Linear transformations

If $\mathcal{V}$ and $\mathcal{W}$ are vector spaces, a \textit{linear transformation}, $T$, from $\mathcal{V}$ to $\mathcal{W}$ is a mapping $T: \mathcal{V} \rightarrow \mathcal{W}$ so that for all vectors $u, v$ in $\mathcal{V}$, and all scalars $c$,

1. $T(u + v) = T(u) + T(v)$
2. $T(cv) = cT(v)$.

Think of a linear transformation as a map that respects the vector space structure. Thus, $T$ not only takes vectors in $\mathcal{V}$ to vectors in $\mathcal{W}$, but also transforms addition of vectors in $\mathcal{V}$ to addition of vectors in $\mathcal{W}$, and scalar multiplication of vectors in $\mathcal{V}$ to scalar multiplication of vectors in $\mathcal{W}$.

\textbf{EXAMPLE.} The mapping $\mathbb{R}^n \rightarrow \mathbb{R}^m$ given by an $m \times n$ matrix is a linear transformation. In fact, as we will now prove, these are the only linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

\textbf{PROPOSITION.} Every linear transformation, $T$, from $\mathbb{R}^n$ to $\mathbb{R}^m$ is given by multiplication by an $m \times n$ matrix.

\textbf{PROOF.} Let $e_1, e_2, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. That is,

\[
e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{etc.} \quad \text{Then, for some } w_1, \ldots, w_n \text{ in } \mathbb{R}^m, \text{ we have}
\]

$T(e_1) = w_1, T(e_2) = w_2, \ldots, T(e_n) = w_n$. Let $M$ be the $m \times n$ matrix with columns $w_1, \ldots, w_n$.

If $v = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ is in $\mathbb{R}^n$, then $v = a_1 e_1 + \cdots + a_n e_n$. Therefore, using the linearity of $T$,

\[
T(v) = T(a_1 e_1 + a_2 e_2 + \cdots + a_n e_n) \\
= a_1 T(e_1) + a_2 T(e_2) + \cdots + a_n T(e_n) \\
= a_1 w_1 + a_2 w_2 + \cdots + a_n w_n \\
= Mv
\]

\textbf{QED}

\textbf{REMARK.} The essence of the proof is that linear transformations are very rigid in the sense that they are totally determined by their values on a basis.
Let’s consider the case $m = 1$, that is linear transformations from $\mathbb{R}^n$ to $\mathbb{R}$. These are also called linear functionals on $\mathbb{R}^n$. The proposition tells us that every linear functional can be realized as multiplication by a $1 \times n$ matrix. But multiplying a vector $\mathbf{v}$ by the matrix \[
\begin{bmatrix}
    a_1 \\
    a_2 \\
    \vdots \\
    a_n
\end{bmatrix}
\] is the same as taking the dot product of $\mathbf{v}$ with the vector \[
\begin{bmatrix}
    a_1 \\
    a_2 \\
    \vdots \\
    a_n
\end{bmatrix}
\]

COROLLARY. Every linear functional, $T$, on $\mathbb{R}^n$ has the form $T(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ for some vector $\mathbf{u}$.

The proposition tells us that, if we are studying the vector spaces $\mathbb{R}^n$, there are no linear transformations other than multiplication by a matrix. Since we already know about these maps, it may seem that, in this case at least, there is nothing to be gained by thinking in terms of linear transformations. However, this is not the case. For example, we may encounter a map defined in a way which is totally different from multiplication by a matrix, but is obviously a linear transformation. In this situation, we know there is a matrix which realizes the map, but without the concept of linear transformation, we wouldn’t have even thought to look for it.

EXAMPLE. Let $T_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the map that rotates counterclockwise through an angle $\alpha$.

It is easy to see geometrically that this is a linear transformation. For example, it does not matter whether we multiply by a scalar and then
rotate or first rotate and then multiply by a scalar. Similarly, addition is preserved. (If we think of addition geometrically, the whole picture gets rotated.) By the proposition, there must be some matrix which does the rotation, say $M_\alpha$. Moreover, to find $M_\alpha$, we just need to figure out $T_\alpha\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$ and $T_\alpha\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$. By simple trigonometry, $T_\alpha\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$ and $T_\alpha\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \cos(\alpha + \pi/2) \\ \sin(\alpha + \pi/2) \end{bmatrix} = \begin{bmatrix} -\sin \alpha \\ \cos \alpha \end{bmatrix}$. Thus, multiplication by $M_\alpha = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$ gives a map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ which rotates counterclockwise through the angle $\alpha$.

EXAMPLE. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the map which reflects through the line $x = 0$.

It is easy to see that this map respects addition and scalar multiplication. To determine the matrix, notice that $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ and $T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Therefore, the matrix $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ reflects through the line $x = 0$. 
EXAMPLE. Let \( T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) be the map which reflects through the line \( x - 2y = 0 \).

It is again easy to see that this map respects addition and scalar multiplication. To determine the matrix, notice that vectors on the line \( x - 2y = 0 \) are not moved by \( T \). Thus, \( T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \). On the other hand, vectors perpendicular to the line, such as \( \begin{bmatrix} -1 \\ 2 \end{bmatrix} \), are mapped into their negatives. That is, \( T\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \). Now,

\[
\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{5} \left( \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right).
\]

Therefore,

\[
T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \frac{1}{5} \left( T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) + 2T\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) \right) = \frac{1}{5} \left( \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right) = \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix}.
\]

Similarly, \( \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ 2 \end{bmatrix} \). Thus,

\[
T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 2 \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix} - \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}.
\]

This is now enough to determine the matrix. Multiplication by \( \begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix} \) reflects vectors in \( \mathbb{R}^2 \) through the line \( x - 2y = 0 \).
EXAMPLE. Consider the map \( T: \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) given by reflection through the plane \( x + y + z = 0 \). This is a linear transformation. It maps vectors perpendicular to the plane to their negatives. Thus,

\[
T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}.
\]

Also, \( T \) does not move vectors on the plane. Thus,

\[
T\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad T\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.
\]

We now know \( T \) on a basis for \( \mathbb{R}^3 \). This allows us to determine \( T \) applied to any vector. To find the matrix representation for \( T \), we should figure out the values of \( T \) applied to elements of the standard basis. For this, we express the standard basis in terms of our other basis. Since

\[
\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3} \left( \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right),
\]

we get

\[
T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3} \left( \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 1/3 \\ -2/3 \\ -2/3 \end{bmatrix}.
\]

Similarly,

\[
\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.
\]

Thus,

\[
T\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/3 \\ -2/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 1/3 \\ -2/3 \end{bmatrix}.
\]

Finally,

\[
\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.
\]

Thus,

\[
T\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1/3 \\ -2/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -2/3 \\ 1/3 \end{bmatrix}.
\]

It follows that multiplication by

\[
\begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{bmatrix}
\]

in \( \mathbb{R}^3 \) across the plane \( x + y + z = 0 \).

REMARK. This type of example also arises in reverse – given the matrix, describe geometrically what it does. The deepest part of this problem is finding the basis on which the transformation acts simply.