1 What Do You Learn from This Note

We will introduce linear transformations here, including its definition and different types of linear transformations.

Basic concept: transformations(变换), linear transformations(线性变换), standard matrix (标准矩阵), onto/surjection(满射), one-to-one/injection(单设, 一对一映射), image (像), pre–image (原像), domain (定义域), codomain (对映域/余定义域/取值空间), range of $T$(值域)

2 Linear Transformations

Let $A$ be an $m \times n$ matrix. Then for any $\vec{x} \in \mathbb{R}^n$, we can obtain another vector $\vec{y} = A\vec{x} \in \mathbb{R}^m$. Thus, we can create a rule using matrix $A$ which associates each vector in $\mathbb{R}^n$ to a unique vector in $\mathbb{R}^m$. For $m = n = 1$, this association turns out to be a linear function $\vec{y} = (a)\vec{x} = a\vec{x}$, which has been discussed extensively in school mathematics.

**Definition 1 (Transformations(变换)).** In general, a function $T$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is called a transformation (or map) from $\mathbb{R}^n$ to $\mathbb{R}^m$, denoted by(记为)

$$T : \mathbb{R}^n \to \mathbb{R}^m, \quad \vec{y} = T(\vec{x}) \quad (\text{or } \vec{x} \mapsto T(\vec{x})).$$

where $\vec{y}$ is called the image (像) of $\vec{x}$ (under $T$) and $\vec{x}$ is called a pre–image (原像) of $\vec{y}$, and here $\mathbb{R}^n$ is the called the domain (定义域) and $\mathbb{R}^m$ is called
the codomain (对映域/余定义域/取值空间). In addition, all images \( T(\vec{x}) \) are called the range of \( T \) (所有像组合，即所有\( T(\vec{x}) \)的组合称为值域). Note that the pre–image of \( \vec{y} \) may not be unique.

In linear algebra, we focus on studying a special type of transformations which has the property called linearity and is defined as follows.

**Definition 2** (linear transformations(线性变换)). A transformation \( T : \mathbb{R}^n \to \mathbb{R}^m \) is called linear if for any \( \vec{v}, \vec{v}_1, \vec{v}_2 \in \mathbb{R}^n \) and \( c \in \mathbb{R} \),
1. \( T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) \);
2. \( T(c\vec{v}) = cT(\vec{v}) \).

Remarks:
1. A linear transformation \( T \) always maps zero vector to zero vector since \( T(\vec{0}) = T(0 \cdot \vec{0}) = 0 \cdot T(\vec{0}) = \vec{0} \).
2. The linearity of \( T \) can be expressed in one equation, i.e. \( T \) is linear iff
   \[
   T(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1T(\vec{v}_1) + c_2T(\vec{v}_2)
   \]
   for any vectors \( \vec{v}_1, \vec{v}_2 \) and scalars \( c_1, c_2 \). The above equation can be generalised as
   \[
   T(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1T(\vec{v}_1) + \cdots + c_nT(\vec{v}_n)
   \]
   by an easy mathematical induction on \( n \).

Examples: Textbook P85(见课程板书).

### 3 The Matrix of A Linear Transformation (线性变换矩阵)

Example: Let \( A \) be any \( m \times n \) matrix. Define \( T_A : \mathbb{R}^n \to \mathbb{R}^m \), \( T_A(\vec{x}) = A\vec{x} \). Then \( T_A \) is a linear transformation. Thus from the view of linear transformation, solving matrix equation \( A\vec{x} = \vec{b} \) is exactly the same as computing the pre–images of \( \vec{b} \) under \( T_A \). The next theorem shows that any linear transformation appears in the form of matrix linear transformation.
Theorem 3. For any linear transformation \( T : \mathbb{R}^n \to \mathbb{R}^m \), there is a unique \( m \times n \) matrix \( A \) such that \( T(\vec{x}) = A\vec{x} \).

Proof. Let \( \vec{e}_i \in \mathbb{R}^n \) be such that the entries of \( \vec{e}_i \) are 0 except the \( i \)-th one which is 1. Thus, for any \( \vec{x} \in \mathbb{R} \), \( \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n \). Define

\[
A = (T(\vec{e}_1) \cdots T(\vec{e}_n)).
\]

(Existence, 存在性) We have \( T(\vec{x}) = A\vec{x} \ since

\[
T(\vec{x}) = T(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n) = x_1T(\vec{e}_1) + \cdots + x_nT(\vec{e}_n) = (T(\vec{e}_1) \cdots T(\vec{e}_n))\vec{x} = A\vec{x}.
\]

(Uniqueness, 唯一性) Suppose that \( A' = (\vec{a}'_1 \cdots \vec{a}'_n) \) and \( T(\vec{x}) = A'\vec{x} \). We shall show \( A' = A \). Since for all \( i = 1, \ldots, n \), we have

\[
T(\vec{e}_i) = A'\vec{e}_i = (\vec{a}'_1 \cdots \vec{a}'_n)\vec{e}_i = 0 \cdot \vec{a}'_1 + \cdots + 1 \cdot \vec{a}'_i + \cdots + 0 \cdot \vec{a}'_n = \vec{a}'_i
\]

So \( A' = (\vec{a}'_1 \cdots \vec{a}'_n) = (T(\vec{e}_1) \cdots T(\vec{e}_n)) = A \). \( \Box \)

Matrix \( A = (T(\vec{e}_1) \cdots T(\vec{e}_n)) \) appearing in the above theorem is called the standard matrix (标准矩阵) for \( T \).

4 Advanced Linear Transformations

Definition 4 (onto/surjection(满射)). A transformation \( T : \mathbb{R}^n \to \mathbb{R}^m \) is said to be onto (or surjective) iff

\[
\forall \vec{y} \in \mathbb{R}^m \exists \vec{x} \in \mathbb{R}^n \ T(\vec{x}) = \vec{y}
\]

in other words, for any \( \vec{y} \in \mathbb{R}^m \), \( \vec{y} \) has at least one pre-image. (即\( \mathbb{R}^m \)中的任意值都可以在定义域中找到原像；但这蕴含了多对一的映射，即定义域中不同的值经过变换后可能是同一个值)
**Definition 5** (one-to-one/injection). A transformation \( T : \mathbb{R}^n \to \mathbb{R}^m \) is said to be one-to-one (or injective) iff
\[
\forall \vec{x}_1, \vec{x}_2 \in \mathbb{R}^n \quad T(\vec{x}_1) = T(\vec{x}_2) \Rightarrow \vec{x}_1 = \vec{x}_2,
\]
in other words, for any \( \vec{y} \in \mathbb{R}^m \), \( \vec{y} \) has at most one pre–image.(即\( \mathbb{R}^m \)中的任意值至多在定义域中有一个原像；但这隐含可能不是满射，即不是\( \mathbb{R}^m \)的所有值都会在定义域中有原像)

**Definition 6** (bijection). A linear transformation \( T : \mathbb{R}^n \to \mathbb{R}^m \) is said to be bijective iff it is both injective and surjective.

The following theorem gives sufficient and necessary conditions for injective and surjective linear transformations.

**Theorem 7.** Let \( T : \mathbb{R}^n \to \mathbb{R}^m \) be a linear transformation and \( A = (\vec{a}_1 \cdots \vec{a}_n) \) be the standard matrix for \( T \). Then
1. \( T \) is injective iff (if and only if, 当且仅当) \( A\vec{x} = \vec{0} \) has only one solution, that is \( \vec{0} \);
2. \( T \) is injective iff \( \vec{a}_1, \ldots, \vec{a}_n \) are linearly independent;
3. \( T \) is subjective iff \( \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} = \mathbb{R}^m \).

**Proof.** 1. Suppose that \( T \) is injective. Then \( A\vec{x} = T(\vec{x}) = \vec{0} \) certainly has only one solution by the definition of injection.

Conversely, suppose that \( A\vec{x} = \vec{0} \) has only one solution and \( T(\vec{x}_1) = T(\vec{x}_2) \) for some \( \vec{x}_1, \vec{x}_2 \in \mathbb{R}^n \). Then
\[
A(\vec{x}_1 - \vec{x}_2) = T(\vec{x}_1 - \vec{x}_2) = T(\vec{x}_1) - T(\vec{x}_2) = \vec{0},
\]
which means \( \vec{x}_1 - \vec{x}_2 \) is a solution of \( A\vec{x} = \vec{0} \). So by uniqueness, \( \vec{x}_1 - \vec{x}_2 = 0 \), that is \( \vec{x}_1 = \vec{x}_2 \).

2. This is obvious since \( A\vec{x} = \vec{0} \) has only one solution equivalent to \( \vec{a}_1, \ldots, \vec{a}_n \) are linearly independent.(这里把\( A\vec{x} = 0 \)展开成\( x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = 0 \)来看)

3. Suppose that \( T \) is surjective. Then for any \( \vec{y} \in \mathbb{R}^m \), there exists \( \vec{x} \in \mathbb{R}^n \) such that \( \vec{y} = T(\vec{x}) = A\vec{x} = (\vec{a}_1 \cdots \vec{a}_n)\vec{x} \), that is \( \vec{y} \in \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} \). So \( \mathbb{R}^m \subseteq \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} \). Obviously, \( \mathbb{R}^m \supseteq \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} \). So \( \mathbb{R}^m = \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} \).

Conversely, suppose that \( \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} = \mathbb{R}^m \). Then for any \( \vec{y} \in \mathbb{R}^m \), there exists \( \vec{x} \in \mathbb{R}^n \) such that \( \vec{y} = (\vec{a}_1 \cdots \vec{a}_n)\vec{x} \), that is \( \vec{y} = A\vec{x} = T(\vec{x}) \). So \( T \) is surjective. \(\square\)
Reference

David C. Lay. Linear Algebra and Its Applications (3rd edition). Pages 73~90

Putti with Birds, by Boucher