Lecture Note on Linear Algebra*

1. Systems of Linear Equations

Wei-Shi Zheng, 2011

1 Why Learning Linear Algebra(线性代数)?

Solving linear equation system is the heart of linear algebra. Linear algebra is widely used in computer science and electronic engineering, e.g. artificial intelligence (人工智能), pattern recognition(模式识别), computer vision(计算机视觉), data mining(数据挖掘), machine learning(机器学习). It is fundamental of many other subjects/courses and research approaches.

2 What Do You Learn from This Note

Basic concept about linear equation(线性方程), system of linear equations(线性方程组), matrix(矩阵) and its solution(解).

3 What Is System of Linear Equations?

Let us begin with an introduction of solving systems of linear equations. We first show some examples of linear equations:

$$x = 0,\ -2x = -1,\ 3x = 4(y - 9),\ x_1 - 1.1x_2 = \frac{1}{3}(x_3 + \mathrm{e}).$$

In general, linear equation is defined as follows:

Definition 1 (linear equation(线性方程)). A linear equation with variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

$$a_1x_1 + \cdots + a_nx_n = b$$

*The first version of this lecture note was edited by Dr. Jialun Huang.

where b and the **coefficients** (系数) $a_1, a_2, \ldots, a_n$ are real (实数) or complex (虚数) numbers known in advance.

Definition 2 (system of linear equations(线性方程组)). A system of linear equations with variables (变量) $x_1, x_2, \ldots, x_n$ is a collection of finite linear equations with variables $x_1, x_2, \ldots, x_n$.

Examples:

$$\left\{\begin{array}{rrcl} -2x_1 & -3x_2 & = & 5 \\ x_1 & +2x_2 & = & 8 \\ 2x_1 & +4x_2 & = & 0 \end{array}\right. , \quad \left\{\begin{array}{rrrcl} x_1 & +4x_2 & -x_3 & = & -1 \\ \frac{2}{5}x_1 & +5x_2 & -x_3 & = & 0 \end{array}\right. .$$

The general form of a system of m linear equations with n variables is:

$$\left\{\begin{array}{rrcrcl} a_{11}x_1 & +a_{12}x_2 & +\cdots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\cdots & +a_{2n}x_n & = & b_2 \\ \cdots & & & & & \\ a_{m1}x_1 & +a_{m2}x_2 & +\cdots & +a_{mn}x_n & = & b_m \end{array}\right. \tag{1}$$

4 Solution of System of Linear Equations

Definition 3 (solution(解)). A list $(s_1, s_2, \ldots, s_n)$ of numbers is called a solution of (1) iff (i.e. if and only if) all the equations in (1) are satisfied by substituting $s_1, s_2, \ldots, s_n$ for $x_1, x_2, \ldots, x_n$. The set of all solutions of (1) is called the **solution set** (解集) of (1). Two systems of linear equations are said to be **equivalent** (等价) if they have the same solution set.

Does a system of linear equations always have a solution? Let us investigate some examples as follows:
Examples:

$$\left\{\begin{array}{rrcl} x_1 & -2x_2 & = & -1 \\ -x_1 & +3x_2 & = & 3 \end{array}\right. , \left\{\begin{array}{rrcl} x_1 & -2x_2 & = & -1 \\ -x_1 & +2x_2 & = & 3 \end{array}\right. , \left\{\begin{array}{rrcl} x_1 & -2x_2 & = & -1 \\ -x_1 & +2x_2 & = & 1 \end{array}\right. .$$

From the above examples, we see that not all systems of linear equations have a solution. Generally speaking, a system of linear equations has either:
1. no solutions (无解), i.e. the solution set is empty (empty set), or
2. exactly one solution(唯一解), i.e. the solution set contains only one element (singleton set), or
3. infinitely many solutions(无限多解), i.e. the solution set contains infinitely many elements (infinite set).

DEFINITION 4 (consistence (相容)). A system of linear equations is said to be consistent if its solution set is non–empty (i.e. either one solution or infinitely many solutions), otherwise it is inconsistent.

5 Matrix

DEFINITION 5 (matrix (矩阵)). A table of numbers with m rows (行) and n columns (列) as above is called an $m \times n$ matrix. we normally use a capital letter such as A, B, X etc. to denote a matrix.

We define the $m \times (n+1)$ matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

to be the corresponding **augmented matrix** (增广矩阵) of system (1), where we call the corresponding m rows and n columns:

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

as the **coefficient matrix** (系数矩阵) of system (1).

Accordingly, we generate a one to one correspondence (一一对应) between systems and matrices.

Example:

$$\left\{\begin{array}{rrrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ -4x_1 & +5x_2 & +9x_3 & = & -9 \end{array}\right. \qquad \longleftrightarrow \qquad \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{pmatrix}.$$

6 Solving a Linear System

We are now introducing a procedure for solving systems of linear equations.

Basic strategy (基本策略). The basic strategy is to replace one system with an equivalent system (i.e. one with the same solution set) that is easier to solve. For example, if a linear system consists of three variables x_1, x_2 and x_3, then use the x_1 term in the first equation of a system to eliminate the x_1 terms in the other equations. Then use the x_x term in the second equation of a system to eliminate the x_1 terms in the other equations, and so on, until you finally obtain a very simple equivalent system of equations.

A system can be transformed into another equivalent system such that some variable is eliminated in some equation by applying an elementary operation(初等变换). After applying a series of elementary operations to the original system, the original system is transformed into an easy–to–solve system. The following is a special example:

$$\left\{\begin{array}{rrrcr} 3x_1 & +x_2 & +x_3 & = & 5 \\ & 2x_2 & +x_3 & = & 3 \\ & & x_3 & = & 1 \end{array}\right. .$$

Furthermore, by applying elementary operations, the above system can be transformed into the trivial system

$$\left\{\begin{array}{rrrcr} x_1 & & & = & 1 \\ & x_2 & & = & 1 \\ & & x_3 & = & 1 \end{array}\right. .$$

Operations. Formally, three types of elementary operations on systems of linear equations are necessary for us to realize the above transformation. Let S bs a linear system consists of m equations and n variables. Let e_i denote the i–th equation of S. The three types of elementary operations are defined

as follows:
1. *Interchange(对换变换)*: Perform exchange between the i–th equation and the j–th equation of S ($e_i \leftrightarrow e_j$);
2. *Scaling(倍乘变换)*: Multiply the i–th equation by a nonzero number λ ($e_i := \lambda e_i$);
3. *Replacement(倍加变换)*: Add the result of multiplying the j–th equation by a number λ to the i–th equation ($e_i := e_i + \lambda e_j$).

To solve the linear system S, we shall perform a sequence of elementary operations, resulting in a set of equivalent linear systems

$$S = S_0 \xrightarrow{op_1} S_1 \xrightarrow{op_2} \cdots \xrightarrow{op_l} S_l$$

systematically, such that S_l is an easy–to–solve or even a trivial system.

Example:

$$\left\{\begin{array}{rrrcl} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ -4x_1 & +5x_2 & +9x_3 & = & -9 \end{array}\right. \xrightarrow{e_3:=e_3+4e_1} \left\{\begin{array}{rrrcl} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ & -3x_2 & +13x_3 & = & -9 \end{array}\right. \xrightarrow{e_3:=e_3+\frac{3}{2}e_1}$$

$$\left\{\begin{array}{rrrcl} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ & & x_3 & = & 3 \end{array}\right. \xrightarrow{e_2:=e_2+8e_3} \left\{\begin{array}{rrrcl} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & & = & 32 \\ & & x_3 & = & 3 \end{array}\right. \xrightarrow{e_2:=\frac{1}{2}e_2}$$

$$\left\{\begin{array}{rrrcl} x_1 & -2x_2 & +x_3 & = & 0 \\ & x_2 & & = & 16 \\ & & x_3 & = & 3 \end{array}\right. \xrightarrow{e_1:=e_1-e_3} \left\{\begin{array}{rrrcl} x_1 & -2x_2 & & = & -3 \\ & x_2 & & = & 16 \\ & & x_3 & = & 3 \end{array}\right. \xrightarrow{e_1:=e_1+2e_2}$$

$$\left\{\begin{array}{rrrcl} x_1 & & & = & 29 \\ & x_2 & & = & 16 \\ & & x_3 & = & 3. \end{array}\right.$$

Theorem 6. Elementary operations are reversible(可逆的). That is, if $S_1 \xrightarrow{op} S_2$ where S_1, S_2 are systems and *op* is an elementary operation, then there is another elementary operation op^{-1} such that $S_2 \xrightarrow{op^{-1}} S_1$.

Proof. Exercise. □

Connection to Row Operations. In terms of matrix representation for a linear system of equations, the corresponding to elementary operations on

linear systems can be directly performed as elementary row operations (行变换) on matrices as follows:
1. *Interchange:* Exchange the i–th row and the j–th row of a matrix ($r_i \leftrightarrow r_j$);
2. *Scaling*: Multiply the i–th row of a matrix by a nonzero number λ ($r_i := \lambda r_i$);
3. *Replacement*: Add the result of multiplying the j–th row by a number λ to the i–th row of a matrix ($r_i := r_i + \lambda r_j$).

DEFINITION 7 (Row Equivalence (行等价)). If matrix A can be transformed into matrix B by applying a series of elementary row operations on A then we say A is row equivalent to B and denote this equivalence by $A \sim B$.

Obviously, $A \sim B$ if and only if their corresponding systems are equivalent.

From now on, we always represent a system of linear equations by its corresponding augmented matrix. We shall use the more convenient matrix language in the remaining part of this course.

Reference

David C. Lay. Linear Algebra and Its Applications (3rd edition). Pages 1∼13.