Linear Algebra

Riddhiman

July 2023

1 Preface

Self study Linear Algebra notes. Source is *Vector Calculus, Linear Algebra, and Differential Geometry* by John and Barbara Hubbard. These notes will start from Section 1.3. Notes for Sections 1.1 and 1.2 will be added later.

2 Matrix Multiplication as Linear Transformation

Definition 2.1 (Linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$)**:** A *linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping such that for all scalars a and for all $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathbb{R}^n$,

$$T(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = T(\vec{\mathbf{v}}) + T(\vec{\mathbf{w}}) \text{ and } T(a\vec{\mathbf{v}}) = aT(\vec{\mathbf{v}}) \tag{2.1}$$

This above was the definition of a linear transformation. The following theorem will relate matrices to linear transformation.

Theorem 2.1 (Matrices and linear transformation)

1. Any $m \times n$ matrix A defines a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is given by matrix multiplication:
$$T(\vec{\mathbf{v}}) = A\vec{\mathbf{v}}. \tag{2.2}$$

2. Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is given by multiplication by the $m \times n$ matrix $[T]$:
$$T(\vec{\mathbf{v}}) = [T]\vec{\mathbf{v}}, \tag{2.3}$$
where the i^{th} column of $[T]$ is $T(\vec{\mathbf{e}}_i)$.

Proof to part 2 of Theorem 2.1. Given that our domain is $\mathbb{R}^n$ we can write any vector $\vec{\mathbf{v}} \in \mathbb{R}^n$ as

$$\vec{\mathbf{v}} = u_1\vec{\mathbf{e}}_1 + u_2\vec{\mathbf{e}}_2 + \ldots + u_n\vec{\mathbf{e}}_n = \sum_{i=1}^{n} u_i\vec{\mathbf{e}}_i$$

Thus,

$$\begin{aligned} T(\vec{\mathbf{v}}) &= T\left(\sum_{i=1}^{n} u_i\vec{\mathbf{e}}_i\right) \\ &= \sum_{i=1}^{n} u_i T(\vec{\mathbf{e}}_i) \end{aligned}$$

Also note that since every i^{th} column of $[T]$ is $T(\vec{\mathbf{e}}_i)$. When $[T]$ is multiplied by $\vec{\mathbf{v}}$, the i^{th} of the resultant column vector can be denoted as $u_1 T(\vec{\mathbf{e}}_1) + \ldots + u_n T(\vec{\mathbf{e}}_n)$ which is just $T(\vec{\mathbf{v}})$. □

2.1 Finding the matrix of a linear transformation

By Theorem 2.1 is it clear that to find the matrix say $[T]$ associated with the linear transformation T we have to find what T does to the basis vectors $\vec{\mathbf{e}}_i$.

For example, consider a straight line passing through the origin of $\mathbb{R}^2$. Also consider two arbitrary vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$. We wish to find the resultant vectors when $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ are reflected about the line. Using the figure below it is easy to see that the act of reflecting vectors about a straight line passing through the origin is a linear transformation. Let this transformation be called T.

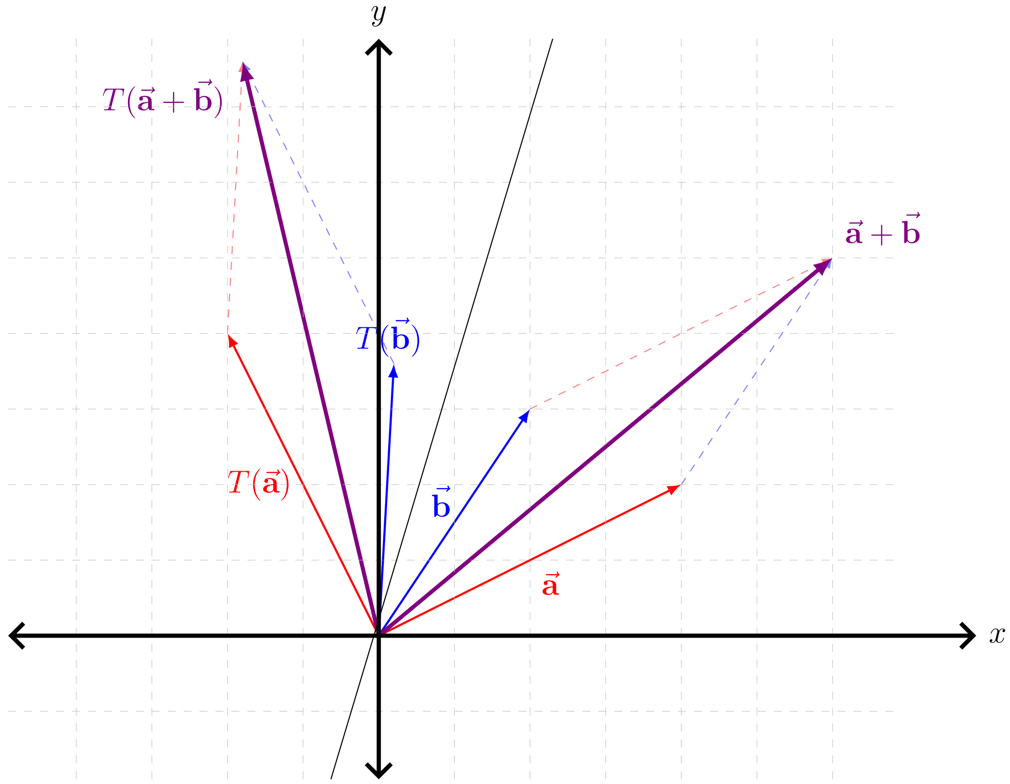


Figure 2.1: Reflection is linear

Recall that the first column of $[T]$ is $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$, and thus the second column is $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)$. Let the line make an angle θ with the x-axis. Then the matrix comes out to be

$$[T] = \begin{bmatrix}\cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta\end{bmatrix}$$

Now just multiply this matrix by any vector to get the reflected vector. This is how matrices behave – they encode linear transformations.

Another example of a linear transformation is the projection matrix. This matrix projects given vectors onto a certain line passing through the origin. If said line is making an angle θ with the x-axis then the matrix is

$$\begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$

In both the cases above, how the where the basis vectors land after applying the transformation was found by basic trigonometry.

2.2 Geometry interpretation of linear transformation

A linear transformation may be applied to entire subsets of $\mathbb{R}^n$ instead of just individual vectors. Here are some transformations, their matrix representations, and how each point in $\mathbb{R}^n$ behaves when transformed.

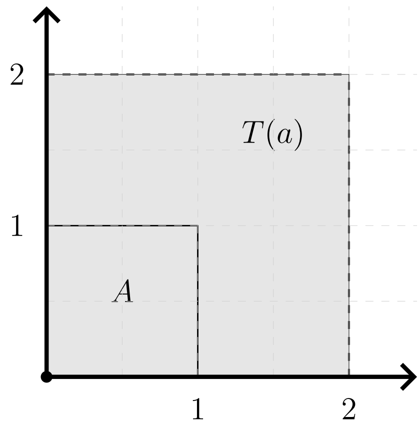
Example 2.1 (Identity transformation)**:** The identity transformation : $\mathbb{R}^n \to \mathbb{R}^n$ is represented by I_n. Applying this transformation to a subset of $\mathbb{R}^n$ leaves it unchanged.

Example 2.2 (Scaling transformation)**:** This transformation T enlarges everything by a factor of a and is given by $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$.

Example 2.3 (Stretching transformation)**:** These transformations are of the form $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. It will stretch the unit square into a rectangle.

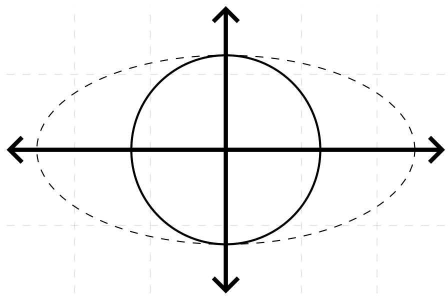
Example 2.4 (Rotation transformation)**:** This is a transformation R which rotates the subset by an angle θ in the counterclockwise around the origin. This is given by

$$[R] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



(a) The scaling transformation given by $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ turns the unit square into the square with side length 2

.



(b) The linear transformation given by $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ stretches the unit circle into an ellipse

Theorem 2.2 (Composition corresponds to matrix multiplication)

Suppose $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^l$ are linear transformations given by the matrices $[S]$ and $[T]$ respectively. Then the $[T \circ S]$ is linear and

$$[T \circ S] = [T][S] \tag{2.4}$$

Proof. First we will prove that $T \circ S$ is indeed a linear transformation. Hence, consider the following

$$\begin{aligned}(T \circ S)(a\vec{\mathbf{v}} + b\vec{\mathbf{w}}) &= T(S(a\vec{\mathbf{v}} + b\vec{\mathbf{w}})) = T(aS(\vec{\mathbf{v}}) + bS(\vec{\mathbf{w}}))\\ &= T(aS(\vec{\mathbf{v}})) + T(bS(\vec{\mathbf{w}}))\\ &= aT(S(\vec{\mathbf{v}})) + T(S(\vec{\mathbf{w}}))\\ &= a(T \circ S)(\vec{\mathbf{v}}) + b(T \circ S)(\vec{\mathbf{w}}).\end{aligned}$$

This shows that $T \circ S$ is a linear transformation has a corresponding matrix $[T \circ S]$. Now we have to prove Equation 2.4. For this we will use the following two facts:

1. $A\vec{\mathbf{e}}_i$ is the i^{th} column of A.
2. The i^{th} column of AB is $A\vec{\mathbf{b}}_i$ where $\vec{\mathbf{b}}_i$ is the i^{th} column of B.

Now,

$$[T \circ S]\vec{\mathbf{e}}_i = (T \circ S)(\vec{\mathbf{e}}_i) = T(S(\vec{\mathbf{e}}_i)) = T([S]\vec{\mathbf{e}}_i) = [T]([S]\vec{\mathbf{e}}_i). \tag{2.5}$$

Notice that by fact (1), $[T \circ S]\vec{\mathbf{e}}_i$ is the i^{th} column of $[T \circ S]$ and that $[S]\vec{\mathbf{e}}_i$ is the i^{th} column of $[S]$, and thus by fact (2), $[T]([S]\vec{\mathbf{e}}_i)$ becomes the i^{th} column of $[T][S]$. Since the columns of $[T \circ S]$ and $[T][S]$ are equivalent, they are the same matrix, which proves Equation 2.4. □

Given below is an example desmostrating composition of linear transformations.

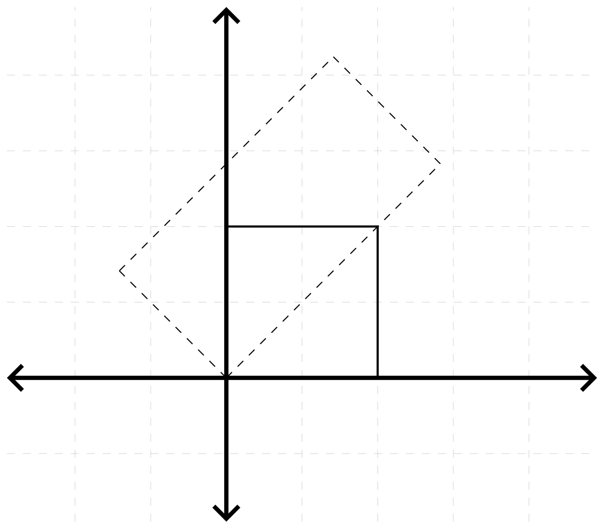


Figure 2.3: Result of applying $\begin{bmatrix}\cos\frac{\pi}{4} & -\sin\frac{\pi}{4}\\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4}\end{bmatrix}\begin{bmatrix}2 & 0\\ 0 & 1\end{bmatrix}$ to the unit square.

2.3 Invertibility of matrices and linear transformations

Proposition 2.3

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if the $m \times n$ matrix $[T]$ is invertible. If it is invertible, then

$$[T^{-1}] = [T]^{-1} \tag{2.6}$$

Before we prove this lets first break it down a bit. This proposition has essentially two statements which we have to prove,

- A linear transformation T is invertible if $[T]$ is invertible.
- $[T]$ is invertible if the linear transformation T is invertible and $[T^{-1}] = [T]^{-1}$.

Proof. We prove the statements one by one.

Assume that $[T]$ is invertible. To prove that $T : \mathbb{R}^n \to \mathbb{R}^m$ is invertible we just have to prove that T is both one-one and onto. Consider a vector $\vec{\mathbf{y}} \in \mathbb{R}^m$ and a transformation $S : \mathbb{R}^m \to \mathbb{R}^n$ such that $[S] = [T]^{-1}$. Thus we can write

$$\vec{\mathbf{y}} = ([T][S])\vec{\mathbf{y}} = T(S(\vec{\mathbf{y}})),$$

This is enough to prove that T is onto, since $\vec{\mathbf{y}}$ can be any vector in $\mathbb{R}^m$ and it always has a solution $S(\vec{\mathbf{y}})$ in $\mathbb{R}^n$. Next we prove that T is one-one. Assume two vectors $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_1 \in \mathbb{R}^n$ and that $T(\vec{\mathbf{x_1}}) = T(\vec{\mathbf{x}}_2)$. Thus,

$$\vec{\mathbf{x}}_1 = ([S][T])\vec{\mathbf{x}}_1 = S(T(\vec{\mathbf{x}}_1)) = S(T(\vec{\mathbf{x}}_2)) = ([S][T])\vec{\mathbf{x}}_2 = \vec{\mathbf{x}}_2.$$

Thus T is one-one. Hence, if $[T]$ is invertible then T is also invertible. But aren't we assuming that S is injective here?

Now we will prove the second statement. Assume that $T : \mathbb{R}^n \to \mathbb{R}^m$ is invertible. We will first prove that $T^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ is also linear (and hence has a matrix $[T^{-1}]$ associated with it). This is proved by the following computation. Let $\vec{\mathbf{y}}_1, \vec{\mathbf{y}}_2 \in \mathbb{R}^m$, then

$$\begin{aligned} T(aT^{-1}(\vec{\mathbf{y}}_1) + bT^{-1}(\vec{\mathbf{y}}_2)) &= aT(T^{-1}(\vec{\mathbf{y}}_1)) + bT(T^{-1}(\vec{\mathbf{y}}_2)) = a\vec{\mathbf{y}}_1 + b\vec{\mathbf{y}}_2 \\ &= T \circ T^{-1}(a\vec{\mathbf{y}}_1 + b\vec{\mathbf{y}}_2). \end{aligned}$$

Since, T is one-one, we can conclude that $aT^{-1}(\vec{\mathbf{y}}_1) + bT^{-1}(\vec{\mathbf{y}}_2) = T^{-1}(a\vec{\mathbf{y}}_1 + b\vec{\mathbf{y}}_2)$ proving that T^{-1} is linear. Now we are left to prove that $[T^{-1}] = [T]^{-1}$.

□

3 The geometry of $\mathbb{R}^n$

3.1 The dot product

Just everyone's favourite dot product. Also known as the *standard inner product*.

Definition 3.1 (Dot product)**:** The *dot product* $\vec{\mathbf{x}} \cdot \vec{\mathbf{y}}$ of two vectors $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{R}^n$ is

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} := x_1y_1 + x_2y_2 + \ldots + x_ny_n. \tag{3.1}$$

The dot product is commutative, distributive and can be denoted as matrix multiplication:

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = \vec{\mathbf{x}}^\top \vec{\mathbf{y}} = \vec{\mathbf{y}}^\top \vec{\mathbf{x}}.$$

. Similarly, matrix multiplication can be viewed in form of dot products.

Now we define length in terms of dot product.

Definition 3.2 (Length of a vector)**:**