Lecture Note on Linear Algebra
14. Linear Independence, Bases and Coordinates

Wei-Shi Zheng,
wszheng@ieee.org, 2011

November 3, 2011

1 What Do You Learn from This Note

Do you still remember the unit vectors we have introduced in Chapter 1:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{1}$$

They are a set of vectors which are the basis (specially the standard basis) of $\mathbb{R}^3$. In this lecture note, we give a more general definition of basis

Basic Concept: Basis(基), standard basis(标准基), coordinate system(坐标系), coordinate vector(坐标向量)，coordinate mapping(坐标映射)

2 Basis(基)

Definition 1 (linear independence)**.** Let V be a vector space. Vectors $\vec{v}_1, \ldots, \vec{v}_n \in V$ are said to be linearly independent iff (if and only if)

$$c_1\vec{v}_1 + \ldots + c_n\vec{v}_n = \vec{0}$$

implies

$$c_1 = \cdots = c_n = 0.$$

The set $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of vectors is called an independent set if $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent.

注：这里linear independence是一个generalized的定义，是指在一般向量空间中的定义。如果向量空间是欧式空间，就是我们第一章的线性独立的定义。在本学期课程，线性空间暂时限制在欧式空间上定义。
Remarks:

1. Vectors $\vec{v}_1, \ldots, \vec{v}_n$ are said to be linearly dependent if they are not linearly independent, or equivalently, there exists $c_1, \ldots, c_n \in \mathbb{R}$ which are not all zero such that
$$c_1\vec{v}_1 + \ldots + c_n\vec{v}_n = \vec{0}.$$
2. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a linearly independent set then any subset of it is linearly independent.
3. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a linearly dependent set then any superset of it, that is any set contains $\{\vec{v}_1, \ldots, \vec{v}_n\}$, is linearly dependent.
4. As in the case of $\mathbb{R}^n$, $\vec{0}$ itself is linearly dependent. So if one of $\vec{v}_1, \ldots, \vec{v}_n$ is $\vec{0}$, then $\vec{v}_1, \ldots, \vec{v}_n$ are linearly dependent.

Definition 2 (basis(基))**.** Let V be a vector space and $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ an indexed set of vectors of V. Then $\mathcal{B}$ is called a basis of V iff
(1) $\mathcal{B}$ is linearly independent;
(2) $V = \operatorname{Span}\mathcal{B}$.
Also define the empty set $\emptyset$ to be the basis of the trivial space $\{\vec{0}\}$, namely a vector space contains only the zero vector.

Example: For any $n \in \mathbb{Z}^+$, $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{R}^n$, which is called the *standard basis* (标准基) of $\mathbb{R}^n$.

Example: For any $n \in \mathbb{Z}^+$, $\{1, \ldots, x^n\}$ forms a basis of $\mathbb{R}_n[x]$. It is obvious that $\mathbb{R}_n[x] = \operatorname{Span}\{1, \ldots, x^n\}$. On the other hand, if $c_0 + c_1x + \cdots + c_nx^n = 0$ then $c_0 = \cdots = c_n = 0$ by equality of polynomials.

Theorem 3. Let V be a vector space and $S = \{\vec{v}_1, \ldots, \vec{v}_n\} \subseteq V$. If $V = \operatorname{Span} S$ then some subset of S forms a basis of V.

Proof.

1. If $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent, then by assumption, S is a basis of V.

2. Otherwise, $\vec{v}_1, \ldots, \vec{v}_n$ is linearly dependent. So there exists $\vec{v}_i \in S$ such that $\vec{v}_i$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n$. Define $S_{n-1} = \{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}$. Then we have $V = \operatorname{Span} S = \operatorname{Span} S_{n-1}$. Now $S_{n-1} \subseteq S$ is either a basis of V or a linear dependent set. For the case that S_{n-1} is linearly dependent, again we can remove some $\vec{v}_j \in S_{n-1}$ from S_{n-1} to obtain $S_{n-2} \subseteq S_{n-1}$ such that $V = \operatorname{Span} S_{n-2}$. Repeat this process and finally, we must obtain a linearly independent subset S_k of S such that $V = \operatorname{Span} S_k$. Thus, S_k is a basis of V.

□

Theorem 4. *If two matrix A and B are row equivalent. Then, the columns of A has exactly the same linear dependence relationships as the columns of B.*

Proof. As A and B are row equivalent, that is A can be row reduced to a matrix B. That is the matrix equations $A\vec{x} = \vec{0}$ and $B\vec{x} = \vec{0}$ have exactly the same set of solutions. That is the columns of A have exactly the same linear dependence relationships as the columns of B. □

Theorem 5. *The pivot columns of a matrix A form a basis for Col$A$.*

Proof. We prove this theorem by Four steps: (1) We need to use the result of the last theorem. Let B be the reduced echelon form of A. The set of pivot columns of B is linearly independent, because no vector in the set is linear combination of the vectors that precede it.
(2) Since A is row equivalent to B, the pivot columns of A are linearly independent as well according to Theorem 4.
(3) The nonpivot columns of B must be the linear combination of the pivot columns of B.
(4) For this same reason, every nonpivot column of A is a linear combination of the pivot columns of A. Thus the nonpivot columns of A may be discarded from the spanning set for Col$A$, by the Spanning Set theorem. This leaves the pivot columns of A as a basis for Col$A$. □

3 Coordinate Systems

We first give the definition of coordinate and then we prove that the representation is unique.

Definition 6 (coordinate)**.** *Suppose $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis for V and $\vec{x}$ is in V. The coordinates of $\vec{x}$ relative to the basis $\mathcal{B}$ are the weights $c_1, \cdots, c_n$ such that $\vec{x} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n$. The vector $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$ is called the **coordinate vector** of $\vec{x}$ with respect to basis $\mathcal{B}$ ($\mathcal{B}$–coordinates of $\vec{x}$) and is denoted by $[\vec{x}]_{\mathcal{B}}$.*

We now prove that the coordinate with respect to basis $\mathcal{B}$ is unique.

Theorem 7. Let V be vector space and $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis of V. Then for each $\vec{x} \in V$, there exists a unique $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$ such that

$$\vec{x} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n = (\vec{b}_1 \ \cdots \ \vec{b}_n) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Proof. The existence of $(c_1 \ \cdots \ c_n)^T \in \mathbb{R}^n$ follows $V = \operatorname{Span} \mathcal{B}$. Now suppose that $(c_1' \ \cdots \ c_n')^T \in \mathbb{R}^n$ and $\vec{x} = c_1'\vec{b}_1 + \cdots + c_n'\vec{b}_n$. Then

$$\vec{0} = \vec{x} - \vec{x} = (c_1\vec{b}_1 + \cdots + c_n\vec{b}_n) - (c_1'\vec{b}_1 + \cdots + c_n'\vec{b}_n) = (c_1 - c_1')\vec{b}_1 + \cdots + (c_n - c_n')\vec{b}_n.$$

Since $\vec{b}_1, \ldots, \vec{b}_n$ are linearly independent, so $c_1 - c_1' = \cdots = c_n - c_n' = 0$. That is $c_i = c_i'$ for all i. The uniqueness follows. $\square$

4 Coordinate Mapping: A Linear Transformation View

Given a basis $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ of $\mathbb{R}^n$, for any $\vec{x} \in \mathbb{R}^n$, its coordinates are $c_1, c_2, \cdots, c_n$. We actually can rewrite the following relations:

$$\vec{x} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n = P_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}. \tag{2}$$

We call $P_{\mathcal{B}}$ the change-of-coordinates matrix from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$.

Note that $P_{\mathcal{B}}$ is invertible (because the corresponding linear transformation is injective and $P_{\mathcal{B}}$ is a square matrix). So we also have:

$$P_{\mathcal{B}}^{-1}\vec{x} = [\vec{x}]_{\mathcal{B}}. \tag{3}$$

From what we have learned from matrix theory, we have the following:

Theorem 8. *Let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis of $\mathbb{R}^n$. Then, the linear transformation $T: P_{\mathcal{B}}^{-1}\vec{x} \to [\vec{x}]_{\mathcal{B}}$ is injective (one-to-one) from $\mathbb{R}^n$ to $\mathbb{R}^n$. We call $P_{\mathcal{B}}^{-1}$ the **Coordinate Mapping**.*

Now we wish to generalize the above theorem for a more general vector space V as follows:

Theorem 9. *Let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis of V. Then, the transformation $T(\vec{x}) = [\vec{x}]_{\mathcal{B}}$ is linear and injective (one-to-one) from V to $\mathbb{R}^n$.*

Proof. There are two steps to prove this theorem. We first need to prove T is a linear transformation. Second we need to prove it is one-to-one.
Step 1: Let $\vec{u}$,$\vec{v}$ be two vectors in V, then we have:

$$\vec{v} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n, \ \vec{u} = d_1\vec{b}_1 + \cdots + d_n\vec{b}_n.$$

So for any scalar e, f, we have

$$\begin{aligned}
&T(e\vec{v}+f\vec{u}) \\
&= T((ec_1+fd_1)\vec{b}_1+\cdots+(ec_n+fd_n)\vec{b}_n) \\
&= \begin{pmatrix} ec_1+fd_1 \\ \vdots \\ ec_n+fd_n \end{pmatrix} \\
a &= e\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + f\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \\
&= e[\vec{v}]_{\mathcal{B}} + f[\vec{u}]_{\mathcal{B}} \\
&= eT(\vec{v}) + fT(\vec{u}).
\end{aligned} \tag{4}$$

Step 2: We need to prove T is one-to-one. Suppose for any $\vec{v}, \vec{u}$,

$$\vec{v} = c_1\vec{b}_1+\cdots+c_n\vec{b}_n,\ \ \vec{u} = d_1\vec{b}_1+\cdots+d_n\vec{b}_n.$$

If $T(\vec{v}) = T(\vec{u})$, then $T(\vec{v}-\vec{u}) = \vec{0}$. That is $c_1 - d_1 = 0, \cdots, c_n = d_n$, and therefore $\vec{v}-\vec{u}$ □

注：下面的概念只要求了解。

Definition 10 (isomorphism(同构))**.** *In general a one-to-one linear transformation from a vector space V onto a vector space W is called an isomorphism from V to W*

5 Change of Basis

Definition 11 (CHANGE–OF–COORDINATES MATRIX,坐标变换)**.** Let V be a vector space of dimension n with bases $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ and $\mathcal{C} = \{\vec{c}_1, \ldots, \vec{c}_n\}$. Then for $i = 1, \ldots, n$, we have

$$\vec{c}_i = (\vec{b}_1\ \cdots\ \vec{b}_n)[\vec{c}_i]_{\mathcal{B}}.$$

It is convenient to write

$$(\vec{c}_1\ \cdots\ \vec{c}_n) = (\vec{b}_1\ \cdots\ \vec{b}_n)([\vec{c}_1]_{\mathcal{B}}\ \cdots\ [\vec{c}_n]_{\mathcal{B}}) = (\vec{b}_1\ \cdots\ \vec{b}_n)[\mathcal{C}]_{\mathcal{B}}.$$

The $n \times n$ matrix $[\mathcal{C}]_{\mathcal{B}} = ([\vec{c}_1]_{\mathcal{B}}\ \cdots\ [\vec{c}_n]_{\mathcal{B}})$ is called the matrix from bases $\mathcal{C}$ to $\mathcal{B}$.

Theorem 12. Let V be a vector space of dimension n with bases $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ and $\mathcal{C} = \{\vec{c}_1, \ldots, \vec{c}_n\}$. Then for any $\vec{v} \in V$,

$$[\vec{v}]_{\mathcal{B}} = [\mathcal{C}]_{\mathcal{B}}[\vec{v}]_{\mathcal{C}}.$$

Proof. We have

$$\vec{v} = (\vec{c}_1 \ \cdots \ \vec{c}_n)[\vec{v}]_{\mathcal{C}} = (\vec{b}_1 \ \cdots \ \vec{b}_n)[\mathcal{C}]_{\mathcal{B}}[\vec{v}]_{\mathcal{C}}$$

On the other hand, $\vec{v} = (\vec{b}_1 \ \cdots \ \vec{b}_n)[\vec{v}]_{\mathcal{B}}$. By the uniqueness of coordinates, we must have $[\vec{v}]_{\mathcal{B}} = [\mathcal{C}]_{\mathcal{B}}[\vec{v}]_{\mathcal{C}}$. $\square$

Remarks: 1. Matrix $[\mathcal{C}]_{\mathcal{B}}$ is invertible, since its rank is n. It is easy to see that $[\mathcal{C}]_{\mathcal{B}}^{-1} = [\mathcal{B}]_{\mathcal{C}}$. So

$$[\vec{v}]_{\mathcal{C}} = [\mathcal{C}]_{\mathcal{B}}^{-1}[\vec{v}]_{\mathcal{B}} = [\mathcal{B}]_{\mathcal{C}}[\vec{v}]_{\mathcal{B}}.$$

2. We know that $\mathcal{E} = \{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^n$. Now let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be any basis of $\mathbb{R}^n$. Then

$$[\mathcal{B}]_{\mathcal{E}} = ([\vec{b}_1]_{\mathcal{E}} \ \cdots \ [\vec{b}_n]_{\mathcal{E}}) = (\vec{b}_1 \ \cdots \ \vec{b}_n).$$

Examples: Textbook P.274, P.275.

THE FEAST IN THE HOUSE OF LEVI, *by Veronese*