S of a vector space V is said to be *linearly independent* if S is not linearly dependent.

We have the following logical reformulation of linear independence:

Theorem

Let $S = \{v_1, \dots, v_r\}$ be a finite subset in V . Then S is said to be linearly independent if the only $a_1, \dots, a_r \in \mathbb{R}$ satisfying

$$a_1v_1 + \dots + a_rv_r = 0$$

are $a_1 = \dots = a_r = 0$.

The above reformulation is more convenient if one thinks to solve some system of linear equations involving a_1, \dots, a_r .

Example

Example

Show that the vectors $(1, 0, 0, 2)$, $(0, 1, 0, 1)$, $(0, 0, 1, 0)$, $(0, 0, 1, 1)$ form a set of linearly independent vectors.

Solution: We have to solve x_1, x_2, x_3, x_4 such that

$$x_1(1, 0, 0, 2) + x_2(0, 1, 0, 1) + x_3(0, 0, 1, 0) + x_4(0, 0, 1, 1) = (0, 0, 0, 0).$$

The corresponding augmented matrix is:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 \end{array} \right)$$

The Gaussian eliminations (here I skip the details) take to the following reduced row echelon form:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right)$$

and so $x_1 = x_2 = x_3 = x_4 = 0$ is the only solution.

Example

Let V be the vector space of all functions from \mathbb{R} to \mathbb{R} . (Recall that we have defined the addition and scalar multiplication before.) Let $f_1(x) = x$ and $f_2(x) = x^2$. Show that $\{f_1, f_2\}$ is a linearly independent set.

Sol: Let $a_1, a_2 \in \mathbb{R}$ such that $a_1 f_1 + a_2 f_2 = 0$. (What is 0 in that vector space?)

- If we substitute $x = 1$, we then have $a_1 + a_2 = 0$.
- If we substitute $x = 2$, we then have $2a_1 + 4a_2 = 0$.

Solving the equations, we have $a_1 = a_2 = 0$. Hence, $\{f_1, f_2\}$ is a linearly independent set.

Relations of linear dependence among subsets

Theorem

Let $S_1 \subset S_2 \subset \mathbb{R}^n$.

- If S_2 is linearly independent, then S_1 is also linearly independent.
- If S_1 is linearly dependent, then S_2 is also linearly dependent.

Basis

Definition

A **basis** for a vector space V is a *linearly independent* subset of V which also *spans* for V .

One may think that a basis is a smallest spanning set for V . This is because if S is a basis for V , then $\text{span}(S') \neq V$ for any proper subset S' in S .

Examples

Example

Let $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, \dots , $\mathbf{e}_n = (0, \dots, 0, 1)$. The set $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis for V , which will be called the *standard basis* for \mathbb{R}^n .

Examples

Example

Let $v_1 = (1, 2)$ and $v_2 = (2, 3)$.

- We first show that $\{v_1, v_2\}$ is linearly independent: For any $a_1, a_2 \in \mathbb{R}$ with

$$a_1 v_1 + a_2 v_2 = (0, 0),$$

we have

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and so

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

- Show $\{v_1, v_2\}$ spans \mathbb{R}^2 : Exercise.

Examples

Example

Let $S = \{(1, 2, 0), (2, 3, 0)\}$. Note that S is a linearly independent set. Then S is a basis for $\text{span}(S)$.

Example

Let $S = \{(x_1, x_2, x_3, x_4) : x_1 + x_2 + x_3 + x_4 = 0\}$. Then, by solving the system of linear equations, we have:

$$\{(-s - t - u, s, t, u) : s, t, u \in \mathbb{R}\}.$$

Then a basis for S is

$$\{(1, -1, 0, 0), (1, 0, -1, 0), (1, 0, 0, -1)\}.$$

Unique linear combination from a basis

One important property for a basis is the following, which will be useful to convert things to matrices later:

Theorem

*Let V be a vector space and let $\beta = \{v_1, \dots, v_n\}$ be a subset of V . Then β is a basis for V if and only if for each vector $u \in V$, u can be **uniquely** written as a linear combination of vectors in β i.e.*

$$u = a_1v_1 + \dots + a_nv_n$$

for some unique scalars a_1, \dots, a_n in \mathbb{R} .

Proof of the theorem

Theorem

Let V be a vector space and let $\beta = \{v_1, \dots, v_n\}$ be a subset of V . Then β is a basis for V if and only if for each vector $u \in V$, u can be uniquely written as a linear combination of vectors in β i.e. $u = a_1v_1 + \dots + a_nv_n$ for some unique scalars a_1, \dots, a_n in \mathbb{R} .

Proof.

By the definition of a spanning set, there exist integers a_1, \dots, a_n in \mathbb{R} such that

$$u = a_1v_1 + \dots + a_nv_n.$$

It remains to show the uniqueness. Suppose $u = b_1v_1 + \dots + b_nv_n$ for another linear combination of vectors in β . Then,

$$a_1v_1 + \dots + a_nv_n = b_1v_1 + \dots + b_nv_n$$

and so

$$(a_1 - b_1)v_1 + \dots + (a_n - b_n)v_n = 0.$$

Since v_1, \dots, v_n are linearly independent, the only solution is

$$a_1 - b_1 = a_2 - b_2 = \dots = a_n - b_n = 0.$$

Hence, $a_1 = b_1, \dots, a_n = b_n$. This shows the uniqueness. □

Number of vectors in a basis

Theorem

Let β be a basis for a vector space V . Let n be the number of vectors in β . Let β' be a set of linearly independent vectors in V . Let m be the number of vectors in β' . Then $m \leq n$.

Proof.

Let $\beta = \{v_1, \dots, v_n\}$. Let $\beta' = \{u_1, \dots, u_m\}$. Assume $m > n$ to arrive at a contradiction. Since β is a basis for V , β particularly spans V . Hence, we can write as:

$$u_j = a_{1j}v_1 + \dots + a_{nj}v_n$$

Now, we consider the solution set (x_1, \dots, x_m) of the equation:

$$x_1u_1 + \dots + x_mu_m = 0$$

We can rewrite as:

$$\sum_{i=1}^n (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{im}x_m)v_i = 0$$

Since v_1, \dots, v_n are linearly independent, we obtain the following system of linear equations:

$$a_{11}x_1 + \dots + a_{1m}x_m = 0$$

⋮

$$a_{n1}x_1 + \dots + a_{nm}x_m = 0$$

Since $m > n$, the solution set for (x_1, \dots, x_m) is infinite. In particular, it contains a non-trivial solution. This contradicts that u_1, \dots, u_m are linearly independent. □

Corollary

Let β_1 and β_2 be any two bases for a vector space V . Then β_1 and β_2 contain the same number of vectors.

Definition

The **number of vectors** in a basis for a vector space V is called the *dimension*. The above discussion implies that the definition of the dimension is independent of a choice of a basis.

Example

The dimension of \mathbb{R}^n is n since we have the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

Example

The dimension of the solution set for the equation

$$x_1 + \dots + x_m = 0$$

is $m - 1$.

Subspaces and dimensions

Corollary

Let W be a vector subspace of a vector space V . Then

$$\dim(W) \leq \dim(V).$$

MATH 2101 Linear Algebra I–Vector Spaces I

Vector Spaces

Our goal now is to abstractize some important properties and natures of \mathbb{R}^n , which involves the addition of two vectors and the scalar multiplication on a vector.

Definition of a vector space

Definition

A vector space V over \mathbb{R} consists of a set on which two operations, called additions and scalar multiplications, are defined in the following ways:

- ① (addition) for $x, y \in V$, a unique element $x + y$ is defined in V :
- ② (scalar multiplication) for $x \in V$ and $a \in \mathbb{R}$, a unique element ax is defined in V .

such that

- ① For any $x, y \in V$, $x + y = y + x$ (commutativity of addition).
- ② For any $x, y, z \in V$, $(x + y) + z = x + (y + z)$. (associativity of addition)
- ③ There exists an element in V , denoted by 0 , such that $x + 0 = 0 + x = x$. (zero element in V)
- ④ For each $x \in V$, there exists an element $y \in V$ such that $x + y = 0$.
- ⑤ For any $x \in V$, $1x = x$.
- ⑥ For any $a, b \in \mathbb{R}$ and $x \in V$, $a(bx) = (ab)x$. (associativity of scalar multiplications)
- ⑦ For any $a, b \in \mathbb{R}$ and $x \in V$, $(a + b)x = ax + bx$.
- ⑧ For any $a \in \mathbb{R}$ and any $x, y \in V$, $a(x + y) = ax + ay$. (Distributivity for scalar multiplications)

We write $-v$ for $(-1)v$, the scalar multiple -1 on v .

Examples

Example

The Euclidean space \mathbb{R}^n is a vector space. The addition and the scalar multiplication are the usual ones: for $(x_1, \dots, x_n) \in \mathbb{R}^n$ and $c \in \mathbb{R}$,

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n),$$

$$c(x_1, \dots, x_n) = (cx_1, \dots, cx_n).$$

Example

The set of $n \times m$ matrices. The addition is the matrix addition. The scalar multiplication is also simply the one defined for matrices before.

Examples

Example

Let V be the set of all functions from \mathbb{R} to \mathbb{R} . The addition is defined as: for $f_1, f_2 \in V$, $(f_1 + f_2)(x) = f_1(x) + f_2(x)$ for $x \in \mathbb{R}$. The scalar multiplication is defined as: for $f \in V$ and $c \in \mathbb{R}$, $(cf)(x) = c(f(x))$ for $c \in \mathbb{R}$.

Vector subspaces

Definition

A non-empty subset S of a vector space V is called a *vector subspace* of V if it is closed under matrix addition and scalar multiplication i.e. for $v_1, v_2 \in S$,

- ① $v_1 + v_2 \in S$;
- ② for any $c \in \mathbb{R}$ and any $v \in S$, $cv \in S$.

The way of defining is to ensure that a vector subspace is *still* a vector space under the same addition and scalar multiplication.

Examples

Example

The set $\{(x, y, 0) : x \in \mathbb{R}, y \in \mathbb{R}\}$ is a vector subspace of \mathbb{R}^3 . Also visualize the vector subspace geometrically.

Example

Let v be a non-zero vector in \mathbb{R}^n . The set $\{cv : c \in \mathbb{R}\}$ is a subspace of \mathbb{R}^n .

Examples

Example

The set $\{(1, 2), (3, 4)\}$ with only two non-zero points is not a vector subspace of \mathbb{R}^2 . In general, a set with only finitely many **non-zero** points is not a vector subspace.

Example

The set of all diagonal $n \times n$ matrices is a vector subspace of the set of all $n \times n$ matrices.

Examples

Example

Determine if the following subsets are subspaces:

- ① $S_1 = \{(1, 2), (2, 3)\};$
- ② $S_2 = \{(2t, 3t, 4t) : t \in \mathbb{R}\};$
- ③ $S_3 = \mathbb{R}^3 \setminus \{(2t, 3t, 4t) : t \in \mathbb{R}\};$
- ④ $S_4 = \{(1 + 2t, 2t, 4t) : t \in \mathbb{R}\} \cup \{(0, 0, 0)\};$
- ⑤ $S_5 = \{(0, 0, 0, 0)\};$
- ⑥ $S_6 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^3 + x_2^3 + x_3^2 = 0\}.$

Example

- ① $S_1 = \{(1, 2), (2, 3)\}$; S_1 is not a subspace since $(0, 0) = 0(1, 2)$ is not in S_1 .
- ② $S_2 = \{(2t, 3t, 4t) : t \in \mathbb{R}\}$; S_2 is a subspace.
- ③ $S_3 = \mathbb{R}^3 \setminus \{(2t, 3t, 4t) : t \in \mathbb{R}\}$; S_3 is not a subspace since $(2, 0, 0), (0, 3, 4) \in S_3$ but $(2, 3, 4)$ is not in S_3 .
- ④ $S_4 = \{(1 + 2t, 2t, 4t) : t \in \mathbb{R}\} \cup \{(0, 0, 0)\}$; S_4 is not a subspace since $(1, 0, 0) \in S_4$, but $(2, 0, 0) \notin S_4$.
- ⑤ $S_5 = \{(0, 0, 0, 0)\}$; S_5 is a subspace, and is called **zero subspace**.
- ⑥ $S_6 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^3 + x_2^3 + x_3^2 = 0\}$; S_6 is not a subspace since $(-1, 0, 1), (0, -1, 1) \in S_6$ but $(-1, -1, 2) \in S_6$.

Vector subspaces in \mathbb{R}^2 and \mathbb{R}^3

We give a geometric description of all vectors subspaces in \mathbb{R}^2 and \mathbb{R}^3 :

Example

Any vector subspace of \mathbb{R}^2 is one of the followings:

- $\{0\}$;
- any line passing through the origin;
- \mathbb{R}^2 .

Example

Any vector subspace of \mathbb{R}^3 is one of the followings:

- $\{0\}$;
- any line passing through the origin;
- any plane passing through the origin;
- \mathbb{R}^3 .

Example 1: Homogeneous system of linear equations and vector spaces

We first recall the following definition.

Definition

Let A be a $n \times m$ matrix and let b be a column vector in \mathbb{R}^n . A system of linear equations $Ax = b$ of n equations and m variables is said to be *homogeneous* if $b = 0_{n \times 1}$.

Theorem

The solution set of a homogeneous system of linear equations is a vector subspace i.e. for an $n \times m$ matrix A , $\{x \in \mathbb{R}^m : Ax = 0\}$ is a vector subspace.

Example 1: Homogeneous system of linear equations and vector spaces

Theorem

The solution set of a homogeneous system of linear equations is a vector subspace of \mathbb{R}^m i.e. for an $n \times m$ matrix A , $\{x \in \mathbb{R}^m : Ax = 0\}$ is a vector subspace of \mathbb{R}^m .

Proof.

Let $S = \{x \in \mathbb{R}^m : Ax = 0\}$. For $v_1, v_2 \in S$, $Av_1 = 0$ and $Av_2 = 0$. Hence $A(v_1 + v_2) = Av_1 + Av_2 = 0 + 0 = 0$. Hence, $v_1 + v_2 \in S$. For $v \in S$ and $c \in \mathbb{R}$, $A(cv) = c(Av) = c(0) = 0$ and so $cv \in S$. □

Example 2: Multiplication with a matrix

Theorem

Let A be an $n \times m$ matrix. The set $S = \{Ax : x \in \mathbb{R}^m\}$ is a vector subspace of \mathbb{R}^n .

Proof.

(Addition) For $y_1, y_2 \in S$, $y_1 = Ax_1$ and $y_2 = Ax_2$ for some $x_1, x_2 \in \mathbb{R}^m$. Then $y_1 + y_2 = Ax_1 + Ax_2 = A(x_1 + x_2) \in S$.

(Scalar multiplication) For $y' \in S$ and $c \in \mathbb{R}$, write $y' = Ax'$ for some $x' \in \mathbb{R}^m$. Then $cy' = c(Ax') = A(cx') \in S$.

Hence, S is a vector subspace. □

Remark

The subspace is related to so-called the column space of a matrix, which will be discussed later.

MATH 2101 Linear Algebra I–System of Linear Equations III

Theoretical aspect: Inverse and Reduced row echelon form

In the last part of this section, we shall study more on reduced row echelon forms for a *square* matrix and related matters.

Theorem

Let A be an $n \times n$ matrix. Then A is invertible if and only if the reduced row echelon form of A is I_n .

Proof.

We carry out the **Jordan-Gaussian elimination** to transfer A to a reduced row echelon form B . Then $E_r \dots E_1 A = B$ for some elementary matrices E_1, \dots, E_r . Since E_1, \dots, E_r are invertible, A is invertible if and only if B is invertible. We have shown that the latter condition is equivalent to I_n . □

Theoretical aspect: Invertible matrix and elementary matrices

Corollary

Every invertible matrix is a product of elementary matrices.

Proof.

As shown in the previous proof,

$$E_r \dots E_1 A = I_n$$

for some elementary matrices E_1, \dots, E_r . Then $A = E_1^{-1} \dots E_r^{-1}$. We have shown before that the inverse of an elementary matrix is still elementary and thus we have shown the corollary.



Computational aspect: finding an inverse

We illustrate the idea on using elementary operations to find an inverse. Let A be an invertible $n \times n$ matrix. Then we form $n \times 2n$ augmented matrix:

$$(A \quad | \quad I_n)$$

Then, $A^{-1}(A \quad | \quad I_n) = (A^{-1}A \quad | \quad A^{-1}) = (I_n \quad | \quad A^{-1})$. Now the idea is to use elementary row operations/Jordan-Gaussian eliminations to change $(A \quad | \quad I_n)$ to $(I_n \quad | \quad A^{-1})$. Then we can **read A^{-1} from the product of elementary matrices.**

Example

Example

Find the inverse of $\begin{pmatrix} 1 & 3 \\ -2 & 4 \end{pmatrix}$. We compute in the following way:

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ -2 & 4 & 0 & 1 \end{array} \right) \xrightarrow{1 \times (2) + 2} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 10 & 2 & 1 \end{array} \right) \xrightarrow{2 \times (\frac{1}{10})} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & \frac{1}{5} & \frac{1}{10} \end{array} \right)$$
$$\xrightarrow{2 \times (-3) + 1} \left(\begin{array}{cc|cc} 1 & 0 & \frac{4}{10} & \frac{-3}{10} \\ 0 & 1 & \frac{2}{10} & \frac{1}{10} \end{array} \right)$$

Then $A^{-1} = \begin{pmatrix} \frac{4}{10} & \frac{-3}{10} \\ \frac{2}{10} & \frac{1}{10} \end{pmatrix}$. Moreover, we can write A^{-1} as a product of elementary matrices:

$$A^{-1} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{10} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

Example

Find the inverse of $\begin{pmatrix} 2 & 4 & 3 \\ 4 & 2 & 1 \\ 0 & 2 & 3 \end{pmatrix}$. We compute in the following way:

$$\begin{pmatrix} 2 & 4 & 3 & 1 & 0 & 0 \\ 4 & 2 & 1 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{1 \times \frac{1}{2}} \begin{pmatrix} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 4 & 2 & \frac{1}{2} & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{1 \times (-4) + 2} \begin{pmatrix} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -6 & -5 & -2 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{2 \times \frac{1}{6}} \begin{pmatrix} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{5}{6} & \frac{2}{6} & \frac{1}{6} & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{2 \times (-2) + 3} \begin{pmatrix} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{5}{6} & \frac{2}{6} & \frac{1}{6} & 0 \\ 0 & 0 & \frac{8}{6} & \frac{-4}{6} & \frac{2}{6} & 1 \end{pmatrix}$$

$$\xrightarrow{3 \times \frac{6}{8}} \begin{pmatrix} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{5}{6} & \frac{2}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \end{pmatrix} \xrightarrow{3 \times (\frac{-5}{6}) + 2} \begin{pmatrix} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{12} & \frac{1}{24} & -\frac{5}{8} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

$$\xrightarrow{3 \times (\frac{-3}{2}) + 1} \begin{pmatrix} 1 & 2 & 0 & \frac{5}{4} & \frac{-3}{8} & \frac{-9}{8} \\ 0 & 1 & 0 & \frac{1}{12} & \frac{1}{24} & -\frac{5}{8} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{4} & \frac{3}{4} \end{pmatrix} \xrightarrow{2 \times (-2) + 1} \begin{pmatrix} 1 & 0 & 0 & \frac{-1}{4} & \frac{9}{24} & \frac{1}{8} \\ 0 & 1 & 0 & \frac{9}{12} & \frac{-24}{24} & -\frac{5}{8} \\ 0 & 0 & 1 & \frac{-1}{2} & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

Hence, the inverse is

$$\begin{pmatrix} \frac{-1}{4} & \frac{9}{24} & \frac{1}{8} \\ \frac{9}{12} & \frac{-24}{24} & -\frac{5}{8} \\ \frac{-1}{2} & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

Exercise

Express the inverse in the previous example in terms of a product of elementary matrices.

MATH 2101 Linear Algebra I–System of Linear Equations I

System of linear equations

A **system of n linear equations in m variables** is a collection of equations:

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = b_2$$

⋮

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = b_n$$

We transform to a matrix form so that we can use *techniques of matrices*. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Then, we can rewrite as a matrix equation:

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (1)$$

We also refer such matrix equation to be a system of n linear equations in m -variables. A **solution** to (1) is a column vector s in \mathbb{R}^n such that $As = b$.

Method of using matrix inversions

Suppose now $m = n$ and A is invertible. Then, we multiply A^{-1} on both sides:

$$A^{-1}A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \implies \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

i.e. $A^{-1}b$ is the solution to the system of linear equations.

Example of using matrix inversions

Example

Solve the following system of linear equations by using A^{-1} :

$$x_1 + 2x_2 - x_3 = 5$$

$$x_1 + x_2 + x_3 = 2$$

$$2x_1 - 2x_2 + x_3 = 0$$

Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & -2 & 1 \end{pmatrix}$. Then we have

$$\det(A) = 9, \quad \text{adj}(A) = \begin{pmatrix} 3 & 1 & -4 \\ 0 & 3 & 6 \\ 3 & -2 & -1 \end{pmatrix}^T.$$

Hence, $A^{-1} = \frac{1}{9} \begin{pmatrix} 3 & 0 & 3 \\ 1 & 3 & -2 \\ -4 & 6 & -1 \end{pmatrix}$. Hence, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A^{-1} \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{15}{9} \\ \frac{11}{9} \\ -\frac{8}{9} \end{pmatrix}$.

A variation of matrix inversion method: Carmer's rule

Carmer's rule is a simpler formula that lightens some computations in matrix inversion:

Theorem

Let A be an $n \times n$ invertible matrix and let b be a column vector in \mathbb{R}^n .

Consider a system of n linear equations in n variables:

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = b.$$

Then the solution to the system is given by:

$$x_i = \frac{\det(A_{\textcolor{red}{i}})}{\det(A)},$$

where A_i is the matrix obtained by replacing the *i-th column* by b .

Proof of Cramer's rule

Theorem

The solution to the system $Ax = b$ is given by:

$$x_i = \frac{\det(A_{\textcolor{red}{i}})}{\det(A)},$$

where $A_{\textcolor{red}{i}}$ is the matrix obtained by replacing the *i-th column* by b .

Proof.

To use the formula:

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A),$$

we reduce computing $A^{-1}b$ to computing $\text{adj}(A)b$. Indeed,

$$(\text{adj}(A)b)_{i1} = (-1)^{i+1} b_1 \det(\tilde{A}_{1i}) + \dots + (-1)^{i+n} b_n \det(\tilde{A}_{ni}) = \det(A_i),$$

where the last equality follows from computing $\det(A_i)$ by using the *i-th column in the definition*.

Example of using Carmer's rule

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and let $b = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$. Solve $Ax = b$.

Solution: $\det(A) = -3$. We also have:

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

Hence, by Carmer's rule,

$$x = \frac{1}{-3} \begin{pmatrix} -3 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Remark on the method of matrix inversion or Carmer's rule

Remark

The matrix inversion or Carmer's rule to solve a system of linear equation only works when **the matrix is invertible**. For example, if $n \neq m$ for a $n \times m$ matrix A , then we cannot solve the equation by above method.

Elementary row operations: motivation

In order to deal with general system of linear equations, we have to develop some operations on and matrices. A basic idea is like **to subtract system of linear equations to get less variables in one equation** e.g. Solve

$$x_1 + x_2 = 1$$

$$x_1 + 2x_2 = 2$$

We subtract the second eqn. from the first one to get $x_2 = 1$ and then get $x_1 = 0$.

Elementary row operations

We shall carry out those operations on matrices:

Definition

Let A be an $n \times m$ matrix. Any one of the following three operations on **rows** (resp. columns) is called an **elementary row operation**:

- ① Type (I) **interchanging** any two rows (resp. columns) of A ;
- ② Type (II) **multiplying** any row (resp. column) of A **by a non-zero scalar**;
- ③ Type (III) **adding a scalar multiple of a row** (resp. column) of A to another row of A .

Examples of elementary row operations

Example

Let $A = \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 7 & 1 & 2 \\ 9 & 2 & 0 & 4 \end{pmatrix}$.

- ① (Type I) interchanging the first and third rows gives:

$$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 7 & 1 & 2 \\ 9 & 2 & 0 & 4 \end{pmatrix} \xrightarrow{1 \leftrightarrow 3} \begin{pmatrix} 9 & 2 & 0 & 4 \\ 0 & 7 & 1 & 2 \\ 1 & 2 & -1 & 5 \end{pmatrix}$$

- ② (Type II) multiplying the second row by $\sqrt{2}$ gives:

$$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 7 & 1 & 2 \\ 9 & 2 & 0 & 4 \end{pmatrix} \xrightarrow{2 \times (\sqrt{2})} \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 7\sqrt{2} & \sqrt{2} & 2\sqrt{2} \\ 9 & 2 & 0 & 4 \end{pmatrix}$$

Examples of elementary row operations

Example

(Type III) adding the scalar multiple 2 of the **first** row to the **third** row gives:

$$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 7 & 1 & 2 \\ 9 & 2 & 0 & 4 \end{pmatrix} \xrightarrow{1 \times (2) + 3} \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 7 & 1 & 2 \\ 11 & 6 & -2 & 14 \end{pmatrix}$$

Elementary matrices

We would like to convert those elementary row operations to *matrix multiplications* so that again we can use techniques of matrices. We first describe the matrices that we need.

Definition

An $n \times n$ **elementary matrix** is a matrix obtained by performing an elementary operation on I_n . An elementary matrix is called of type I, II or III according to its corresponding elementary operation.

Example

- ① $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ is an elementary matrix performing the **interchanging** the **first and third row**. (Type I)
- ② $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ is an elementary matrix performing the **scalar multiplication of 2 on the second row** (Type II).
- ③ $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$ is an elementary matrix performing **the addition of the scalar multiple -2 of the first row to the third row** (Type III).

Non-example of elementary matrices

Example

(Non-example) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ are not elementary.

Exercise

Write down the elementary matrices for the following elementary row operations on 4×4 matrices:

- Interchanging the first and fourth rows.
- Adding the second row to the third row.
- Scalar multiplication of 2 on the first row.

Matrix multiplications of elementary matrix= elementary row operations

Theorem

Let A be an $n \times m$ matrix.

- ① Suppose E is an elementary matrix. Then EA is the matrix obtained from A by the corresponding *elementary row operation*.
- ② Conversely, if B is a matrix obtained from A by an *elementary row operation*, then $B = EA$ for the elementary matrix corresponding to that operation.

Example

$$\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4(-2) & 5(-2) & 6(-2) \end{pmatrix}.$$

(LHS: matrix multiplication EA . RHS: row operations)

Example

Let $E = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, which corresponds to the row elementary operation for adding the scalar multiple 2 of the second row to the first row. Let $A = \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix}$. Then

$$EA = \begin{pmatrix} u_1 + 2v_1 & u_2 + 2v_2 & u_3 + 2v_3 \\ v_1 & v_2 & v_3 \end{pmatrix}$$

Invertibility of elementary matrices

Theorem

*Elementary matrices are **invertible**. Moreover, the inverse of an elementary matrix is still an elementary matrix of the same type.*

Proof.

We just illustrate the idea of the proof. **Finding an inverse is the same as finding an elementary operation to get back to the original matrix.**

- In type I, if the elementary operation interchanges the k -th row and the l -th row, then applying **the same elementary operation once more** (i.e. interchanging the k -th and the l -th rows again) gives back to the original matrix.
- In type II, if the elementary operation multiplies the k -th row by a scalar a , then the elementary operation **multiplying the k -th row by the scalar $\frac{1}{a}$** gives back to the original matrix.
- In type III, if the elementary operation adds the scalar multiple a of the k -th row to the l -th row, then the elementary operation for **adding the scalar multiple $-a$ of the k -th row to the l -th row** gives back to the original matrix.

General form of elementary matrices

We have the following general description of elementary matrices:

Remark

Elementary matrices E for each type take the following form:

- (Type I) **Interchanging k -th and l -th rows:** the diagonal entries $E_{ii} = 1$ for $i \neq k, l$ and $E_{ii} = 0$ for $i = k$ or l ; $E_{kl} = E_{lk} = 1$; and all other entries are zero.
- (Type II) **Scalar multiplication a on the k -th row:** $E_{ii} = 1$ for $i \neq k$ and $E_{kk} = a$ and all other entries are zero.
- (Type III) **Adding a scalar multiple a of the k -th row to the l -th row:** all the diagonal entries are 1 i.e. $E_{ii} = 1$ for any i , and $E_{lk} = a$, and all other entries are zero.

Operations on elementary matrices

Exercise

Determine if the followings still give an elementary matrix:

- Addition of two elementary matrices
- Multiplication of two elementary matrices
- Transpose of an elementary matrix

MATH 2101 Linear Algebra I–Determinants and inverses

2×2 determinant

Determinant of a matrix is to compute a scalar value from a matrix. It has many applications.

Definition

(2×2 matrix) The determinant of a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $ad - bc$.

Example

$$\det \begin{pmatrix} 3 & 4 \\ 2 & 9 \end{pmatrix} = 27 - 8 = 19$$

Remark

The area formed by the square $(0, 0)^T, (1, 0)^T, (0, 1)^T, (1, 1)^T$ is 1. Let A be a 2×2 matrix. The area of the parallelogram formed by

$$A(0, 0)^T, A(1, 0)^T, A(0, 1)^T, A(1, 1)^T$$

is given by $|\det(A)|$

Definition for general cases

Definition

($n \times n$ matrix) The *determinant* of an $n \times n$ matrix A is defined inductively as follows. We first define a **minor** \tilde{A}_{ij} to be the matrix obtained by **deleting the i -th row and j -th column**. Fixing certain j , we then define:

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij})$$

or fixing certain i , define

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij}).$$

The above definition is *independent of a choice of a row or a column.*

Example

Example

Let $A = \begin{pmatrix} 3 & 4 & 1 \\ 2 & 9 & 1 \\ 0 & -1 & -5 \end{pmatrix}$. The minors for the **second** column is:

$$\tilde{A}_{12} = \begin{pmatrix} 2 & 1 \\ 0 & -5 \end{pmatrix}, \quad \tilde{A}_{22} = \begin{pmatrix} 3 & 1 \\ 0 & -5 \end{pmatrix}, \quad \tilde{A}_{32} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

Then

$$\begin{aligned} \det(A) &= (-1)^{1+2}(4)\det\begin{pmatrix} 2 & 1 \\ 0 & -5 \end{pmatrix} + (-1)^{2+2}(9)\det\begin{pmatrix} 3 & 1 \\ 0 & -5 \end{pmatrix} + (-1)^{2+3}(-1)\det\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \\ &= (-1)(4)(-10) + (9)(-15) - (1)1 \end{aligned}$$

Determinant for identity matrices

Theorem

$$\det(I_n) = 1$$

$$\det(0_{n \times n}) = 0.$$

Proof.

Inductively, we have

$$\det(I_n) = (-1)^{1+1}(1)\det(I_{n-1}) = 1$$

We leave $\det(0_{n \times n}) = 0$ for an exercise.



Exercise on determinants

Exercise

Find

$$\det \begin{pmatrix} 3 & 4 & 0 & 0 \\ 2 & 9 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & -1 & 0 & -5 \end{pmatrix}$$

Exercise

Let A be a 3×3 matrix. Prove that there exists $t \in \mathbb{R}$ such that $\det(A - tI_3) = 0$. (Hint: Consider $\det(A - xI_3)$ as a polynomial in x and solve for x .)

Property 1: Switching rows

We shall not show that the definition is independent of a choice of a row or a column, but we shall use this fact to show the followings:

Theorem

(Property 1: switching rows) Let A be an $n \times n$ matrix. If B is obtained from A by switching two rows, then

$$\det(A) = -\det(B).$$

Proof.

We first consider that the **two rows are consecutive** i.e. B is obtained from A by interchanging the i -th and $(i + 1)$ -th rows. Now, we compute $\det(B)$ by using the $i + 1$ -th row and so we have:

$$\det(B) = \sum_{j=1}^n (-1)^{i+1+j} \tilde{B}_{i+1,j}.$$

But, we have that $\tilde{B}_{i+1,j} = \tilde{A}_{i,j}$. On the other hand, by computing $\det(A)$ by using the i -th row,

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \tilde{A}_{ij}.$$

The general case follows from switching two consecutive rows multiple times. □

Example

Theorem

(Property 1: switching rows) Let A be an $n \times n$ matrix. If B is obtained from A by switching two rows, then

$$\det(A) = -\det(B).$$

Example

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

Property II: Two identical rows

Corollary

(Property 2: two identical rows) Let A be an $n \times n$ matrix. If A has two **identical** rows, then $\det(A) = 0$.

Proof.

By using the previous theorem, we have $\det(A) = -\det(A)$ by switching the two identical rows. Hence $2\det(A) = 0$ and so $\det(A) = 0$. □

Example

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

Property: adding a scalar multiple on a row vector

Theorem

(Adding a scalar multiple on a row vector) Let v_1, \dots, v_n be row vectors in \mathbb{R}^n and let u be another row vector in \mathbb{R}^n . Let $k \in \mathbb{R}$. Then

$$\det \begin{pmatrix} v_1 \\ \vdots \\ v_{r-1} \\ v_r + ku \\ v_{r+1} \\ \vdots \\ v_n \end{pmatrix} = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{r-1} \\ v_r \\ v_{r+1} \\ \vdots \\ v_n \end{pmatrix} + k \det \begin{pmatrix} v_1 \\ \vdots \\ v_{r-1} \\ u \\ v_{r+1} \\ \vdots \\ v_n \end{pmatrix}$$

Proof.

The main idea of the proof is to choose the r -th row for computing the determinant. The details are left as an exercise.



Example

Example

Let $v_1 = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 & 2 & 1 \end{pmatrix}$. Let $u = \begin{pmatrix} 1 & 4 & -3 \end{pmatrix}$. Then

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

Example

$$\det \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} = \det \begin{pmatrix} 2 & 3 \\ 3 & 7 \end{pmatrix}$$

Exercise

Find the determinant of the matrix

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

Exercise

Show that

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

Transpose and determinants

Theorem

(Taking transpose) Let A be an $n \times n$ matrix. Then $\det(A^T) = \det(A)$.

Proof.

We shall prove by an induction. When $n = 1$, it is clear. Write $A = (a_{ij})$. We first pick the **first row** for computing $\det(A)$. Then we have:

$$\det(A) = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det(\tilde{A}_{1j})$$

Write $A^T = (b_{ij})$. We now pick the **first column** for computing $\det(A^T)$ and we have:

$$\det(A^T) = \sum_{i=1}^n (-1)^{i+1} b_{i1} \det(\tilde{A}^T_{i1}).$$

Note that $A_{1x}^T = \tilde{A}^T_{x1}$ and $a_{1x} = b_{x1}$. Inductively, we have $\det(A_{1x}^T) = \det(A_{1x})$. Combining the formulas, we see that two expressions coincide. □

From rows to columns

One can use the previous result to switch rows and columns to obtain:

Theorem

Let A be an $n \times n$ matrix. Then

- ① If two columns of A are switched to obtain B , then $\det(B) = -\det(A)$.
- ② If two columns of A are identical, then $\det(A) = 0$.
- ③ Let v_1, \dots, v_n be column vectors in \mathbb{R}^n and let u be another column vector in \mathbb{R}^n . Let $k \in \mathbb{R}$. Then

$$\begin{aligned}& \det(v_1 \ \dots \ v_{r-1} \ \ v_r + ku \ \ v_{r+1} \ \ \dots \ \ v_r) \\&= \det(v_1 \ \dots \ v_{r-1} \ \ v_r \ \ v_{r+1} \ \ \dots \ \ v_n) \\&\quad + k \det(v_1 \ \dots \ v_{r-1} \ \ u \ \ v_{r+1} \ \ \dots \ \ v_n)\end{aligned}$$

Determinant for scalar multiplications

Theorem

(Scalar multiplication) For an $n \times n$ matrix A and a scalar c , $\det(cA) = c^n \det(A)$.

Proof.

We shall prove by an induction on n . When $n = 1$, the statement is trivial. We now consider $n \geq 2$. Then, by definitions,

$$\det(cA) = \sum_{i=1}^n (-1)^{i+j} (ca_{ij}) \det(\tilde{A}_{ij}).$$

By inductive hypothesis, we have that:

$$\det(\tilde{A}_{ij}) = c^{n-1} \det(\tilde{A}_{ij}).$$

Hence, $\det(cA) = \sum_{i=1}^n (-1)^{i+j} c^n (a_{ij} \det(\tilde{A}_{ij}))$.



Multiplicative property

Theorem

(Multiplicative property) Let A, B be two $n \times n$ -matrices. Then
 $\det(AB) = \det(A) \cdot \det(B)$.

Exercise

An $n \times n$ matrix A is said to be *orthogonal* if $AA^T = I_n$. Prove that if Q is orthogonal, then $\det(Q) = \pm 1$.

How about determinant on matrix additions?

There is no formula between determinant and matrix additions! In general,

$$\det(A + B) \neq \det(A) + \det(B).$$

Example

$$\det\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \neq \det\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \det\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Definition of an inverse

Definition

An $n \times n$ matrix A is said to be *invertible* if there exists a matrix B such that $AB = BA = I_n$. We say call B to be the *inverse* of A . We shall write the inverse of A to be A^{-1} .

Remark

The inverse of an invertible matrix A is unique. If C is another inverse of A , then $C(AB) = CI_n = C$ and $(CA)B = I_nB = B$. By associativity,
 $C = C(AB) = (CA)B = B$.

Determinant of A^{-1}

Theorem

Let A be an invertible $n \times n$ -matrix. Then $\det(A) \neq 0$. Moreover,

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

Proof.

$$AA^{-1} = I_n \xrightarrow{\text{taking } \det} \det(AA^{-1}) = \det(I_n) \xrightarrow{\det(I_n)=1} \det(AA^{-1}) = 1$$

$$\begin{aligned} &\text{use } \det(AB) = \det(A)\det(B) \\ &\xrightarrow{\quad} \det(A)\det(A^{-1}) = 1 \end{aligned}$$

$$\text{Hence, } \det(A^{-1}) = \frac{1}{\det(A)}. \quad \square$$

Proof

Theorem

Let A be an $n \times n$ matrix. Define the adjugate of A as:

$$\text{adj}(A) = \begin{pmatrix} (-1)^{1+1} \det \tilde{A}_{11} & \dots & (-1)^{1+n} \det \tilde{A}_{1n} \\ \vdots & & \vdots \\ (-1)^{n+1} \det \tilde{A}_{n1} & \dots & (-1)^{n+n} \det \tilde{A}_{nn} \end{pmatrix}^T$$

If $\det(A) \neq 0$, then $\frac{1}{\det A} \cdot \text{adj}(A)$ is the inverse of A .

Proof.

One needs to use the following two formulas: when multiplying the i -th row of A with i -th column in $\text{adj}(A)$,

$$\sum_{x=1}^n (-1)^{i+x} a_{ix} \det \tilde{A}_{ix} = \det(A)$$

when multiplying the i -th row with i' -th column in $\text{adj}(A)$ ($i' \neq i$),

$$\sum_{x=1}^n (-1)^{i+x} a_{ix} \det \tilde{A}_{i'x} = 0$$

Detecting invertibility of a matrix by computing determinant

It is *not* *flesible* to check the invertibility of a matrix from definitions. The determinant provides a convenient way to do so:

Theorem

Let A be an $n \times n$ matrix. Then A is *invertible* if and only if $\det(A) \neq 0$.

A formula for computing A^{-1}

Theorem

Let A be an $n \times n$ matrix. Define the *adjugate of A* as:

$$\text{adj}(A) = \begin{pmatrix} (-1)^{1+1}\det\tilde{A}_{11} & \dots & (-1)^{1+n}\det\tilde{A}_{1n} \\ \vdots & & \vdots \\ (-1)^{n+1}\det\tilde{A}_{n1} & \dots & (-1)^{n+n}\det\tilde{A}_{nn} \end{pmatrix}^T,$$

where \tilde{A}_{ij} is the (i,j) -minor matrix of A . If $\det(A) \neq 0$, then

$$\frac{1}{\det A} \cdot \text{adj}(A)$$

is the inverse of A .

Example

Example

Determine if the following matrices are invertible. If it is invertible, find the inverse.

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

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

Solution:

- $\det(A) = \det(B) = 0$ (why? explain it!) and so A and B are not invertible.
- $\det(C) = 1 \neq 0$ and so C is invertible. Then

$$\text{adj}(C) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix}^T, C^{-1} = \frac{1}{1} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -3 \\ 0 & -1 & 2 \end{pmatrix}.$$

Inverse with other operations

Theorem

Let A, B be an invertible $n \times n$ matrix. Let $c \in \mathbb{R}$. Then

- ① $(AB)^{-1} = B^{-1}A^{-1};$
- ② $(cA)^{-1} = c^{-1}A^{-1};$
- ③ $(A^T)^{-1} = (A^{-1})^T;$
- ④ $(A^{-1})^{-1} = A.$

Proof.

We only check for the left inverses.

- ① $(B^{-1}A^{-1})(AB) = B^{-1}I_nB = B^{-1}B = I_n$
- ② $(c^{-1}A^{-1})(cA) = c^{-1}cA^{-1}A = I_n$
- ③ $(A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n$
- ④ $\textcolor{blue}{A} A^{-1} = I_n$

Applications of determinants and inverses

Determinants and matrix inversions are useful in solving system of linear equations e.g. Carmer's rule. We will soon talk about this in next section. To give you an idea, let A be an $n \times n$ matrix and b be a column vector in \mathbb{R}^n . If we want to *solve* a column vector $x \in \mathbb{R}^n$ such that $Ax = b$ and A is invertible, then the *solution* to the equation is:

$$x = A^{-1}b$$

since $A(A^{-1}b) = I_n b = b$.