Linear Transformations
Reading: Lay 1.8 and 1.9

September 13, 2013

1 Linear Transformations

Let's begin by recalling the definition of a linear transformation.

Definition 1.1. A linear transformation is a mapping T (that is, a function) from $\mathbb{R}^n$ to $\mathbb{R}^m$ (for some n and m) with the following two properties:

1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T;

2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

Recall that the function T defined by $T(\mathbf{x}) = A\mathbf{x}$ (where A is a matrix) is a linear transformation.

Example 1.2. The function T defined by

$$T\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = \begin{bmatrix} v_2 \\ v_1 \end{bmatrix}$$

is a linear transformation. Indeed,

$$\begin{aligned} T\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}\right) &= T\left(\begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} v_2 + w_2 \\ v_1 + w_1 \end{bmatrix} \\ &= \begin{bmatrix} v_2 \\ v_1 \end{bmatrix} + \begin{bmatrix} w_2 \\ w_1 \end{bmatrix} \\ &= T\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}\right). \end{aligned}$$

This proves the first property of linear transformations. The second property is proved similarly.

In fact, T is just the function defined by $T(\mathbf{v}) = A\mathbf{v}$, where

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Example 1.3. Let the vector

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

and define a mapping T by

$$T(\mathbf{v}) = \mathbf{v} + \mathbf{a}$$

Then T has domain $\mathbb{R}^3$ and codomain $\mathbb{R}^3$. T is not a linear transformation. The easiest way to see this is to note that

$$T(\mathbf{0}) = \mathbf{a} + \mathbf{0} = \mathbf{a} \neq \mathbf{0}.$$

Since a linear transformation always maps the zero vector to the zero vector, we see T cannot be a linear transformation.

2 Matrix of a linear transformation

First, I am going to introduce a standard notation.

Definition 2.1. Consider $\mathbf{R}^n$ for some n. We denote by $\mathbf{e}_i$ the vector with a 1 in the ith position and a 0 everywhere else. So for instance, in $\mathbb{R}^2$, we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 01 \end{bmatrix}.$$

Note that the actual vector denoted by, for instance, $\mathbf{e}_1$ is dependent on which $\mathbb{R}^n$ we are working in.

In this section, we will show that every linear transformation T actually has the form $T(\mathbf{v}) = A\mathbf{v}$ for some matrix A. The technique is best illustrated by an example.

Example 2.2. Let T be a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$. Suppose that

$$T(\mathbf{e}_1) = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Find a matrix A such that $T(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^2$.

Note that, if

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

then

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2.$$

So we have, by the properties of linear transformations, that

$$\begin{aligned} T(\mathbf{v}) &= v_1 T(\mathbf{e}_1) + v_2 T(\mathbf{e}_2) \\ &= v_1 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{v}. \qquad (1) \end{aligned}$$

Since $\mathbf{v}$ was an arbitrary vector, we see that T has the form of multiplication by a matrix A, where A is the matrix appearing on line (**??**) above.

The above example illustrates the general principle. It is not hard to turn the reasoning of the example into a proof of the "exists" part of the following theorem (the "unique" part is left as an exercise in Lay. We do it at the end of these notes):

Theorem 2.3. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that*

$$T(\mathbf{v}) = A\mathbf{v} \quad \textit{for all } \mathbf{v} \in \mathbb{R}^n.$$

In fact, A is the $m \times n$ matrix whose columns are given by the images of the vectors $\mathbf{e}_i$:

$$A = [T(\mathbf{e}_1)\ T(\mathbf{e}_2) \ldots T(\mathbf{e}_n)]. \qquad (2)$$

We call the matrix A appearing in (**??**) the **standard matrix for the linear transformation** T.

At this point, Lay shows a number of examples of linear transformations. You should read this, but it would be quite boring to present in class. So we will skip ahead a bit.

3 "Onto", "One-to-one"

There are two properties of functions which you have probably seen versions of in calculus which will be important in our study of linear transformations.

Definition 3.1. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is called **onto** $\mathbb{R}^m$ if every $\mathbf{b} \in \mathbb{R}^m$ is the image of at least one $\mathbf{x}$ in $\mathbb{R}^n$.

That is, a mapping T is onto if, and only if, the equation $T(\mathbf{x}) = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{R}^n$. Using the standard matrix for T allows us to state this in another way:

Theorem 3.2. *Let T be a linear transformation and A be its standard matrix. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$.*

Proof. T is onto if and only if $T(\mathbf{x}) = \mathbf{b}$ has a solution for every vector $\mathbf{b}$. Using the standard matrix, this is equivalent to saying that the equation

$$A\mathbf{x} = \mathbf{b}$$

has a solution for every $\mathbf{b}$. Using the definition of the product $A\mathbf{x}$, we see that $A\mathbf{x}$ is a linear combination of the columns of A. Thus, T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$. □

Definition 3.3. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is called **one-to-one** if each $\mathbf{b} \in \mathbb{R}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$.

That is, a mapping T is one-to-one if, and only if, for every $\mathbf{b}$ the equation $T(\mathbf{x}) = \mathbf{b}$ has either one solution or no solutions.

As with most of what we have learned in Chapter 1 of Lay, determining whether some T is one-to-one, onto, or both can be boiled down to computing the RREF for appropriate matrices.

Example 3.4. Let T be the linear transformation with standard matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Is T one-to-one? Does T map $\mathbb{R}^3$ onto $\mathbb{R}^3$?

Recall that the columns A span $\mathbb{R}^3$ if and only if there is a pivot position in each row of A. We compute the RREF of A, which turns out to be

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

There is a pivot position in each row, so the columns of A span $\mathbb{R}^3$. Therefore, T is onto.

This RREF also shows us that T is one-to-one. It is a little trickier to explain why. Note that T is one-to-one if and only if the equation $A\mathbf{x} = \mathbf{b}$ has no more than one solution for every $\mathbf{b}$. Now, when we check for solutions of $A\mathbf{x} = \mathbf{b}$, we write the augmented matrix which looks like A except which has $\mathbf{b}$ appended as an extra column. When computing the RREF of this matrix, we will find three pivots (because we found three pivots in the computation of the RREF of A). So there will be no free variables. This implies that there is at most one solution.

The method we used above to check whether T was one-to-one is perhaps a little opaque. Fortunately, a simpler way to check whether a linear transformation is one-to-one is provided by the following theorem.

Theorem 3.5. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.*

So to check whether a given T is one-to-one, we just have to check that there are no nontrivial solutions to $A\mathbf{x} = \mathbf{0}$, where A is the standard matrix for T.

Proof of Theorem **??**. Suppose T is one-to-one. Then the equation $T(\mathbf{x}) = \mathbf{0}$ has only one solution. Since the trivial solution is guaranteed to be a solution, it follows that $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.

On the other hand, suppose that $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution, and assume that T is not one-to-one. Then there exists some $\mathbf{b}$ and two vectors $\mathbf{u} \neq \mathbf{v}$ such that $T(\mathbf{u}) = T(\mathbf{v}) = \mathbf{b}$. Using the preceding equation gives us

$$T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0}.$$

But since T is linear, this implies

$$T(\mathbf{u} - \mathbf{v}) = \mathbf{0}.$$

Since $\mathbf{u} - \mathbf{v} \neq \mathbf{0}$, this implies that $T(\mathbf{x}) = \mathbf{0}$ has a nontrivial solution, a contradiction. Therefore, T must be one-to-one. □

Using the standard matrix for T, we can restate the preceding theorem.

Theorem 3.6. *A linear transformation T is one-to-one if and only if the columns of its standard matrix A are linearly independent.*

4 A "challenging" problem

In this section, we will briefly work the "hard" problem from Lay (Section 1.9, problem 33) of showing that the standard matrix for a linear transformation is unique. So suppose a linear transformation T has two standard matrices A and B. That is,

$$A\mathbf{x} = T(\mathbf{x}) = B\mathbf{x} \quad \text{for all } \mathbf{x}. \tag{3}$$

Number the columns of the matrices:

$$A = \begin{bmatrix}\mathbf{a}_1 \ \mathbf{a}_2 \ldots \mathbf{a}_n\end{bmatrix}, \quad B = \begin{bmatrix}\mathbf{b}_1 \ \mathbf{b}_2 \ldots \mathbf{b}_n\end{bmatrix}.$$

Now, by (**??**), we have that $A\mathbf{e}_1 = B\mathbf{e}_1$. Since $A\mathbf{e}_1 = \mathbf{a}_1$ and similarly for B, we have that $\mathbf{a}_1 = \mathbf{b}_1$. Repeating this for each column gives that $\mathbf{a}_i = \mathbf{b}_i$ for every number i. This implies that $A = B$. □