

Lecture 1 Systems of Linear Equations, Matrices

Definition 1.1 An example of a *linear equation* is

$$x + 2y = 4.$$

Generally, we say that

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

where a_1, \dots, a_n, b are real constants, is a linear equation in n variables x_1, \dots, x_n . The n variables are sometimes also called *unknowns*.

Example 1.2 We are most familiar with linear equations with 2 or 3 unknowns. For example,

$$x + 2y = 4$$

is a straight line in the two dimensional space. While

$$z = 0$$

or equivalently $0x + 0y + z = 0$ is a plane in the xy plane contained in the three dimensional space.

Definition 1.3 A set of real numbers s_1, \dots, s_n is said to be a *solution* of the linear equation

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

if

$$a_1s_1 + a_2s_2 + \dots + a_ns_n = b.$$

Exercise 1.4 In 20 seconds, write down as many solutions of (a) $x + 2y = 4$ and (b) $z = 0$ as you can.

Definition 1.5 The set of all solutions of the equation is called the *solution set* of the equation and an expression that gives us all these solutions is called the *general solution* of the equation.

Example 1.6 The solution set of $x + 2y = 4$ can be written as

$$\left\{ \left(t, \frac{4-t}{2} \right) \mid t \in \mathbb{R} \right\}.$$

t is sometimes called a *parameter*. A general solution of the equation can be written as

$$\begin{cases} x = t \\ y = \frac{4-t}{2}, \quad t \in \mathbb{R}. \end{cases}$$

Notice that by substituting different values of t into $(t, \frac{4-t}{2})$, we can get different solutions of the equation. How many solutions are there?

Exercise 1.7 Write down the solution set and a general solution of $z = 0$.

Definition 1.8 A finite set of linear equations in the variables x_1, x_2, \dots, x_n is called a *system of linear equations* (or a *linear system*):

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

where $a_{11}, a_{12}, \dots, a_{mn}$ and b_1, b_2, \dots, b_m are real constants. We say the above linear system has n unknowns and m equations. The definitions of *solution set* and *general solution* for a linear equation can be extended to a linear system.

Example 1.9 Consider the following system of linear equations:

$$\begin{cases} 4x_1 - x_2 + 3x_3 = -1 \\ 3x_1 + x_2 + 9x_3 = -4 \end{cases}$$

$x_1 = 1, x_2 = 2, x_3 = -1$ is a solution of the system but $x_1 = 1, x_2 = 8, x_3 = 1$ is not. A general solution of the linear system is

$$\begin{cases} x_1 = \frac{-5-12t}{7} \\ x_2 = \frac{-13-27t}{7} \\ x_3 = t, \quad t \in \mathbb{R}. \end{cases}$$

Remark 1.10 Notice that there is one parameter in the general solution of the linear system above. How many parameters are there in the general solution of $z = 0$? Later on, we will develop a systematic method to obtain a general solution of a linear system (if it exists).

Remark 1.11 Not all linear systems have solutions. Can you write down a linear system with 2 equations and 2 unknowns that has no solutions?

Definition 1.12 A linear system that has no solutions is said to be *inconsistent* while one that has at least one solution is said to be *consistent*

Discussion 1.13 In the xy -plane, the linear system

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$$

represents two straight lines. Discuss the possible geometrical positions of the two lines and the respective solution sets.

Exercise 1.14 Write down a linear system representing three planes in the three dimensional space. Discuss possible geometrical positions of the three planes and the respective solution sets.

Theorem 1.15 Every linear system has either *no solution*, *exactly one solution*, or *infinitely many solutions*.

Definition 1.16 A linear system

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

can be represented by a rectangular array of numbers

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right).$$

This array is called the *augmented matrix* of the linear system.

Example 1.17 The augmented matrix of the linear system

$$\begin{cases} 4x_1 - x_2 + 3x_3 = -1 \\ 3x_1 + x_2 + 9x_3 = -4 \end{cases}$$

is

$$\left(\begin{array}{ccc|c} 4 & -1 & 3 & -1 \\ 3 & 1 & 9 & -4 \end{array} \right).$$

Definition 1.18 In general, any rectangular array of numbers is called a *matrix* (plural *matrices*). The numbers in the array are called the *entries* of the matrix. The *size* of the matrix is given by $m \times n$ where m is the number of rows and n is the number of columns in the matrix. The (i, j) -entry of the matrix is the number in the i th row and j th column.

Definition 1.19 A *column matrix* (or column vector) is a matrix with only one column while a *row matrix* (or row vector) is one with only one row.

Notation 1.20 Instead of writing a $m \times n$ matrix \mathbf{A} as

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

we may instead write $\mathbf{A} = (a_{ij})_{m \times n}$ where a_{ij} is the (i, j) -entry of \mathbf{A} . If there is no ambiguity regarding the size of the matrix, we simply write $\mathbf{A} = (a_{ij})$.

Definition 1.21 The following are some some special types of matrices:

1. A matrix is *square* if it has the same number of columns and rows. A square matrix of size $n \times n$ is said to be of *order* n .
2. A *diagonal* matrix is a square matrix \mathbf{A} where $a_{ij} = 0$ whenever $i \neq j$. If \mathbf{A} is of order n , then the *diagonal* of \mathbf{A} is the sequence of entries $a_{11}, a_{22}, \dots, a_{nn}$. The entries a_{ii} are called the *diagonal* entries while $a_{ij}, i \neq j$ are the *non-diagonal* entries.
3. A *scalar* matrix is a diagonal matrix such that $a_{11} = a_{22} = \dots = a_{nn} = c$ for some constant c . If $c = 1$, then the scalar matrix is called the *identity* matrix. We use \mathbf{I}_n to denote an identity matrix of order n . If there is no ambiguity, we may also use \mathbf{I} to denote an identity matrix.
4. A *zero* matrix is one where $a_{ij} = 0$ for all i, j . We denote the $m \times n$ zero matrix by $\mathbf{0}_{m \times n}$ or simply $\mathbf{0}$.
5. A square matrix \mathbf{A} is *symmetric* if $a_{ij} = a_{ji}$ for all i, j .
6. A square matrix \mathbf{A} is *upper triangular* if $a_{ij} = 0$ whenever $i > j$. \mathbf{A} is said to be *lower triangular* if $a_{ij} = 0$ whenever $i < j$. Both lower and upper triangular matrices are called *triangular* matrices.

Definition 1.22 Two matrices are said to be *equal* if (a) they have the same size; (b) each of their corresponding entries are equal. That is, two matrices \mathbf{A} and \mathbf{B} are equal if and only if the sizes of \mathbf{A} and \mathbf{B} are the same and also $a_{ij} = b_{ij}$ for all i, j .

Definition 1.23 Let $\mathbf{A} = (a_{ij})_{m \times n}$, $\mathbf{B} = (b_{ij})_{m \times n}$ and c be a constant. We define the matrices $\mathbf{A} + \mathbf{B}$, $\mathbf{A} - \mathbf{B}$ and $c\mathbf{A}$ as follows.

1. *Matrix addition:* $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})_{m \times n}$.
2. *Matrix subtraction:* $\mathbf{A} - \mathbf{B} = (a_{ij} - b_{ij})_{m \times n}$.
3. *Scalar multiplication:* $c\mathbf{A} = (ca_{ij})_{m \times n}$.

Example 1.24 Suppose \mathbf{A} and \mathbf{B} are matrices of the same size. Prove that $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.

Solution Since \mathbf{A} and \mathbf{B} are matrices of the same size, $\mathbf{A} + \mathbf{B}$ and $\mathbf{B} + \mathbf{A}$ also have the same size. Next, consider the (i, j) -entry of $\mathbf{A} + \mathbf{B}$.

$$\begin{aligned} (i, j)\text{-entry of } \mathbf{A} + \mathbf{B} &= a_{ij} + b_{ij} \\ &= b_{ij} + a_{ij} \\ &= (i, j)\text{-entry of } \mathbf{B} + \mathbf{A} \end{aligned}$$

Thus we have shown that $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.

Exercise 1.25 If \mathbf{A} , \mathbf{B} and \mathbf{C} are matrices of the same size and a, b are scalars. Prove that the following statements are true: (a) $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$; (b) $a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$; (c) $(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$; (d) $a(b\mathbf{A}) = (ab)\mathbf{A} = b(a\mathbf{A})$.

Definition 1.26 Suppose \mathbf{A} is a $m \times n$ matrix. Then the *transpose* of \mathbf{A} , denoted by \mathbf{A}^T , is a $n \times m$ matrix obtained from \mathbf{A} by interchanging the rows and the columns. Note that if \mathbf{A} is a square matrix, then $\mathbf{A} = \mathbf{A}^T$ if and only if \mathbf{A} is symmetric.

Example 1.27 If

$$\mathbf{A} = \begin{pmatrix} 1 & -4 & 3 & 0 \\ 2 & 0 & 0 & 3 \end{pmatrix},$$

then

$$\mathbf{A}^T = \begin{pmatrix} 1 & 2 \\ -4 & 0 \\ 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

Definition 1.28 Let $\mathbf{A} = (a_{ij})_{m \times n}$ and $\mathbf{B} = (b_{ij})_{n \times p}$ be two matrices. Then the *product* \mathbf{AB} is defined to be a $m \times p$ matrix whose (i, j) -entry is given by

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj},$$

for $i = 1, \dots, m$, $j = 1, \dots, p$.

Remark 1.29 From the definition above, we see that for \mathbf{AB} to be defined, the number of columns in \mathbf{A} must equal to the number of rows in \mathbf{B} .

Example 1.30 Suppose $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 0 & -2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 1 & 4 \\ 3 & -2 & 1 \end{pmatrix}$. Then

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} 1 \cdot 0 + 1 \cdot 3 & 1 \cdot 1 + 1 \cdot (-2) & 1 \cdot 4 + 1 \cdot 1 \\ 1 \cdot 0 + 3 \cdot 3 & 1 \cdot 1 + 3 \cdot (-2) & 1 \cdot 4 + 3 \cdot 1 \\ 0 \cdot 0 + (-2) \cdot 3 & 0 \cdot 1 + (-2) \cdot (-2) & 0 \cdot 4 + (-2) \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -1 & 5 \\ 9 & -5 & 7 \\ -6 & 4 & -2 \end{pmatrix} \end{aligned}$$

Exercise 1.31 Let $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$. Compute both \mathbf{AB} and \mathbf{BA} .

Discussion 1.32 From the exercise above, we see that it is unclear what we mean when we say "multiply \mathbf{A} and \mathbf{B} ". Thus, we refer to \mathbf{AB} as the *pre-multiplication* of \mathbf{A} to \mathbf{B} and \mathbf{BA} as the *post-multiplication* of \mathbf{A} to \mathbf{B} .

Discussion 1.33 In algebra, we often use the following argument: "If $xy = 0$, then either $x = 0$ or $y = 0$ ". Now, if \mathbf{A} and \mathbf{B} are matrices such that \mathbf{AB} is defined. Furthermore, suppose $\mathbf{AB} = \mathbf{0}$. Does this imply either $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$?

Theorem 1.34 Rules of matrix multiplication and manipulation.

1. If \mathbf{A} , \mathbf{B} and \mathbf{C} are $m \times p$, $p \times q$ and $q \times n$ matrices respectively, then $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
2. If \mathbf{A} , \mathbf{B}_1 and \mathbf{B}_2 are $m \times p$, $p \times n$ and $p \times n$ matrices respectively, then $\mathbf{A}(\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{AB}_1 + \mathbf{AB}_2$.
3. If \mathbf{A} , \mathbf{C}_1 and \mathbf{C}_2 are $p \times n$, $m \times p$ and $m \times p$ matrices respectively, then $(\mathbf{C}_1 + \mathbf{C}_2)\mathbf{A} = \mathbf{C}_1\mathbf{A} + \mathbf{C}_2\mathbf{A}$.
4. If c is a scalar and \mathbf{A} , \mathbf{B} are $m \times p$, $p \times n$ matrices respectively, then $a(\mathbf{AB}) = (a\mathbf{A})\mathbf{B} = \mathbf{A}(a\mathbf{B})$.
5. If \mathbf{A} is a $m \times n$ matrix, then $\mathbf{A}\mathbf{0}_{n \times q} = \mathbf{0}_{m \times q}$ and $\mathbf{0}_{p \times m}\mathbf{A} = \mathbf{0}_{p \times n}$.
6. If \mathbf{A} is a $m \times n$ matrix, then $\mathbf{A}\mathbf{I}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A}$.
7. $(\mathbf{A}^T)^T = \mathbf{A}$.
8. If \mathbf{A} and \mathbf{B} are both $m \times n$ matrices, then $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.
9. If a is a scalar, then $(a\mathbf{A})^T = a\mathbf{A}^T$.

10. If \mathbf{A} and \mathbf{B} are $m \times n$ and $n \times p$ matrices respectively, then $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$.

Notation 1.35 Let

$$\mathbf{A} = (a_{ij})_{m \times p} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix}$$

where $\mathbf{a}_i = (a_{i1} \ a_{i2} \ \cdots \ a_{ip})$ is the i th row of \mathbf{A} . Let

$$\mathbf{B} = (b_{ij})_{p \times n} = (\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n)$$

where $\mathbf{b}_j = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix}$ is the j th column of \mathbf{B} . Then

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \mathbf{b}_1 & \mathbf{a}_1 \mathbf{b}_2 & \cdots & \mathbf{a}_1 \mathbf{b}_n \\ \mathbf{a}_2 \mathbf{b}_1 & \mathbf{a}_2 \mathbf{b}_2 & \cdots & \mathbf{a}_2 \mathbf{b}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_m \mathbf{b}_1 & \mathbf{a}_m \mathbf{b}_2 & \cdots & \mathbf{a}_m \mathbf{b}_n \end{pmatrix} \quad (1)$$

where $\mathbf{a}_r \mathbf{b}_s = (a_{r1} \ a_{r2} \ \cdots \ a_{rp}) \begin{pmatrix} b_{1s} \\ b_{2s} \\ \vdots \\ b_{ps} \end{pmatrix} = \left(\sum_{k=1}^p a_{rk} b_{ks} \right)$ which is a 1×1

matrix. Also we can write

$$\mathbf{AB} = (\mathbf{Ab}_1 \ \mathbf{Ab}_2 \ \cdots \ \mathbf{Ab}_n), \quad (2)$$

where $\mathbf{A}\mathbf{b}_j$ is the j th column of $\mathbf{A}\mathbf{B}$, or

$$\mathbf{A}\mathbf{B} = \begin{pmatrix} \mathbf{a}_1\mathbf{B} \\ \mathbf{a}_2\mathbf{B} \\ \vdots \\ \mathbf{a}_m\mathbf{B} \end{pmatrix}, \quad (3)$$

where $\mathbf{a}_i\mathbf{B}$ is the i th row of $\mathbf{A}\mathbf{B}$.

Exercise 1.36 Let

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

$$\mathbf{a}_1 = (1 \ 2 \ 3), \quad \mathbf{a}_2 = (4 \ 5 \ 6),$$

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}.$$

Note that

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = (\mathbf{b}_1 \ \mathbf{b}_2).$$

Work out the RHS of (1), (2) and (3) to show that they are all equal to

$$\begin{pmatrix} 2 & 1 \\ 8 & 7 \end{pmatrix}.$$

Remark 1.37 Using Notation 1.35, any linear system

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

can be written in matrix form as

$$\mathbf{Ax} = \mathbf{b}$$

where

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