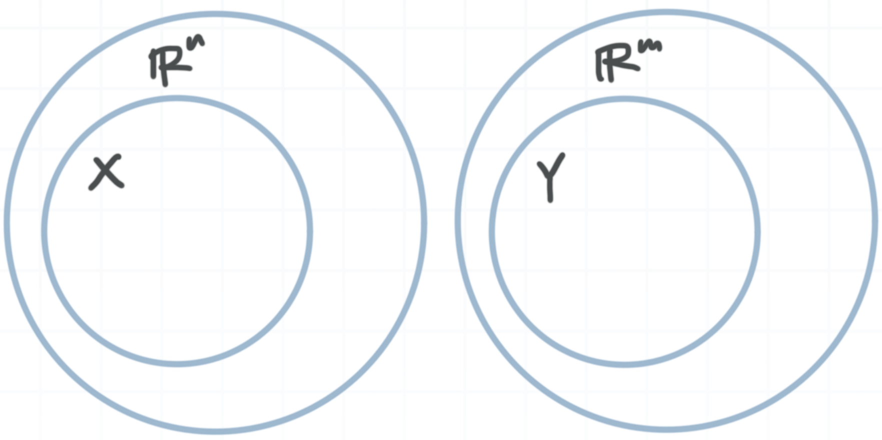
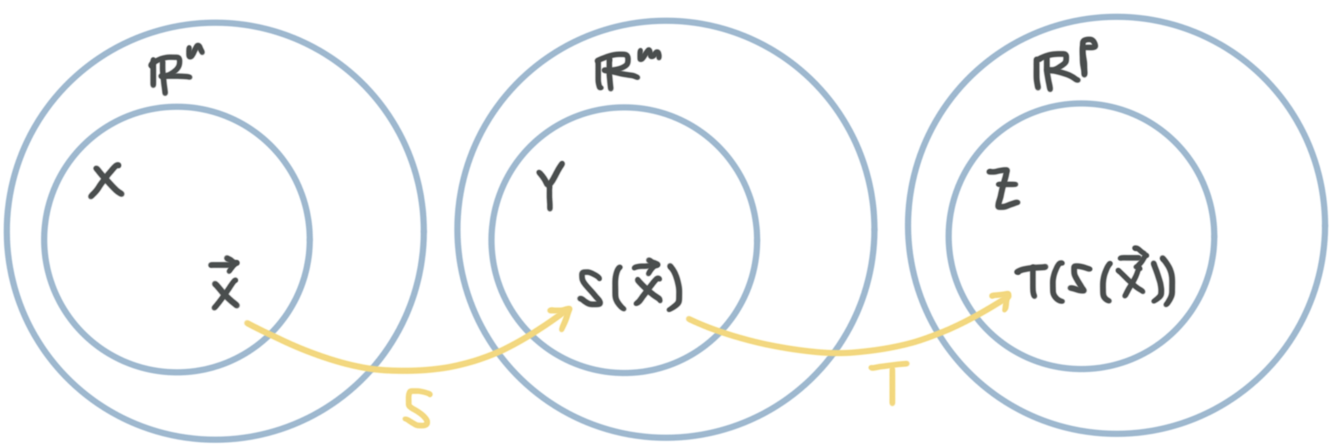
Compositions of linear transformations

We've already talked about a linear transformation as transforming vectors in a set X to vectors in a set Y. For instance, we can say that the set X is a subset of $\mathbb{R}^n$ and the set Y is a subset of $\mathbb{R}^m$.



We'll also say that S is a linear transformation that transforms vectors from X to Y, $S : X \to Y$. But let's then say that we want to take vectors from the subset Y and transform them into vectors in a subset Z, which is contained in $\mathbb{R}^p$ space. If T is a linear transformation that transforms vectors from Y to Z, $T : Y \to Z$,



then we can start to think about a composition of transformations. Because at this point, if we want to transform directly from X to Z, we can do so using a composition of the transformations S and T. S will take us from X to Y, and then T will take us from Y to Z, so a composition of T with S will take us all the way from X to Z.

For instance, let's say we have a vector $\vec{x}$ that lies in the subset X, and we want to transform it into Z. We could first transform it into Y using the transformation $S: X \to Y$, and we'd get $S(\vec{x})$. Then we could transform this transformed vector from Y into Z using the transformation $T: Y \to Z$, and we'd get $T(S(\vec{x}))$. We can also write $T(S(\vec{x}))$ as $T \circ S(\vec{x})$, so

$$T \circ S: X \to Z$$

is the composition of T with S.

Compositions as linear transformations

If we know that the transformations S and T are linear transformations, then we can say that the composition of the transformations $T(S(\vec{x}))$ is also a linear transformation. And we know this is true, because we can see that the composition is closed under addition,

$$T \circ S(\vec{x} + \vec{y}) = T(S(\vec{x} + \vec{y}))$$

$$T \circ S(\vec{x} + \vec{y}) = T(S(\vec{x}) + S(\vec{y}))$$

$$T \circ S(\vec{x} + \vec{y}) = T(S(\vec{x})) + T(S(\vec{y}))$$

$$T \circ S(\vec{x} + \vec{y}) = T \circ S(\vec{x}) + T \circ S(\vec{y})$$

and closed under scalar multiplication.

$$T(S(c\vec{x})) = T(S(c\vec{x}))$$

$$T(S(c\vec{x})) = T(cS(\vec{x}))$$

$$T(S(c\vec{x})) = cT(S(\vec{x}))$$

Additionally, because S and T are linear transformations, they can each be written as a matrix-vector product.

Let's say the transformation S is given by the matrix-vector product $S(\vec{x}) = A\vec{x}$, where A is an $m \times n$ matrix, and the transformation T is given by the matrix-vector product $T(\vec{x}) = B\vec{x}$, where B is a $p \times m$ matrix. Then because the composition is also a linear transformation, it can also be written as a matrix-vector product, and that matrix-vector product is

$$T \circ S(\vec{x}) = T(S(\vec{x})) = T(A\vec{x}) = BA\vec{x} = C\vec{x}$$

where C is a $p \times n$ matrix.

Looking at this equation, we can see that as long as we can find the matrix-matrix product BA, we'll be able to find the transformation of any vector $\vec{x}$ that's in the subset X, transformed all the way through the subset Y and into the subset Z, simply by multiplying the matrix-matrix product BA by the vector $\vec{x}$ that we want to transform.

Example

The transformation $S : X \to Y$ transforms vectors in the subset X into vectors in the subset Y. The transformation $T : Y \to Z$ transforms vectors in the subset Y into vectors in the subset Z. The subsets X, Y, and Z are all in

$\mathbb{R}^2$. Find a matrix that represents the composition of the transformations $T \circ S$, and then use it to transform $\vec{x}$ in X into its associated vector $\vec{z}$ in Z.

$$S(\vec{x}) = \begin{bmatrix} 3x_1 \\ x_1 - x_2 \end{bmatrix}$$

$$T(\vec{x}) = \begin{bmatrix} 2x_1 + 4x_2 \\ -x_2 \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

We need to start by representing each transformation as a matrix-vector product. Because in both cases we're transforming *from* $\mathbb{R}^2$, we'll use the I_2 identity matrix,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Applying the transformation S to each column of the identity matrix gives

$$S\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3(1) \\ 1-0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$S\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3(0) \\ 0-1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

So the transformation S can be written as the matrix-vector product

$$S(\overrightarrow{x}) = \begin{bmatrix} 3 & 0 \\ 1 & -1 \end{bmatrix} \overrightarrow{x}$$

Applying the transformation T to each column of the I_2 identity matrix gives

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2(1) + 4(0) \\ -0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2(0) + 4(1) \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

So the transformation T can be written as the matrix-vector product

$$T(\overrightarrow{y}) = \begin{bmatrix} 2 & 4 \\ 0 & -1 \end{bmatrix} \overrightarrow{y}$$

If we call the matrix from S

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

and we call the matrix from T

$$B = \begin{bmatrix} 2 & 4 \\ 0 & -1 \end{bmatrix}$$

then the composition of the transformations can be written as the matrix-vector product

$$T(S(\overrightarrow{x})) = BA\overrightarrow{x}$$

$$T(S(\vec{x})) = \begin{bmatrix} 2 & 4 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & -1 \end{bmatrix} \vec{x}$$

Now we can multiply the matrices. Remember that matrix multiplication (which is *not* commutative) is only defined when the number of columns in the first matrix is equivalent to the number of rows in the second matrix. In this case, the first matrix has two columns, and the second matrix has two rows, so the product of the matrices is defined.

$$C = \begin{bmatrix} 2 & 4 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & -1 \end{bmatrix}$$

$$C = \begin{bmatrix} 2(3) + 4(1) & 2(0) + 4(-1) \\ 0(3) - 1(1) & 0(0) - 1(-1) \end{bmatrix}$$

$$C = \begin{bmatrix} 6 + 4 & 0 - 4 \\ 0 - 1 & 0 + 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 10 & -4 \\ -1 & 1 \end{bmatrix}$$

This matrix will allow us to transform any vector $\vec{x}$ from the original subset X through the subset Y and into the subset Z. In other words, it lets us take vectors straight from X all the way to Z.

The composition of the transformations can be written as

$$T(S(\vec{x})) = \begin{bmatrix} 10 & -4 \\ -1 & 1 \end{bmatrix} \vec{x}$$

The problem asks us to transform $\vec{x} = (3,4)$, so we simply find the matrix-vector product.

$$T\left(S\left(\begin{bmatrix}3\\4\end{bmatrix}\right)\right)=\begin{bmatrix}10 & -4\\-1 & 1\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix}$$

$$T\left(S\left(\begin{bmatrix}3\\4\end{bmatrix}\right)\right)=\begin{bmatrix}10(3)-4(4)\\-1(3)+1(4)\end{bmatrix}$$

$$T\left(S\left(\begin{bmatrix}3\\4\end{bmatrix}\right)\right)=\begin{bmatrix}30-16\\-3+4\end{bmatrix}$$

$$T\left(S\left(\begin{bmatrix}3\\4\end{bmatrix}\right)\right)=\begin{bmatrix}14\\1\end{bmatrix}$$

Therefore, we can say that the vector $\overrightarrow{x}=(3,4)$ in the subset X is transformed into the vector $\vec{z}=(14,1)$ in the subset Z.