Lesson 3
System of Linear Equations

Christopher Ting

✉: cting@hiroshima-u.ac.jp
Christopher Ting
☎: +81 082-424-7268
: http://cting.x10host.com/

April 17, 2020 Version 1.2

Table of Contents

Inspiring Motivation

Considering the inconceivable complexity of processes even in a simple cell, it is little short of a miracle that the simplest possible model—namely, a linear equation between two variables—actually applies in quite a general number of cases.

—**Ludwig von Bertalanffy**

Learning Outcomes

- Explain the reasons for why we need to study system of linear equations.
- Define and connect the following concepts: matrix-vector formulation, augmented matrix, simplified matrix, dimension, and rank.
- Solve a system of linear equations algebraically by the elimination method, while maintaining a connection to the augmented matrix.
- Interpret how the elimination method corresponds to performing row operations on an augmented matrix.

Learning Outcomes (Cont'd)

- Elaborate the basic matrices, their properties, and how they can be applied in simplifying a matrix.
- Assess whether a matrix is in its (reduced) row echelon or simplified form.
- Generate the solutions by performing the row reduction or Gaussian elimination algorithm.
- Describe the conditions for the system of linear equations to have solutions and their uniqueness.
- Develop an intuitive picture of how the dimension of a square matrix, its regularity, its inverse, and its simplified form, are connected.

What is Linear Algebra?

- Linear $+$ Algebra
- Linear: straight lines, flat planes, hyperplanes
- An **operator** L is said to be **linear** if it has two properties:
  1. *Homogeneity:* “Constant c can be pulled out"

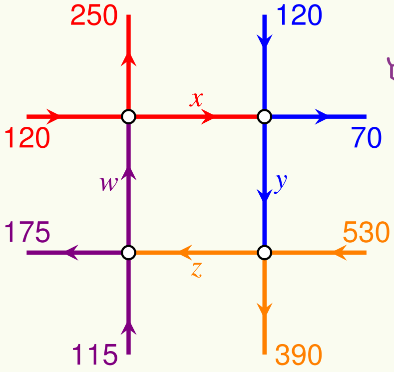
$$L(cv) = cL(v)\,.$$

  2. *Additivity:* “The operator L can be distributed"

$$L(v+u) = L(v) + L(u)\,.$$

- Linear algebra is the study of vectors and linear functions.

Application: Monitoring the Traffic Flow



- The numbers and arrows indicate the number of cars per hour flowing along each street.
- Since the number of cars entering each intersection has to equal the number of cars leaving that intersection, we obtain a system of linear equations:

$$\begin{cases} w + 120 = x + 250, \\ x + 120 = y + 70, \\ y + 530 = z + 390, \\ z + 115 = w + 175. \end{cases}$$

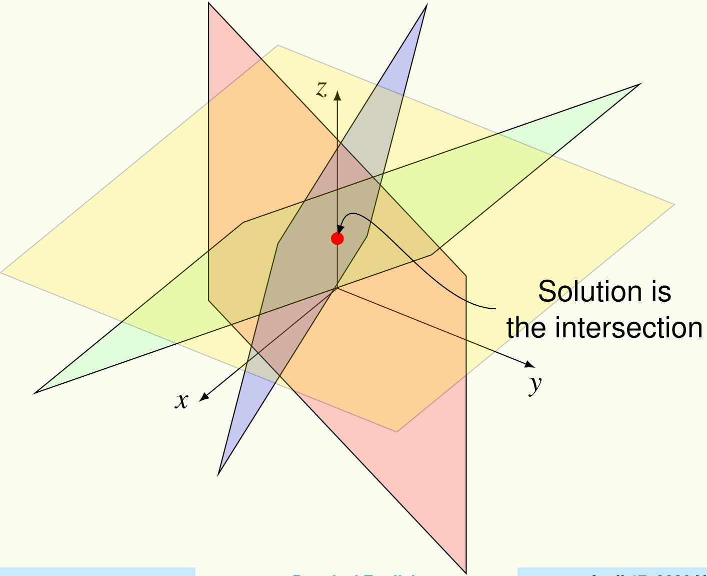
Basic Idea

- An equation in the unknowns $x, y, z, \ldots$ is called **linear** if both sides of the equation are a sum of (constant) multiples of $x, y, z, \ldots$ plus an optional constant.

- A **system of linear equations** is a collection of several linear equations, such as the traffic flows expressed in the standard form:

$$\begin{cases} w - x = 130, \\ x - y = -50, \\ y - z = -140, \\ z - w = 60. \end{cases}$$

Geometric Illustration



Solution is
the intersection

System of Linear Equations

Definition 2.1 (System of Linear Equations).

A **system of linear equations** is a collection of m equations in the variable quantities $x_1, x_2, x_3, \ldots, x_n$ of the form,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = b_3, \\ \qquad\qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

where the values of a_{ij}, b_i, and x_j $\left(1 \leq i \leqq m, 1 \leq j \leqq n\right)$ are from the set of complex numbers.

Solution of a System of Linear Equations

Definition 2.2 (Solution of a System of Linear Equations).

A **solution** of a system of linear equations in n variables $x_1, x_2, \ldots, x_n$ is an ordered list of n complex numbers, $s_1, s_2, \ldots s_n$ such that if we substitute s_1 for x_1, s_2 for x_2, $\cdots$, and s_n for x_n, then for every equation of the system, the left side will equal the right side, i.e. each equation is true simultaneously.

Definition 2.3 (Solution Set of a System of Linear Equations).

The **solution set** of a linear system of equations is the set that contains every solution to the system, and nothing more.

- The solution set can have infinitely many elements.
- It can be an empty set as there are no solutions.

Example of a System of Equations

- Given the system of linear equations,

$$\begin{aligned} x_1 + 2x_2 + x_4 &= 7, \\ x_1 + x_2 + x_3 - x_4 &= 3, \\ 3x_1 + x_2 + 5x_3 - 7x_4 &= 1, \end{aligned}$$

  we have $n = 4$ variables and $m = 3$ equations. Also,

$$\begin{array}{lllll} a_{11} = 1 & a_{12} = 2 & a_{13} = 0 & a_{14} = 1 & b_1 = 7 \\ a_{21} = 1 & a_{22} = 1 & a_{23} = 1 & a_{24} = -1 & b_2 = 3 \\ a_{31} = 3 & a_{32} = 1 & a_{33} = 5 & a_{34} = -7 & b_3 = 1 \end{array}$$

- One solution is $x_1 = -2$, $x_2 = 4$, $x_3 = 2$, $x_4 = 1$.
- But it is not the only one. For example, another solution is $x_1 = -12$, $x_2 = 11$, $x_3 = 1$, $x_4 = -3$.
- More solutions can be found.

Coefficient Matrix, Variable and Constant Vectors

Definition 2.4 (Matrix-Vector Form).

With respect to the system of linear equations (1), let

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

Then (1) can be written as

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

- $\boldsymbol{A}$: **coefficient matrix**
- $\boldsymbol{x}$: **variable vector**
- $\boldsymbol{b}$: **constant vector**

Augmented Matrix

Definition 2.5 (Augmented Matrix).

Given the coefficient matrix $\boldsymbol{A}$ and the constant matrix $\boldsymbol{b}$, the **augmented matrix** is defined as

$$\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] = \left[\begin{array}{ccc|c} \boldsymbol{a}_{11} & \cdots & \boldsymbol{a}_{1n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{a}_{m1} & \cdots & \boldsymbol{a}_{mn} & b_m \end{array}\right].$$

- The augmented matrix has the dimension of $m \times (n+1)$.

Equation Operations

Definition 2.6 (Equation Operations).

Given a system of linear equations, the following three operations will transform the system into a different system, and each operation is known as an **equation operation**.

1. Swap the locations of two equations in the list of equations.
2. Multiply each term of an equation by a nonzero quantity.
3. Multiply each term of an equation by some quantity, and add these terms to another equation, on both sides of the equality, which obtains a new equation. Leave the former equation the same after this operation, but replace the latter equation by the new one.

Example: Augmented Matrix

- Consider the following system of linear equations:

$$\begin{aligned} 2x_1 - x_2 &= 0, \\ -x_1 + x_2 - 2x_3 &= 4, \\ 3x_1 - 2x_2 + x_3 &= -2. \end{aligned}$$

- The augmented matrix is

$$\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] = \left[\begin{array}{ccc|c} 2 & -1 & 0 & 0 \\ -1 & 1 & -2 & 4 \\ 3 & -2 & 1 & -2 \end{array}\right].$$

- It is a 3 by 4 matrix.

Example: Transformation Step 1

1. Multiplying the first equation by $\dfrac{1}{2}$ corresponds to multiplying the first row of the matrix by $\dfrac{1}{2}$, which is denoted by $R_1 \longleftarrow \frac{1}{2}R_1$.

$$\begin{aligned} x_1 - \tfrac{1}{2}x_2 &= 0 \\ -x_1 + x_2 - 2x_3 &= 4 \\ 3x_1 - 2x_2 + x_3 &= -2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ -1 & 1 & -2 & 4 \\ 3 & -2 & 1 & -2 \end{array}\right].$$

Example: Transformation Step 2

2. Then, we add the first equation to the second equation. This corresponds to adding the first row of the matrix to the second row of the matrix, which is denoted by $R_2 \longleftarrow R_2 + R_1$. We obtain

$$\begin{aligned} x_1 - \tfrac{1}{2}x_2 &= 0 \\ \tfrac{1}{2}x_2 - 2x_3 &= 4 \\ 3x_1 - 2x_2 + x_3 &= -2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -2 & 4 \\ 3 & -2 & 1 & -2 \end{array}\right].$$

Example: Transformation Step 3

3. Next, we add -3 times the first equation to the third equation. This corresponds to adding -3 times the first row of the matrix to the third row of the matrix, which is notated by $R_3 \leftarrow R_3 - 3R_1$. The result is

$$\begin{aligned} x_1 - \tfrac{1}{2}x_2 &= 0 \\ \tfrac{1}{2}x_2 - 2x_3 &= 4 \\ \color{red}{-\tfrac{1}{2}x_2 + x_3} &\color{red}{= -2} \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -2 & 4 \\ \color{red}{0} & \color{red}{-\frac{1}{2}} & \color{red}{1} & \color{red}{-2} \end{array}\right].$$

Example: Transformation Steps 4 and 5

4. Adding the second row of the matrix to the third row of the matrix, which is notated by $R_3 \longleftarrow R_3 + R_2$, we obtain

$$\begin{aligned} x_1 - \tfrac{1}{2}x_2 &= 0 \\ \tfrac{1}{2}x_2 - 2x_3 &= 4 \\ -x_3 &= 2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -2 & 4 \\ 0 & 0 & -1 & 2 \end{array}\right].$$

5. $R_3 \longleftarrow -R_3$

$$\begin{aligned} x_1 - \tfrac{1}{2}x_2 &= 0 \\ \tfrac{1}{2}x_2 - 2x_3 &= 4 \\ x_3 &= -2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -2 & 4 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

Example: Transformation Steps 6 and 7

6. $R_2 \leftarrow R_2 + 2R_3$

$$\begin{aligned} x_1 - \tfrac{1}{2}x_2 &= 0 \\ \tfrac{1}{2}x_2 &= 0 \\ x_3 &= -2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -\tfrac{1}{2} & 0 & 0 \\ 0 & \tfrac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

7. $R_1 \longleftarrow R_1 + R_2$

$$\begin{aligned} x_1 &= 0 \\ \tfrac{1}{2}x_2 &= 0 \\ x_3 &= -2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & \tfrac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

Example: Transformation Step 8

8. $R_2 \leftarrow 2R_2$

$$\begin{aligned} x_1 &= 0 \\ x_2 &= 0 \\ x_3 &= -2 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

- So the augmented matrix becomes $\left[\boldsymbol{I} \,\middle|\, \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}\right]$
- Note that this is not the only way to transform $\boldsymbol{A}$ into $\boldsymbol{I}$.

Definition of Basic Matrix Transformations

Definition 2.7 (Basic Matrix Transformations).

The following 3 operations are called the **basic transformations** with respect to the row of a matrix:

(i) interchanging one row vector with the other row vector,
(ii) multiplying one row with a non-zero constant,
(iii) adding to a row another row that has been multiplied with a non-zero constant.

Moreover, the basic transformations with respect to the column of a matrix can be correspondingly defined by changing the row vectors in these 3 operations to column vectors.

〰 The matrices that correspond to these 3 operations are called the **basic matrices**.

Definition 1 Basic Matrix P_{ij}

$$\begin{array}{cc}
 & \begin{array}{ccccccccccc} & & & \text{column } i & & & & \text{column } j & & & \\ & & & \downarrow & & & & \downarrow & & & \end{array} \\
\boldsymbol{P}_{ij} = & \left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]
\begin{array}{l} \\ \\ \\ \text{row } i \\ \\ \\ \\ \text{row } j \\ \\ \\ \\ \end{array}
\end{array}$$

Definition 2 Basic Matrix $Q_i(c), c \neq 0$

$$\boldsymbol{Q}_i(c) = \begin{array}{cc} \begin{array}{ccccccc} &&&\text{column } i&&& \\ &&&\downarrow&&& \end{array} & \\ \begin{bmatrix} 1 &&&&&& \\ &\ddots&&&&& \\ &&1&&&& \\ &&&c&&& \\ &&&&1&& \\ &&&&&\ddots& \\ &&&&&&1 \end{bmatrix} & \begin{array}{c} \\ \\ \\ \text{row } i \\ \\ \\ \\ \end{array} \end{array}$$

Definition 3 Basic Matrix $\boldsymbol{R}_{ij}(d), d \neq 0$

$$
\begin{array}{rcl}
 & \begin{array}{ccccccccccc} & & & \text{column } i & & & & \text{column } j & & & \\ & & & \downarrow & & & & \downarrow & & & \end{array} & \\
\boldsymbol{R}_{ij}(d) = & \left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 1 & & & & d & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right] & \begin{array}{l} \\ \\ \\ \text{row } i \\ \\ \\ \\ \text{row } j \\ \\ \\ \\ \end{array}
\end{array}
$$

Matrix Operations by Basic Matrices

- Matrix operation (i) in Definition 2.7 corresponds to

$$\boldsymbol{A} \longrightarrow \boldsymbol{P}_{ij}\boldsymbol{A}.$$

- Matrix operation (ii) is the transformation

$$\boldsymbol{A} \longrightarrow \boldsymbol{Q}_{i}(c)\boldsymbol{A}.$$

- Matrix operation (iii) is implemented by

$$\boldsymbol{A} \longrightarrow \boldsymbol{R}_{ij}(d)\boldsymbol{A}.$$

Corresponding Transformation Matrix

- For Step 1 in Slide 17, the basic transformation matrix is
$$\boldsymbol{Q}_1\left(\frac{1}{2}\right) = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
- The reason is
$$\boldsymbol{Q}_1\left(\frac{1}{2}\right)\boldsymbol{A} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & -2 \\ 3 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ -1 & 1 & -2 \\ 3 & -2 & 1 \end{bmatrix},$$
which is the new $\boldsymbol{A}$ matrix after transformation.

Corresponding Transformation Matrix (Cont'd)

- After Step 2, the new $\boldsymbol{A}$ matrix is $\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -2 \\ 3 & -2 & 1 \end{bmatrix}$.
- For Step 3 in Slide 19, the basic transformation matrix is

$$\boldsymbol{R}_{31}(-3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}.$$

- The reason is

$$\boldsymbol{R}_{31}(-3)\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -2 \\ 3 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -2 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix},$$

which is the new $\boldsymbol{A}$ matrix after transformation.

Theorem 2.8 on Basic Matrices

Theorem 2.8 (Basic Matrices).

The basic matrices are regular, and their inverses are some basic matrices.

Proof.

By simple calculations, it can be established that

$$\boldsymbol{P}_{ij}\boldsymbol{P}_{ij} = \boldsymbol{Q}_i(c)\boldsymbol{Q}_i\left(c^{-1}\right) = \boldsymbol{R}_{ij}(d)\boldsymbol{R}_{ij}(-d) = \boldsymbol{I}.$$

It follows that

$$\boldsymbol{P}_{ij}^{-1} = \boldsymbol{P}_{ij}, \quad \boldsymbol{Q}_i^{-1}(c) = \boldsymbol{Q}\left(c^{-1}\right), \quad \boldsymbol{R}_{ij}^{-1}(d) = \boldsymbol{R}_{ij}(-d).$$

□

Remarks

- Two ways of solving a system of linear equations:

System of Linear Equations $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \quad \overset{\text{Matrix Operations}}{\longrightarrow \cdots\cdots\cdots\cdots \longrightarrow} \quad \boldsymbol{I}\boldsymbol{x}=\boldsymbol{x}=\widehat{\boldsymbol{b}}$

Augmented Matrix $\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] \quad \underset{\text{Basic Transformation}}{\longrightarrow \cdots\cdots\cdots\cdots \longrightarrow} \quad \left[\boldsymbol{I}\middle|\widehat{\boldsymbol{b}}\right]$

- The method based on the augmented matrix is called **the Gauss-Jordan elimination**.

- Note that these two methods work only if the system has a unique set of solutions.

Example

- Represent the system of linear equations as vector-matrix equation and augmented matrix.

$$\left.\begin{array}{rcrcr} x & + & y & = & 27 \\ 2x & - & y & = & 0 \end{array}\right\} \Longleftrightarrow \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 27 \\ 0 \end{bmatrix} \Longleftrightarrow \left[\begin{array}{cr|r} 1 & 1 & 27 \\ 2 & -1 & 0 \end{array}\right].$$

- With the first equation replaced by the sum of the two equations, we obtain

$$\left.\begin{array}{rcrcr} 3x & + & 0 & = & 27 \\ 2x & - & y & = & 0 \end{array}\right\} \Longleftrightarrow \begin{bmatrix} 3 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 27 \\ 0 \end{bmatrix} \Longleftrightarrow \left[\begin{array}{cr|r} 3 & 0 & 27 \\ 2 & -1 & 0 \end{array}\right].$$

- Let the new first equation be the old first equation divided by 3:

$$\left.\begin{array}{rcrcr} x & + & 0 & = & 9 \\ 2x & - & y & = & 0 \end{array}\right\} \Longleftrightarrow \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix} \Longleftrightarrow \left[\begin{array}{cr|r} 1 & 0 & 9 \\ 2 & -1 & 0 \end{array}\right].$$

Example (Cont'd)

- Replace the second equation by the second equation minus two times the first equation:

$$\left.\begin{array}{rcrcr} x & + & 0 & = & 9 \\ 0 & - & y & = & -18 \end{array}\right\} \Longleftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 9 \\ -18 \end{bmatrix} \Longleftrightarrow \left[\begin{array}{rr|r} 1 & 0 & 9 \\ 0 & -1 & -18 \end{array}\right].$$

- Let the new second equation be the old second equation multiplied by -1:

$$\left.\begin{array}{rcrcr} x & + & 0 & = & 9 \\ 0 & + & y & = & 18 \end{array}\right\} \Longleftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 9 \\ 18 \end{bmatrix} \Longleftrightarrow \left[\begin{array}{rr|r} 1 & 0 & 9 \\ 0 & 1 & 18 \end{array}\right].$$

- The solution is $x = 9$ and $y = 18$.

Main Component and Simplified Matrix

Definition 3.1 (Main Component).

Among all the elements of a row vector, the leftmost non-zero element is defined as the **main component** of the row vector.

Definition 3.2 (Simplified Matrix).

The **simplified matrix** is defined as the matrix that satisfies the following 4 conditions:

(1) For every row that is a null row vector, it is always below a row vector that is not a null vector.

(2) The main component of each row is located more to the right as one moves downward on the matrix.

(3) The main component is 1 for the non-null row vector.

(4) The column vector that contains the main component of a row vector has all the elements equal to zero except the row that contains the main component.

Illustration of a Simplified Matrix and Its Rank

$$\begin{bmatrix}
0\cdots0 & 1 & *\cdots* & 0 & *\cdots* & & 0 & * & \cdots & * \\
 & 0 & 0\cdots0 & 1 & *\cdots* & & \vdots & \vdots & \vdots & \vdots \\
 & & & & & \ddots & & & & \\
 & & & 0 & 0\cdots0 & & 0 & * & \cdots & * \\
 & & & & & & 1 & * & \cdots & * \\
 & & & & & & 0 & 0 & \cdots & 0 \\
 & & & & & & \vdots & \vdots & \cdots & \vdots \\
 & & & & & & 0 & 0 & \cdots & 0
\end{bmatrix}$$

Definition 3.3 (Rank of a Simplified Matrix).

With respect to a simplified matrix, its **rank** is defined as the sum of the main components.

Remarks

- The simplified form is a literal translation of 簡約形 in the Japanese textbook. It is called the **reduced echelon form** in English textbooks.
- The Japanese-translated “main component” 主成分 is called the “**leading term**”.
- Example of a simplified matrix: $\begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
- Not a simplified matrix:

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$. The left-most nonzero entry in the second row is not equal to 1

$\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The leading 1 in the third row is not the only nonzero entry in the column containing it.

Simplified Matrix and Rank

Theorem 3.4 (Transformation to Simplified Matrix).

For any matrix $\boldsymbol{A}$, by applying the matrix operations a finite number of times with respect to the rows, $\boldsymbol{A}$ will surely be transformed into a simplified matrix $\boldsymbol{B}$.

Theorem 3.5 (Uniqueness of Simplified Matrix).

The simplified matrix of a $m \times n$ matrix is unique, i.e., it does not depend on the procedure of simplification.

Definition 3.6 (Rank of Matrix).

For any matrix $\boldsymbol{A}$, its **rank** is the rank of its simplified matrix $\boldsymbol{B}$.

Theorem 3.7 (Upper Limit of Rank).

For any matrix $m \times n$ matrix $\boldsymbol{A}$, $\mathrm{Rank}\boldsymbol{A} \leq m$ and $\mathrm{Rank}\boldsymbol{A} \leq n$.

Proof of Theorem 3.4

- (a) For a given matrix $\boldsymbol{A}$, by repeating basic transformation (i), move the rows with all elements equal to 0 in such a way that (1) of Definition 3.2 is satisfied.
- (b) Apply basic transformation (ii) to normalize the main component of the first row to 1.
- (c) Apply basic transformation (iii) such that the elements of the column that contains the main component of the first row become zero.
- (d) Repeat (b) and (c) to the next row so that the transformed matrix satisfies the conditions (3) and (4) of Definition 3.2.
- (e) For rows that contain non-zero elements, apply basic transformation (i) so that (2) of Definition 3.2 is satisfied. □

Theorems of Existence and Uniqueness

Theorem 3.8 (Condition for the Existence of Solutions).

For any $m \times n$ *matrix* $\boldsymbol{A}$ *and an* m*-dimensional vector* $\boldsymbol{b}$*, the existence of the solution set for the system of linear equations* $\boldsymbol{Ax} = \boldsymbol{b}$ *is equivalent to the condition that*

$$\text{Rank}\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] = \text{Rank}\,\boldsymbol{A}.$$

Theorem 3.9 (Condition for the Uniqueness of Solutions).

For any $m \times n$ *matrix* $\boldsymbol{A}$ *and an* m*-dimensional vector* $\boldsymbol{b}$*, the solution to the system of linear equations* $\boldsymbol{Ax} = \boldsymbol{b}$ *is unique if and only if*

$$\text{Rank}\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] = \text{Rank}\,\boldsymbol{A} = n.$$

System of First-Order Equations

Definition 3.10 (System of First-Order Equations).

With respect to a $m \times n$ matrix $\boldsymbol{A}$, the system of first-order equations

$$\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$$

is called the **system of first-order equations**. The solution $\boldsymbol{x} = \mathbf{0}$ is called the **trivial solution**. All other solutions are called the **non-trivial solutions**.

Theorem 3.11 (Trivial Solution).

Let $\boldsymbol{A}$ be an $m \times n$ matrix. Only trivial solution is possible for $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ if and only if the rank of $\boldsymbol{A}$ is n.

Proof.

The augmented matrix is $\left[\boldsymbol{A} \middle| \mathbf{0}\right]$. Now, $\mathrm{Rank}\left[\boldsymbol{A} \middle| \mathbf{0}\right] = \mathrm{Rank}\,\boldsymbol{A}$. □

Non-trivial Solutions

Theorem 3.12 (Existence of Non-Trial Solutions).

Let $\boldsymbol{A}$ *be a* $m \times n$ *matrix for which* $m < n$*. Then there exist non-trivial solutions of* $\boldsymbol{Ax} = \mathbf{0}$*.*

Proof.

From Theorem 3.7, $\mathrm{Rank}\boldsymbol{A} \leqq m < n$. From Theorem 3.11, $\boldsymbol{Ax} = \mathbf{0}$ cannot have trivial solutions. Thus the solutions of $\boldsymbol{Ax} = \mathbf{0}$ are non-trivial. □

Example of Trivial Solution

Solve the system of first-order equations: $\begin{cases} 2x+3y=0, \\ 4x+5y=0. \end{cases}$

- In the matrix form, the system of first-order equations is

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

- Simplification of the coefficient matrix:

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \xrightarrow{R_1 \longleftarrow \frac{1}{2}R_1} \begin{bmatrix} 1 & \frac{3}{2} \\ 4 & 5 \end{bmatrix} \xrightarrow{R_2 \longleftarrow 4R_1 - R_2} \begin{bmatrix} 1 & \frac{3}{2} \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 \longleftarrow R_1 - \frac{3}{2}R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}_2$$

- Hence, $\mathrm{Rank}\,\boldsymbol{A} = \mathrm{Rank}\,\boldsymbol{I}_2 = 2$.
- By Theorem 3.11, the solution is trivial, i.e., $\begin{bmatrix} x & y \end{bmatrix}' = \begin{bmatrix} 0 & 0 \end{bmatrix}$.

Example of Non-Trivial Solution

Solve the system of first-order equations: $\begin{cases} 2x + 3y + 4z = 0, \\ 4x + 5y + 6z = 0. \end{cases}$

- Simplification of the coefficient matrix:

$$\begin{bmatrix} 2 & 3 & 4 \\ 4 & 5 & 6 \end{bmatrix} \xrightarrow{R_1 \longleftarrow \frac{1}{2} R_1} \begin{bmatrix} 1 & \frac{3}{2} & 2 \\ 4 & 5 & 6 \end{bmatrix} \xrightarrow{R_2 \longleftarrow 4R_1 - R_2} \begin{bmatrix} 1 & \frac{3}{2} & 2 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{R_1 \longleftarrow R_1 - \frac{3}{2} R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}.$$

- Hence, the system of first-order equations becomes

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Longleftrightarrow \quad \begin{cases} x \quad - \ z = 0, \\ \quad y + 2z = 0. \end{cases}$$

Example of Non-Trivial Solution (Cont'd)

- Since the number of rows is 2 and the number of columns is 3, by Theorem 3.12, non-trivial solution exists.
- We find that $x = z$ and $y = -2z$.
- Let c be any arbitrary number. The solution is

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

Theorem 4.1

Theorem 4.1 (Regularity and Rank).

For an n-dimensional square matrix $\boldsymbol{A}$, the following statements are equivalent:

(1) $\boldsymbol{A}$ *is regular.*

(2) $\mathrm{Rank}\boldsymbol{A} = n$.

(3) *The simplified form of $\boldsymbol{A}$ is the n-dimensional identity matrix $\boldsymbol{I}_n$.*

Proof of Theorem 4.1

Proof.

- (1) $\Longrightarrow$ (2)
  Multiply both sides of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ by $\boldsymbol{A}^{-1}$ and we obtain $\boldsymbol{x} = \boldsymbol{0}$, which is a trivial solution. By Theorem 3.11, the rank of $\boldsymbol{A}$ is n.
- (2) $\Longrightarrow$ (3)
  It is obvious from the definition of rank.
- (3) $\Longrightarrow$ (1)
  Let $\boldsymbol{P}_i$, $i = 1, \ldots, m$ be the basic operation matrices. Given that the simplified form of $\boldsymbol{A}$ is $\boldsymbol{I}_n$, we have $\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m\boldsymbol{A} = \boldsymbol{I}_n$. Now, since each $\boldsymbol{P}_i$ has an inverse, we then have $\boldsymbol{A} = \boldsymbol{P}_m^{-1}\boldsymbol{P}_{m-1}^{-1}\cdots\boldsymbol{P}_1^{-1}$. By Theorem 2.8, $\boldsymbol{P}_i^{-1} = \boldsymbol{P}_i$, $i = 1, \ldots, m$. Let $\boldsymbol{T} := \boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m$. We thus obtain $\boldsymbol{T}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{T} = \boldsymbol{I}_n$. It follows that $\boldsymbol{T}$ is the inverse of $\boldsymbol{A}$ and $\boldsymbol{A}$ must be regular.

□

Basic Operation Matrices and Regular Matrix

Theorem 4.2 (Product of Basic Operation Matrices).

All the regular matrices can be expressed as a product of basic operation matrices.

Theorem 4.3 (Commutativity).

For any square matrices $\boldsymbol{A}$ *and* $\boldsymbol{X}$*,*

$$\boldsymbol{AX} = \boldsymbol{I} \qquad \Longleftrightarrow \qquad \boldsymbol{XA} = \boldsymbol{I}.$$

Theorem 4.4 (Square Matrix's Simplified Form).

If an n*-dimensional square matrix* $\boldsymbol{A}$ *is regular, the simplified form of* $\left[\boldsymbol{A}\middle|\boldsymbol{I}_n\right]$ *is* $\left[\boldsymbol{I}_n\middle|\boldsymbol{A}^{-1}\right]$.

Proof of Theorem 4.4

- According to Theorem 4.1, the simplified form of a regular matrix is the identify matrix.
- Since $\boldsymbol{A}$ is regular, we have $\boldsymbol{P}_1 \cdots \boldsymbol{P}_m \boldsymbol{A} = \boldsymbol{I}_n$, where $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots \boldsymbol{P}_m$ are some basic matrices.
- Thus, $\boldsymbol{A}^{-1} = \boldsymbol{P}_1 \boldsymbol{P}_2 \cdots \boldsymbol{P}_m$.
- Under this scenario, we shall simplify $\left[\boldsymbol{A} \middle| \boldsymbol{I}_n\right]$ as follows:

$$\left[\boldsymbol{A} \middle| \boldsymbol{I}_n\right] \longleftrightarrow \left[\boldsymbol{P}_1 \boldsymbol{P}_2 \cdots \boldsymbol{P}_m \boldsymbol{A} \middle| \boldsymbol{P}_1 \boldsymbol{P}_2 \cdots \boldsymbol{P}_m\right] = \left[\boldsymbol{I}_m \middle| \boldsymbol{A}^{-1}\right].$$

Takeaways

- Vocabulary and concepts: augmented matrix, leading term (main component) of a matrix, and reduced echelon (simplified) matrix
- Gauss-Jordan elimination method and basic matrices
- Condition for the existence of solutions: $\mathrm{Rank}\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] = \mathrm{Rank}\,\boldsymbol{A}$
- Condition for the uniqueness of solutions: $\mathrm{Rank}\left[\boldsymbol{A}\middle|\boldsymbol{b}\right] = \mathrm{Rank}\,\boldsymbol{A} = n$
- Intuitive meanings of a system of first-order equations:
  - trivial solution: all hyperplanes intersect at the point of origin.
  - Non-trivial solution: some variables are redundant.
- An n-dimensional square matrix is regular if and only if it has the rank of n.

Keywords