Lecture Note on Linear Algebra

3. Vector Equations

Wei-Shi Zheng, 2011

1 What Do You Learn from This Note

Recall that we have ever said that $\mathbb{R}^2$, which is the set of plane vectors, is a concrete (具体) example of vector space. In this lecture, we will look at some detailed properties of $\mathbb{R}^2$. Further, we will generalize the properties of $\mathbb{R}^2$ to $\mathbb{R}^n$, which is the set of vectors of dimension n over $\mathbb{R}$. We shall see quickly that any system of linear equations is equivalent to a so called vector equation. The study of $\mathbb{R}^n$ will help you to understand the abstract concept of vector spaces which are the major subject studied in linear algebra.

Basic concept: column vector (列向量), linear combination (线性组合), Span (张)

2 Vectors

Definition 1 (vectors (向量) in $\mathbb{R}^2$). A 2×1 matrix

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

is called a column vector (列向量) (or simply a vector) of dimension 2 over $\mathbb{R}$, where v_1, v_2 are real numbers. The set of all such vectors is denoted by $\mathbb{R}^2$. Similarly, we can define row vectors.

Let $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ and $c \in \mathbb{R}$. We have the following definitions on $\mathbb{R}^2$:

1. Equality(相等): we say $\vec{u}$ and $\vec{v}$ are equal, written $\vec{u} = \vec{v}$, iff $u_1 = v_1$ and $u_2 = v_2$.

2. Addition(加): the vector $\begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}$ is called the sum of $\vec{u}$ and $\vec{v}$ and is denoted by $\vec{u} + \vec{v}$.

3. Scalar Multiplication(数乘): the vector $\begin{pmatrix} c \cdot u_1 \\ c \cdot u_2 \end{pmatrix}$ is called the scalar multiple of $\vec{u}$ by scalar c and is denoted by $c\vec{u}$.

Example: Let $\vec{u} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 2 \\ -5 \end{pmatrix}$, find $4\vec{u}$, $-3\vec{v}$ and $4\vec{u} + (-3)\vec{v}$.

Remarks: 1. The vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is called the zero vector and is denoted by $\vec{0}$.
2. For the sake of simplicity, we normally write $-\vec{u}$ for $(-1)\vec{u}$ which is called the additive inverse of $\vec{u}$ and $\vec{u} - \vec{v}$ for $\vec{u} + (-1)\vec{v}$ which is called the difference of $\vec{u}$ by $\vec{v}$.

Theorem 2. Let $\vec{u}$, $\vec{v}$, $\vec{w} \in \mathbb{R}^2$ and c, $d \in \mathbb{R}$. Then
1. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ (Additive Associativity,加法结合率);
2. $\vec{0} + \vec{u} = \vec{u} + \vec{0}$ (Additive Identity加法单位元素);
3. $(-\vec{u}) + \vec{u} = \vec{u} + (-\vec{u}) = \vec{0}$ (Additive Inverse,加法逆运算);
4. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ (Additive commutativity,加法交换性);
(properties 1–4 are called the properties of addition for abelian group (阿贝尔群))
5. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$;
6. $(c + d)\vec{u} = c\vec{u} + d\vec{u}$;
7. $c(d\vec{u}) = (cd)\vec{u}$;
8. $1\vec{u} = \vec{u}$.

Proof. Exercise. □

Suppose that we have created a coordinate system (坐标系统) on a plane. Geometrically, vector $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ have 2 interpretations on this plane:

1. A point on the plane (平面上的一点), vector $\vec{v}$ can be interpreted as a point with coordinate (v_1, v_2). This is a one to one correspondence between points and vectors.

2. A directed segment(线段) (or arrow) with start point (s_1, s_2) and end point (e_1, e_2) such that $v_1 = e_1 - s_1$, $v_2 = e_2 - s_2$. Note that this correspondence is one to many (一对多) rather that one to one (一对一).

Parallelogram Rule (平行四边形法则) for addition:

We now define vectors in $\mathbb{R}^n$ in the same way as in $\mathbb{R}^2$.

Definition 3 (vectors in $\mathbb{R}^n$). A $n \times 1$ matrix

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

is called a column vector (or simply a vector) of dimension n over $\mathbb{R}$, where $v_1, \dots, v_n$ are real numbers. The set of all such vectors is denoted by $\mathbb{R}^n$. (Similarly, we can define row vectors).

Analogous to $\mathbb{R}^2$, we can define 'Equality', 'Addition', and 'Scalar Multiplication' on $\mathbb{R}^n$ in a similar way and Theorem 2 also holds for $\mathbb{R}^n$.

Next, we shall see how to connect vectors in $\mathbb{R}^n$ to systems of linear equations, we first introduce the following:

3 Linear Combinations (线性组合)

Definition 4 (linear combination(线性组合)). Given $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in \mathbb{R}^m$ and $c_1, c_2, \dots, c_n \in \mathbb{R}$. Then we can define a new vector

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n \Big(= \sum_{i=1}^{n} c_i\vec{v}_i \Big),$$

which is called a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ with coefficients (or weights) $c_1, c_2, \dots, c_n$.

Example: Let $\vec{u} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 2 \\ -5 \end{pmatrix}$. Then

$$4\vec{u} - 3\vec{v} = \begin{pmatrix} 4 - 3 \cdot 2 \\ 4(-2) - 3(-5) \end{pmatrix} = \begin{pmatrix} -2 \\ 7 \end{pmatrix}$$

is a linear combination of $\vec{u}$ and $\vec{v}$.

Example: Define $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \vec{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^m$. Then we have $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = c_1\vec{e}_1 + \cdots + c_m\vec{e}_m$. This means that any vector in $\mathbb{R}^m$ is a linear combination of $\vec{e}_1, \ldots, \vec{e}_m$.

Example: Let $\vec{a}_1 = \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix}$, $\vec{a}_2 = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} 7 \\ 4 \\ -3 \end{pmatrix}$. Determine whether or not $\vec{b}$ is a linear combination of $\vec{a}_1$ and $\vec{a}_2$.

Solution. Suppose that $x_1, x_2 \in \mathbb{R}$ such that $\vec{b} = x_1\vec{a}_1 + x_2\vec{a}_2$. Then we have

$$\begin{pmatrix} 7 \\ 4 \\ -3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{pmatrix}$$

By equality of vectors, we have

$$\left\{ \begin{array}{rrcr} x_1 & +2x_2 & = & 7 \\ -2x_1 & +5x_2 & = & 4 \\ -5x_1 & +6x_2 & = & -3 \end{array} \right. .$$

So if we regard x_1, x_2 as unknowns, then $\vec{b}$ is a linear combination of $\vec{a}_1$ and $\vec{a}_2$ if and only if the above system of linear equations with augmented matrix $(\vec{a}_1\ \vec{a}_2\ \vec{b})$ is consistent.

Definition 5 (vector equation (向量方程)). Let $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n, \vec{b} \in \mathbb{R}^m$. Then $x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}$ is called a vector equation with variables $x_1, x_2, \ldots, x_n$.

From the above example, it is easy to see that the solution set of $x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}$ is the same as the system of linear equations with augmented matrix $(\vec{a}_1\ \vec{a}_2\ \cdots\ \vec{a}_n\ \vec{b})$.

One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed (固定) set $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathbb{R}^m$ of vectors.

Definition 6 (Span). Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathbb{R}^m$. The set of all linear combinations of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is called the set generated (or spanned, 张成) by $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ and is denoted by $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$, which is a subset of $\mathbb{R}^m$. We also say $\vec{w} \in \mathbb{R}^m$ can be generated by $\vec{v}_1, v_2, \ldots, \vec{v}_n$ if $\vec{w} \in \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$.

Theorem 7. Let $W = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Then
1. $\vec{0} \in W$;
2. $\vec{v}_i \in W$ for $i = 1, \ldots, n$;
3. $\vec{w} \in W$ implies $\lambda\vec{w} \in W$, where $\lambda \in \mathbb{R}$;
4. $\vec{w}_1, \vec{w}_2 \in W$ implies $\vec{w}_1 + \vec{w}_2 \in W$.

Proof. 1. & 2. are obvious.
3. Suppose $\vec{w} = \sum_{i=1}^n c_i v_i$. Then $\lambda\vec{w} = \sum_{i=1}^n (\lambda c_i)\vec{v}_i \in W$.
4. Suppose $\vec{w}_1 = \sum_{i=1}^n c_{1i}\vec{v}_i$ and $\vec{w}_2 = \sum_{i=1}^n c_{2i}\vec{v}_i$. Then $\vec{w}_1 + \vec{w}_2 = \sum_{i=1}^n (c_{1i} + c_{2i})\vec{v}_i \in W$. □

Reference

David C. Lay. Linear Algebra and Its Applications (3rd edition). Pages 28∼40.

The Lady of Shallot, *by Waterhouse*