

L06LinearTransformations

May 31 - June 1, 2018

1 Linear Transformations

So far we've been treating the matrix equation

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

as simply another way of writing the vector equation

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}.$$

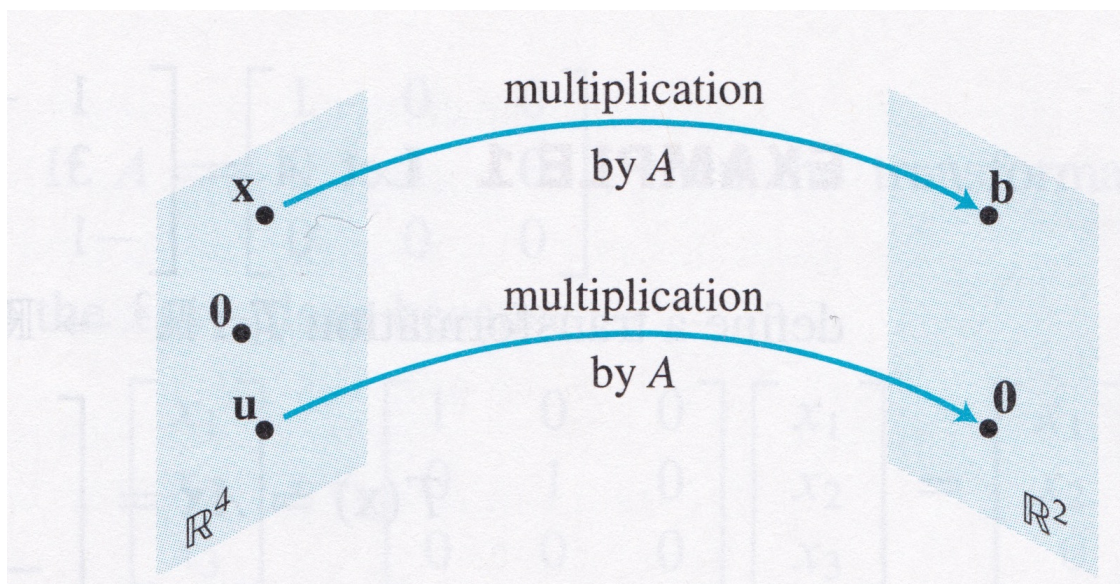
However, we'll now think of the matrix equation in a new way: we will think of A as "acting on" the vector \mathbf{x} to form a new vector \mathbf{b} .

For example, let's let $A = \begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}$. Then we find:

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In other words, if $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$, then A *transforms* \mathbf{x} into \mathbf{b} .

Likewise, if $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}$, then A transforms \mathbf{u} into the $\mathbf{0}$ vector.



This gives a **new** way of thinking about solving $A\mathbf{x} = \mathbf{b}$. We are “searching” for the vectors \mathbf{x} in \mathbb{R}^4 that are transformed into \mathbf{b} in \mathbb{R}^2 under the “action” of A .

We have moved out of the familiar world of functions of one variable: we are now thinking about functions that transform a vector into a vector.

Or, put another way, functions that transform multiple variables into multiple variables.

Some terminology:

A **transformation** (or **function** or **mapping**) T from \mathbb{R}^n to \mathbb{R}^m is a rule that assigns to each vector \mathbf{x} in \mathbb{R}^n a vector $T(\mathbf{x})$ in \mathbb{R}^m .

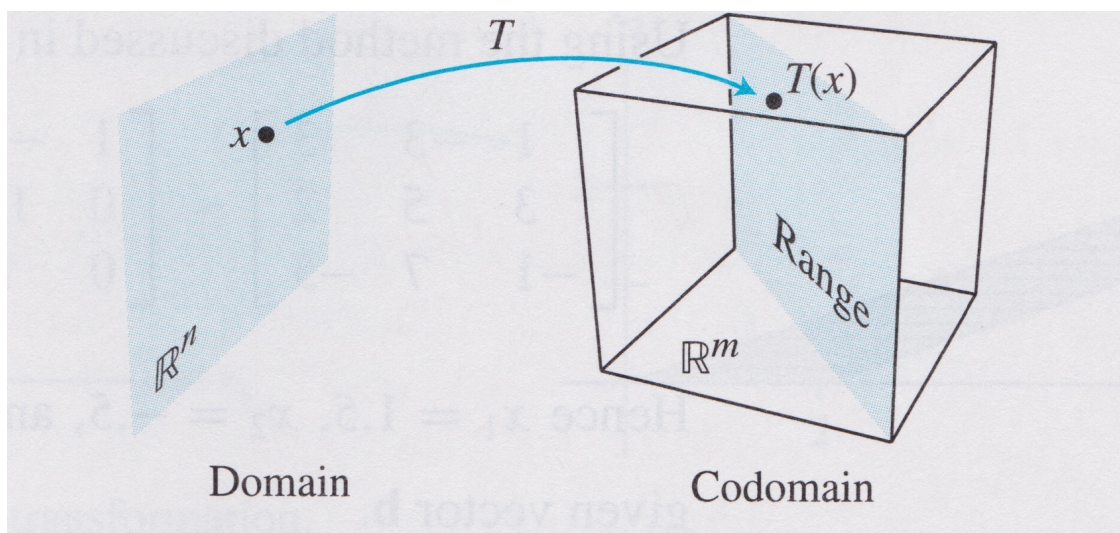
The set \mathbb{R}^n is called the **domain** of T , and \mathbb{R}^m is called the **codomain** of T .

The notation:

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

indicates that the domain of T is \mathbb{R}^n and the codomain is \mathbb{R}^m .

For \mathbf{x} in \mathbb{R}^n , the vector $T(\mathbf{x})$ is called the **image** of \mathbf{x} (under T). The set of all images $T(\mathbf{x})$ is called the **range** of T .



1.1 Question Time! Q6.1

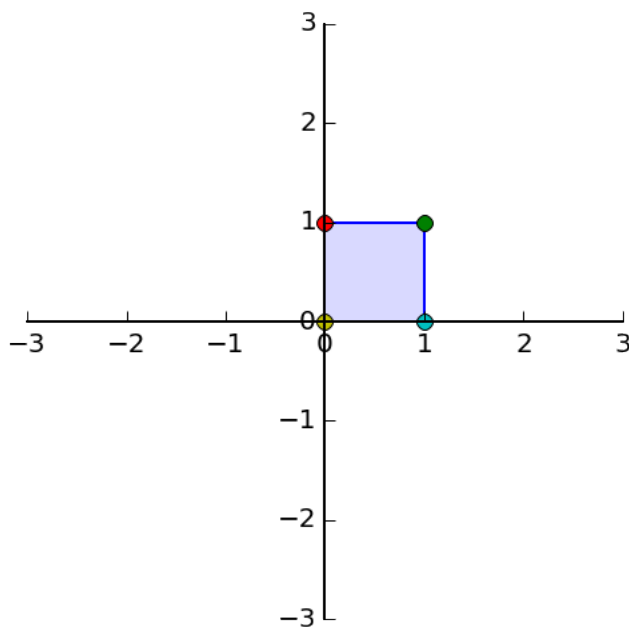
Let's do an example. Let's say I have these points in \mathbb{R}^2 :

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Where are these points located?

```
In [6]: square = np.array([[0.0,1,1,0],[1,1,0,0]])  
        ax = dm.plotSetup(-3,3,-3,3)  
        print(square)  
        dm.plotSquare(square)
```

```
[[ 0.  1.  1.  0.]  
 [ 1.  1.  0.  0.]]
```



Now let's transform each of these points according to the following rule. Let

$$A = \begin{bmatrix} 1 & 1.5 \\ 0 & 1 \end{bmatrix}.$$

We define $T(\mathbf{x}) = A\mathbf{x}$. Then we have

$$T : \mathbb{R}^2 \rightarrow \mathbb{R}^2.$$

What is the image of each of these points under T ?

$$A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$$

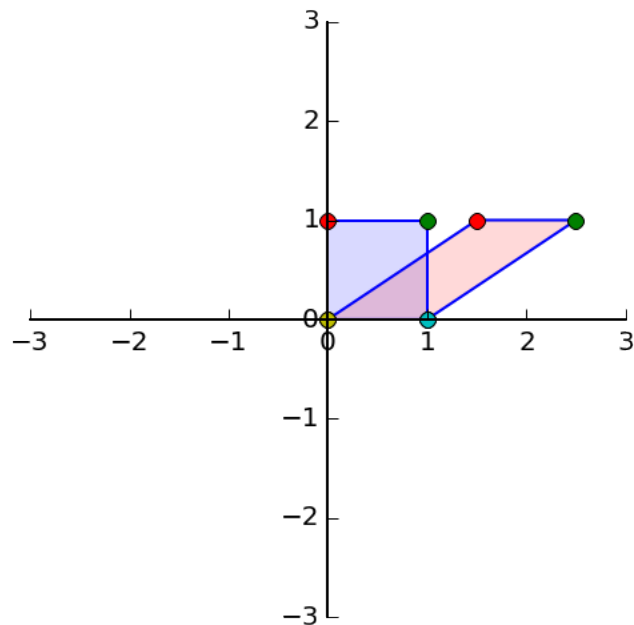
$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1 \end{bmatrix}$$

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

$$A \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

```
In [8]: ax = dm.plotSetup(-3,3,-3,3)
print("square = "); print(square)
dm.plotSquare(square)
#
# create the A matrix
A = np.array([[1.0, 1.5],[0.0,1.0]])
print("A matrix = "); print(A)
#
# apply the shear matrix to the square
ssquare = np.zeros(np.shape(square))
for i in range(4):
    ssquare[:,i] = dm.AxVS(A,square[:,i])
print("sheared square = "); print(ssquare)
dm.plotSquare(ssquare,'r')
```

```
square =
[[ 0.  1.  1.  0.]
 [ 1.  1.  0.  0.]]
A matrix =
[[ 1.  1.5]
 [ 0.  1. ]]
sheared square =
[[ 1.5  2.5  1.  0. ]
 [ 1.  1.  0.  0. ]]
```



This sort of transformation, where points are successively slid sideways, is called a **shear** transformation. We'll be looking at a number of examples in python.

Notice how we multiplied each column of 'square' by A .

In python/numpy, every matrix has a `.dot()` method. This method multiplies the matrix by every column of its argument, returning the results as a matrix.

That is:

`A.dot(B)`

where A and B are matrices, will multiply A by every column of B , and the resulting vectors will be formed into a matrix.

1.2 Linear Transformations

By the properties of matrix-vector multiplication, we know that the transformation $\mathbf{x} \mapsto A\mathbf{x}$ has the properties that

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \quad \text{and} \quad A(c\mathbf{u}) = cA\mathbf{u}$$

for all \mathbf{u}, \mathbf{v} in \mathbb{R}^n and all scalars c .

We are now ready to define one of the most fundamental concepts in the course: the concept of a *linear transformation*.

(You are now finding out why the subject is called linear algebra!)

Definition. A transformation T is **linear** if: 1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in the domain of T ; and 2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all \mathbf{u} in the domain of T .

1.3 Question Time! Q6.2

To fully grasp the significance of what a linear transformation is, don't think of just matrix-vector multiplication. Think of T as a function in more general terms.

The definition above captures a *lot* of functions that are not matrix-vector multiplication. For example, think of:

$$T(x) = \int_0^1 x(t) dt$$

Is T a linear function?

1.4 Properties of Linear Transformations

A key aspect of a linear transformation is that it **preserves the operations of vector addition and scalar multiplication**.

For example: for vectors \mathbf{u} and \mathbf{v} , one can either: 1. Transform them both according to $T()$, then add them, or: 2. Add them, and then transform the result according to $T()$.

One gets the same result either way. The transformation does not affect the addition.

This leads to two important facts.

If T is a linear transformation, then

$$T(\mathbf{0}) = \mathbf{0}$$

and

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

In fact, if a transformation satisfies the second equation for all \mathbf{u}, \mathbf{v} and c, d , then it must be a linear transformation. Both of the rules defining a linear transformation derive from this single equation.

Example.

Given a scalar r , define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(\mathbf{x}) = r\mathbf{x}$.

(T is called a **contraction** when $0 \leq r \leq 1$ and a **dilation** when $r > 1$.)

Let $r = 3$, and show that T is a linear transformation.

Solution.

Let \mathbf{u}, \mathbf{v} be in \mathbb{R}^2 and let c, d be scalars. Then

$$T(c\mathbf{u} + d\mathbf{v}) = 3(c\mathbf{u} + d\mathbf{v})$$

$$= 3c\mathbf{u} + 3d\mathbf{v}$$

$$= c(3\mathbf{u}) + d(3\mathbf{v})$$

$$= cT(\mathbf{u}) + dT(\mathbf{v})$$

Thus T is a linear transformation because it satisfies the rule $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$.