Slide 1: Linear Systems and Matrices I

MAT2040 Linear Algebra

Introduction: what is linear algebra?

- The topic has been studied since 17th century.
- Widely used in mathematics, engineering, natural science, computer science, economics, etc.
- Linear algebra= Linear + algebra.

Introduction: what is linear algebra?

Geometrically, linear means “straight” or “flat”.

Linear means that only addition and (scalar) multiplication are involved.

In 2D, a “linear equation” represents a line. In 3D, a “linear equation” means a plane.

Algebra is to study the mathematical symbols and the rules of manipulating these symbols.

Examples for linear and nonlinear equations

Example of linear equations:

(i) $x + 3y - 2z = 8$

(ii) $4x_1 + 5x_2 - x_3 + x_4 + x_5 = 0$

(iii) $9a - 2b + 7c + 2d = -7$

Example of nonlinear equations:

(i) $xy + 5yz = 13.$

(ii) $x_1 + x_2^3 - x_3x_4x_5^2 = 0$

(iii) $\tan(a) + \log(b) - cd = 0$

Solution type for linear system

Example 1.1

Solving the system of equations

$$2x_1 + 3x_2 = 3, \qquad (1)$$
$$x_1 - x_2 = 4 \qquad (2)$$

Answer. Equation (1) - 2× equation (2) gives

$$5x_2 = -5,$$
$$x_2 = -1.$$

Substituting it into back into (1), one gets

$$2x_1 + 3(-1) = 3.$$

Thus

$$x_1 = 3.$$

Geometrical Interpretation for 2×2 system

Example 1.2

$$(i) \quad \begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 2 \end{aligned}$$

It has unique solution $x_1 = 2, x_2 = 0$.

$$(ii) \quad \begin{aligned} x_1 + x_2 &= 2 \\ x_1 + x_2 &= 1 \end{aligned}$$

It has no solution.

$$(iii) \quad \begin{aligned} x_1 + x_2 &= 2 \\ -x_1 - x_2 &= -2 \end{aligned}$$

It has infinitely many solutions.

Geometrical Interpretation for 2×2 system

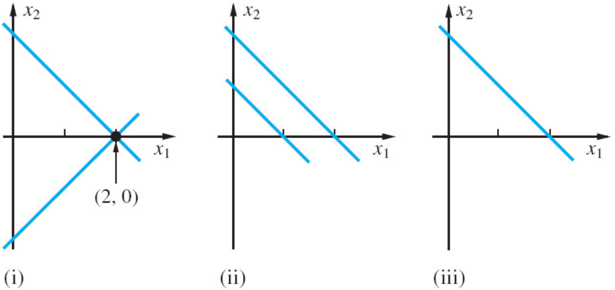

Figure: Geometrical meaning of the solutions

Solving the linear system is one of central topics in Linear algebra

All above mentioned linear system is 3×3 or 2×2 system.

How about solving a general linear system with m equations and n variables? This is one of central topics in this course.

Definition 1.3 (System of Linear Equations)

A system of linear equations is a collection of m equations in the variables (unknowns) $x_1, x_2, x_3, \cdots, x_n$ of the form,

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

where $a_{ij}, b_i (1 \leq i \leq m, 1 \leq j \leq n)$ are given real numbers, x_j $1 \leq j \leq n$ is the unknown variables (also belong to the set of real numbers)

Remark: In this the course, we mainly focus on the case that all coefficients $a_{ij}, b_i (1 \leq i \leq m, 1 \leq j \leq n)$ are real numbers.

Example 1.4

$$
\begin{aligned}
x_1 - x_2 + 2x_3 &= 1, \\
2x_1 + x_2 + x_3 &= 8, \\
x_1 + x_2 &= 5.
\end{aligned}
$$

This is a linear system with 3 equations and 3 variables, and $m = 3, n = 3$. Moreover,

$$
\begin{aligned}
a_{11} = 1, a_{12} = -1, a_{13} = 2, b_1 &= 1, \\
a_{21} = 2, a_{22} = 1, a_{23} = 1, b_2 &= 8, \\
a_{31} = 1, a_{32} = 1, a_{33} = 0, b_3 &= 5.
\end{aligned}
$$

Solution of the System of Linear Equations

Given the system of linear equations:

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

where $a_{ij}, b_i (1 \leq i \leq m, 1 \leq j \leq n)$ are given real numbers, x_j $1 \leq j \leq n$ is the unknown variables (also belong to the set of real numbers)

If $m > n$, then the system is called **overdetermined system**.

If $m < n$, then the system is called **underdetermined system**.

overdetermined system and **underdetermined system** are just English name of the linear system, nothing related to their solution.

Definition 1.5 (Solution and Solution Set)

A **solution** of the system of linear equations is an ordered set $(s_1, s_2, \cdots, s_n)$ such that if we let $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$, then for every equation of the system the left-hand side (LHS) will equal the right-hand side (RHS), i.e. all the equations are true simultaneously.

The **solution set** of a linear system of equations is the set which contains every solution of the system, and nothing more. **Solution set** $S = \{(s_1, s_2, \cdots, s_n) | (s_1, s_2, \cdots, s_n) \text{ is a solution}\}$.

Possibilities for Solution Sets

Example 1.6
(i) The following linear system has **only one** solution.

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

The solution set is $(x_1, x_2) = \{(3, -1)\}$.
(ii) The following linear system has **infinitely many** solutions.

$$\begin{aligned} 2x_1 + 3x_2 &= 3, \\ 4x_1 + 6x_2 &= 6 \end{aligned}$$

The solution set is $(x_1, x_2) = \{(t, \frac{3-2t}{3}) | t \in \mathcal{R}\}$.

Possibilities for Solution Sets

(iii) The following linear system has **no** solution.

$$\begin{aligned} 2x_1 + 3x_2 &= 3, \\ 4x_1 + 6x_2 &= 10 \end{aligned}$$

The solution set is $(x_1, x_2) = \emptyset$, where $\emptyset$ is the empty set.

Remark: It is impossible for a system of linear equations to have exactly 2 solutions.

Equivalent Systems

Definition 1.7 (Equivalent Systems) Two linear systems are equivalent if their have the same solution sets.

Definition 1.8 (Equation Operations) Given a system of linear equations, three equation operations are defined as follows:
1. Swap the location of ith equation with the jth equation. Denote: $R_i \leftrightarrow R_j$. The reverse operation is $R_i \leftrightarrow R_j$.
2. Multiply each term of ith equation by a nonzero constant α. Denote: $R_i \rightarrow \alpha R_i$. The reverse operation is $R_i \rightarrow \frac{1}{\alpha} R_i$.
3. Multiply each term of ith equation by a constant β, and add these terms to jth equation, while keep the ith equation unchanged. ($i \neq j$). Denote: $R_j \rightarrow \beta R_i + R_j$. The reverse operation is $R_j \rightarrow -\beta R_i + R_j$.

Equation Operations Preserve Solution Sets

Proposition 1.9 If we apply one of the three equation operations of Definition 1.8 to a system of linear equations, then the original system and the transformed system are equivalent.

Observations

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

When solving the linear systems, we can take all the constants (coefficients) from a linear system in a nice way to form something called matrix, and then use this matrix to systematically solve the equations.

Definition 1.10 (**Matrix**) A $m \times n$ matrix is a rectangular array of numbers with m rows and n columns in the form:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \triangleq (a_{ij})_{m\times n}$$

or

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

where $a_{ij} \in \mathbb{R}$ or $a_{ij} \in \mathbb{C}$ ($\mathbb{R}$ is the set of real numbers, $\mathbb{C}$ is the set of complex numbers).

(i) For above matrix, a_{ij} is the (i, j)-**entry** of matrix A or the (i, j) **element** of matrix A.

(ii) Matrices usually denoted by $A, B, C, \cdots$.

(iii) When $a_{ij} \in \mathbb{R}$, we call A is a **real matrix**. When $a_{ij} \in \mathbb{C}$, we call A is a **complex matrix**.

(iv) When $m = n$, we call the matrix is **square matrix**.

(v) When all the entries of the matrix are zeros, the matrix is called **zero matrix**, denoted by $\mathbf{O}$.

Example 1.11

$$B = \begin{bmatrix} 1 & 2 & 5 & 3 \\ 1 & 0 & 6 & 1 \\ 4 & 2 & 2 & 2 \end{bmatrix}$$

is a matrix with $m = 3$ rows and $n = 4$ columns, where

$$b_{12} = 2, b_{23} = 6, b_{32} = 2, etc.$$

Definition 1.12 (row vector or column vector)

Let A be an $m \times n$ matrix, if $m = 1$, then A is a row vector. If $n = 1$, then A is a column vector.

(i) Columns vectors are denoted by $\mathbf{u}$, or $\underline{\mathrm{u}}$ or $\underset{\sim}{\mathrm{u}}$.

(ii) Row vectors are denoted by $\vec{\mathbf{u}}$ (this follows the notation of the textbook — this is not standard notation).

For row vector $\vec{\mathbf{u}}$, the element u_{1j} will be simply denoted by u_j (j-th entry of $\vec{\mathbf{u}}$). In the case of column vector $\mathbf{u}$, the element u_{i1} will be simply denoted by u_i (i-th entry of $\mathbf{u}$).

(iii) If $m = n = 1$, then $A = (a_{11})$. We will treat $A = a_{11}$ as a normal real number.

(iv) If all the entries of a vector are 0, then it is a zero vector, denoted by $\mathbf{0}$.

Remark. In most of this course, we will discuss the column vectors.

Definition 1.12 (Coefficient matrix) Given a linear system

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

The **coefficient matrix** of this linear system is the $m \times n$ matrix

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

Definition 1.13 (**Matrix Representation of a Linear System**) Let A be an $m \times n$ matrix and $\mathbf{b}$ be column vector of size m, where

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

The notation $A\mathbf{x} = \mathbf{b}$ (will be defined rigourously in later lectures) denotes the linear system

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

And the above linear system can be represented as $A\mathbf{x} = \mathbf{b}$

Definition 1.14 (**Augmented Matrix**) If the linear system has m equations and n variables, then the **augmented matrix** of the system is a $m \times (n+1)$ matrix, whose whose first n columns are the columns of A and whose last $(n+1)$-th column is the column vector $\mathbf{b}$. This matrix will be written as $[A \,|\, \mathbf{b}]$.

The augmented matrix for the above linear system is

$$[A \,|\, \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Example 1.15

$$
\begin{aligned}
2x_1 + 3x_2 - 5x_3 + 4x_4 - x_5 &= 6,\\
7x_1 - 2x_2 - 3x_3 - 4x_4 + 5x_5 &= 1,\\
-x_1 + x_2 + 9x_3 - 2x_4 + 3x_5 &= 0,
\end{aligned}
$$

has the coefficient matrix

$$
A = \begin{bmatrix} 2 & 3 & -5 & 4 & -1 \\ 7 & -2 & -3 & -4 & 5 \\ -1 & 1 & 9 & -2 & 3 \end{bmatrix}
$$

We can gather the RHS in a vector as follows:

$$
\mathbf{b} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}
$$

The augmented matrix then is

$$[A \,|\, \mathbf{b}] = \left[\begin{array}{rrrrr|r} 2 & 3 & -5 & 4 & -1 & 6 \\ 7 & -2 & -3 & -4 & 5 & 1 \\ -1 & 1 & 9 & -2 & 3 & 0 \end{array}\right]$$

and we can write the linear system as $A\mathbf{x} = \mathbf{b}$, where A is the coefficient matrix and $\mathbf{b}$ is as defined on the previous slide.

Definition 1.16 (Elementary Row Operations)

1. Swap the location of ith row with the jth row. Notation: $R_i \leftrightarrow R_j$. The reverse operation is $R_i \leftrightarrow R_j$.

2. Multiply each element of ith row by a nonzero constant α. Notation: $R_i \to \alpha R_i$ $(\alpha \neq 0)$.
The reverse operation is $R_i \to \frac{1}{\alpha} R_i$ $(\alpha \neq 0)$

3. Multiply each element of ith row by a constant β, and add these terms to jth row, while keep the ith row unchanged. $(i \neq j)$. Notation: $R_j \to \beta R_i + R_j$.
The reverse operation is $R_j \to -\beta R_i + R_j$.

Remark. The reverse row operation belongs to the same type of elementary row operation.

Definition 1.17 (**Row-Equivalent Matrices**) Two matrices are said to be row equivalent if one can obtained from the other by a sequence of elementary row operations.

Theorem 1.18 (**Equivalent Linear Systems**)
Consider two linear systems. If the augmented matrices of the two linear systems are row-equivalent, then the two linear systems are equivalent and they have the same solution set. Row operations preserve the solution set of the linear system.

Solving $n \times n$ linear systems

Basic idea:
Forward substitution:

1. Begin with a system of equations, represent the system by an augmented matrix.

2. Perform elementary row operations and try to get a "simpler" matrix (the upper triangular matrix).

Backward substitution:
3. Convert back to a simpler system of equations and then solve that system (using back substitution), knowing that its solutions are those of the original system.

Example 1.19 The augmented matrix for the linear system

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 4, \\ x_1 + 3x_2 + 3x_3 &= 5, \\ 2x_1 + 6x_2 + 5x_3 &= 6. \end{aligned}$$

is

$$[A \,|\, \mathbf{b}] = \left[\begin{array}{ccc|c} \boxed{1} & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 \\ 2 & 6 & 5 & 6 \end{array}\right] \begin{array}{l} \textit{pivotal} \quad \textit{row} \\ \\ \\ \end{array}$$

Here we can chose 1 as the pivot (the first nonzero entry in the pivotal row) to eliminate the entries in the first column and row 2&3.

$$
\left[\begin{array}{ccc|c} \boxed{1} & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 \\ 2 & 6 & 5 & 6 \end{array}\right] \quad \textit{pivotal} \quad \textit{row}
$$

$$
\xrightarrow[R_3 \to -2R_1 + R_3]{R_2 \to -R_1 + R_2} \left[\begin{array}{ccc|c} 1 & 2 & 2 & 4 \\ 0 & \boxed{1} & 1 & 1 \\ 0 & 2 & 1 & -2 \end{array}\right] \quad \textit{pivotal} \quad \textit{row}
$$

$$
\xrightarrow{R_3 \to -2R_2 + R_3} \left[\begin{array}{ccc|c} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & -4 \end{array}\right]
$$

The above process is called **Gaussian elimination**.

This augmented matrix corresponds to the linear system

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

We can solve this linear system easily using **back substitution:**

$$x_3 = 4, x_2 = -3, x_1 = 2.$$

We can also view back substitution as elementary row operations!

$$\left[\begin{array}{ccc|c} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & -4 \end{array}\right] \xrightarrow{R_3 \to -R_3} \left[\begin{array}{ccc|c} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & \boxed{1} & 4 \end{array}\right] \quad \textit{pivotal} \quad \textit{row}$$

$$\xrightarrow[R_1 \to -2R_3 + R_1]{R_2 \to -R_3 + R_2} \left[\begin{array}{ccc|c} 1 & 2 & 0 & -4 \\ 0 & \boxed{1} & 0 & -3 \\ 0 & 0 & 1 & 4 \end{array}\right] \quad \textit{pivotal} \quad \textit{row}$$

$$\xrightarrow{R_1 \to -2R_2 + R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

The solution can be red-off from the last column.