MATH 2030: MATRICES

Introduction to Linear Transformations

We have seen that we may describe matrices as symbol with simple algebraic properties like matrix multiplication, addition and scalar addition. In the particular case of matrix-vector multiplication, i.e., $A\mathbf{x} = \mathbf{b}$ where A is an $m \times n$ matrix and $\mathbf{x}, \mathbf{b}$ are $n \times 1$ matrices (column vectors) we may represent this as a transformation on the space of column vectors, that is a function $F(\mathbf{x}) = \mathbf{b}$, where $\mathbf{x}$ is the independent variable and $\mathbf{b}$ the dependent variable. In this section we will give a more rigorous description of this idea and provide examples of such matrix transformations, which will lead to the idea of a *linear transformation*.

To begin we look at a matrix-vector multiplication to give an idea of what sort of functions we are working with

$$A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

then matrix-vector multiplication yields

$$A\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

We have taken a 2×1 matrix and produced a 3×1 matrix. More generally for any $\begin{bmatrix} x \\ y \end{bmatrix}$ we may describe this transformation as a matrix equation

$$\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}.$$

From this product we have found a formula describing how A transforms an arbitrary vector in $\mathbb{R}^2$ into a new vector in $\mathbb{R}^3$. Expressing this as a transformation T_A we have

$$T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}.$$

From this example we can define some helpful terminology. A **transformation**[1] T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each each vector $\mathbf{v} \in \mathbb{R}^n$ a unique vector $T(\mathbf{v}) \in \mathbb{R}^m$. The **domain** of T is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$, and we write this as $T : \mathbb{R}^n \to \mathbb{R}^m$. For a vector $\mathbf{v}$ in the domain of T, the vector in the codomain $T(\mathbf{v})$ is called the **image** of $\mathbf{v}$ under T. The set of all possible images $T(\mathbf{v})$ for all $\mathbf{v} in \mathbb{R}^n$ is called the **range** of T. In the previous example the domain of T_A is $\mathbb{R}^2$ and the

[1] or **mapping** or **function**

codomain is $\mathbb{R}^3$, so $T_A : \mathbb{R}^2 \to \mathbb{R}^3$. The image of $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is $\mathbf{w} = T(\mathbf{v}) = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$.
The image of T_A consists of all vectors in the codmain of the form

$$T_A \begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

this describes an arbitrary linear combination of the column vectors of A. We conclude that the image consists of the column space of A. Geometrically we may see this as a plane in $\mathbb{R}^3$ through the origin with the column vectors of A as direction vectors. Notice that $T_A(\mathbf{x}) \subset \mathbb{R}^3$ where $\mathbf{x}$ is any vector in $\mathbb{R}^2$

Linear Transformations. The previous example T_A is a special case of a more general type of transformation called a *linear transformation*. We provide a less rigorous definition, that summarizes the key ideas that the transformation "respect" vector operations of addition and scalar multiplication.

Definition 0.1. A transformation $T : \mathbb{R}^m \to \mathbb{R}^m$ is called a **linear transformation** if

(1) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$.
(2) $T(c\mathbf{v}) = cT(\mathbf{v})$ for all $\mathbf{v}$ in $\mathbb{R}^n$ and all scalars c.

Example 0.2. Consider once again the transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

we will show this is indeed a linear transformation. Define $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} w \\ z \end{bmatrix}$ then compute $T(\mathbf{u} + \mathbf{v})$,

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} w \\ z \end{bmatrix}\right) = T\left(\begin{bmatrix} x + w \\ y + z \end{bmatrix}\right) = \begin{bmatrix} x + w \\ 2(x + w) - 3(y + z) \\ 3(x + w) + 4(y + z) \end{bmatrix} = \begin{bmatrix} x + w \\ (2x - 3y) + (2w - 3z) \\ (3x + 4y) + (3w + 4z) \end{bmatrix}$$

Looking at the far-right hand side we may write this as

$$\begin{bmatrix} x \\ (2x - 3y) \\ (3x + 4y) \end{bmatrix} + \begin{bmatrix} w \\ (2w - 3z) \\ (3w + 4z) \end{bmatrix} = T \begin{bmatrix} x \\ y \end{bmatrix} + T \begin{bmatrix} w \\ z \end{bmatrix} = T(\mathbf{u}) + T(\mathbf{v})$$

To show the second property, consider $T(c\mathbf{v})$ for some scalar c:

$$T\left(c \begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right) = \begin{bmatrix} cx \\ 2cx - cy \\ 3cx + 4cy \end{bmatrix} = \begin{bmatrix} cx \\ c(2x - y) \\ c(3x + 4y) \end{bmatrix} = c \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = c \begin{bmatrix} x \\ y \end{bmatrix}.$$

the second property holds, this is indeed a linear transformation.

Although the linear transformation T in the previous example arose as a matrix transformation T_A, one may go backwards and recover the matrix A from the definition of T given in the example. Notice that

$$T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x\\2x-y\\3x+4y\end{bmatrix} = x\begin{bmatrix}1\\2\\3\end{bmatrix} + y\begin{bmatrix}0\\-1\\4\end{bmatrix} = \begin{bmatrix}1 & 0\\2 & -1\\3 & 4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

where this is just the matrix-vector multiplication of A with an arbitrary vector in the domain. In general a matrix transformation is equivalent to a linear transformation, according to the next theorem

Theorem 0.3. *Let A be an $m \times n$ matrix. Then the matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by*

$$T_A(\mathbf{x}) = A\mathbf{x},\ \mathbf{x} \in \mathbb{R}^n$$

is a linear transformation.

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in the domain, and c a scalar, then $T_A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T_A(\mathbf{u}) + T_A(\mathbf{v})$ and $T_A(c\mathbf{v}) = cA\mathbf{v} = cT_A(\mathbf{v})$. Thus T_A is a linear transformation. □

Example 0.4. Q: Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that sends each point to its reflection in the x-axis. Show that F is a linear transformation.
A: This transformations send each point (x, y) to a new coordinate $(x, -y)$, and so we may write $F\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x\\-y\end{bmatrix}$ To show this is linear notice that

$$\begin{bmatrix}x\\-y\end{bmatrix} = x\begin{bmatrix}1\\0\end{bmatrix} + y\begin{bmatrix}0\\-1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

Thus $F\mathbf{x} = A\mathbf{x}$ showing that this is a matrix transformation and hence a linear transformation by the previous theorem.

Example 0.5. Q:Let $R : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that rotates each point by an angle of $\pi/4$ (90 degrees) counterclockwise about the origin. Show that F is a linear transformation.
A: Plotting this on the plane, we see that R takes any point (x, y) in the plane and sends it too $(-y, x)$, and so as a transformation

$$R\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}-y\\x\end{bmatrix} = x\begin{bmatrix}0\\1\end{bmatrix} + y\begin{bmatrix}-1\\0\end{bmatrix} = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

So R is described by a matrix transformation and therefore is a linear transformation.

Recalling that if we multiply a matrix by standard basis vectors we find the columns of the original matrix, we can use this fact to show that *every* linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ arises as a matrix transformation.

Theorem 0.6. *Let T: $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is a matrix transformation, and more specifically $T = T_A$ where A is the $m \times n$ matrix*

$$A = [T(bfe_1)|T(\mathbf{e}_2)|\cdots|T(\mathbf{e}_n)].$$

Proof. Let $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ be the standard basis vectors in $\mathbb{R}^n$ and let $\mathbf{x}$ be a vector in $\mathbb{R}^n$, so that $\mathbf{x} = x^1\mathbf{e}_1 + ... + x^n\mathbf{e}_n$. Noting that $T(\mathbf{e}_i)$ for $i = 1, ..., n$ are column vectors in $\mathbb{R}^m$, we denote $A = [T(bfe_1)|T(\mathbf{e}_2)|\cdots|T(\mathbf{e}_n)]$ be the $m \times n$ matrix with these vectors as its columns, then

$$T(\mathbf{x}) = T(x^1\mathbf{e}_1 + ... + x^n\mathbf{e}_n) = [T(bfe_1)|T(\mathbf{e}_2)|\cdots|T(\mathbf{e}_n)] \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} = A\mathbf{x}.$$

□

The matrix in the proof of the last theorem is called the **standard matrix of the linear transformation T**.

Example 0.7. Q: Show that a rotation about the origin through an angle θ defines a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ and find its standard matrix. A: Let R_θ be the rotation, we will prove this geometrically. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in the plane, then the parallelogram rule determines the new vector $\mathbf{u} + \mathbf{v}$. If we now apply R_θ the parallelogram is rotated by an angle of θ and so the diagonal of the parallelogram defined by $R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$ Hence $R_\theta(\mathbf{u}+\mathbf{v}) = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$. Similarly if we apply a rotation to $\mathbf{v}$ and $c\mathbf{v}$ by a fixed angle of θ we find $R_\theta(\mathbf{v})$ and $R_\theta(c\mathbf{v})$, however as rotations do not affect lengths we must have $R_\theta(c\mathbf{v}) = cR_\theta(\mathbf{v})$.

We conclude that R_θ is a linear transformation, and we may apply the standard basis vectors of $\mathbb{R}^2$ to this transformation to determine its standard matrix. Using trigonometry we find that

$$R \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} cos\theta \\ sin\theta \end{bmatrix}.$$

Equivalently we find that the second standard basis vector is mapped to

$$R \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -sin\theta \\ cos\theta \end{bmatrix}.$$

Thus the standard matrix for R_θ will be

$$\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

Example 0.8.
- Show that the transformation P:$\mathbb{R}^2 \to \mathbb{R}^2$ that projects a point onto the x-axis is a linear transformation and find its standard matrix.
- More generally, if ℓ is a line through the origin in $\mathbb{R}^2$, show that the transformation $P_\ell : \mathbb{R}^2 \to \mathbb{R}^2$ that projects a point onto ℓ is a linear transformation and find its standard matrix.

A:

- P sends the point (x, y) to the point $(x, 0)$ and so
$$P \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
Thus the transformation matrix for P is just $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.
- The line ℓ has direction vector $\mathbf{d}$, then for any vector $\mathbf{v}$, the transformation P_ℓ is given by $proj_{\mathbf{d}}(\mathbf{v})$ - the projection of $\mathbf{v}$ onto $\mathbf{d}$,
$$proj_{\mathbf{d}}(\mathbf{v}) = \left(\frac{\mathbf{d} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\right) \mathbf{d}.$$

To show P_ℓ is linear consider the sum

$$\begin{aligned} P_\ell(\mathbf{u}+\mathbf{v}) &= \left(\frac{\mathbf{d}\cdot(\mathbf{u}+bv)}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{d}. \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}+\mathbf{d}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{d}. \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{d}+\left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{d}. \end{aligned}$$

.

the last line is just $P_\ell(\mathbf{u})+P_\ell(\mathbf{v})$. Similarly $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$, proving that P_ℓ is indeed a linear transformation.

To determine its standard matrix, we denote $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$, the projection onto the standard basis is just

$$\begin{aligned} P_\ell(\mathbf{e}_1) &= \frac{d_1}{d_1^2+d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \\ P_\ell(\mathbf{e}_2) &= \frac{d_2}{d_1^2+d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \end{aligned}$$

implying that the standard basis is of the form

$$A = \frac{1}{d_1^2+d_2^2}\begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix}.$$

New Linear Transformations from Old. If T:$\mathbb{R}^m \to \mathbb{R}^n$ and S:$\mathbb{R}^n \to \mathbb{R}^p$ are linear transformations, then we may follow T by S to form the **composition** of the two transformations, denoted $S \circ T$. Notice that in order for $S \circ T$ to make sense, the codomain of T and the domain of S must be the same, and the resulting transformation $S \circ T$ goes from $\mathbb{R}^m$ to $\mathbb{R}^p$, that is it maps from the domain of T to the codomain of S. The formal definition of this new function is given as

$$S \circ T(\mathbf{v}) = S(T(\mathbf{v}))$$

We would like to have this new function be a linear transformation, which it is, and we may demonstrate this by showing that $S \circ T$ satisfies the definition of a linear transformation. We will do this by showing that it is a matrix transformation.

Theorem 0.9. *Let T: $\mathbb{R}^m \to \mathbb{R}^n$ and S $\mathbb{R}^n \to \mathbb{R}^p$ be linear transformations. Then $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$ is a linear transformation. Moreover, their standard matrices are related by $[S \circ T] = [S][T]$.*

Proof. Let $[S] = A$ and $[T] = B$, so that A is an $m \times n$ matrix and B a $n \times p$ matrix; if $\mathbf{v}$ is a vector in $\mathbb{R}^m$ we simply compute

$$S \circ T(\mathbf{v}) = S(T(\mathbf{v})) = S(B\mathbf{v}) = A(B\mathbf{v}) = (AB)\mathbf{v}$$

Thus the effect of $S \circ T$ is to multiply vectors by AB, from which it follows immediately that $S \circ T$ is a matrix transformation and hence a linear transformation with the transformation rule $[S \circ T] = [S][T]$. □

Example 0.10. Q:Consider the linear transformation T:$\mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x-y \\ 3x+4y \end{bmatrix}$$

and the linear transformation defined S $\mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2y_1 + y_3 \\ 3y_2 - y_3 \\ y_1 - y_2 \\ y_1 + y_2 + y_3 \end{bmatrix}$$

Find $S \circ T : \mathbb{R}^2 \to \mathbb{R}^3$.
A: Calculating the matrices of each transformation and computing their product we find

$$[S \circ T] = [S][T] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}.$$

It follows that the corresponding transformation is then

$$(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [S \circ T]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5x_1 + 4x_2 \\ 3x_1 - 7x_2 \\ -x_1 + x_2 \\ 6x_1 + 3x_2 \end{bmatrix}$$

Example 0.11. Q: Find the standard matrix of the transformation that first rotates a point 90 degrees counterclockwise about the origin and then reflects the result in the x-axis.
A: The rotation matrix $[R]$ and reflection matrix $[F]$ were given in previous examples as

$$[R] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ [F] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

composing the two we find the desired transformation

$$[R \circ R] = [F][R] = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Inverse of Linear Transformations. Consider the effect of a 90 degree counterclockwise rotation about the origin followed by a 90 degree clockwise rotation about the origin. The cumulative effect of these two transformations is the **identity transformation** I, that is, no change at all ($I(\mathbf{v}) = \mathbf{v}$). If we denote R_{90} and R_{-90} for the respective transformations this means $(R_{-90} \circ R_{90}(\mathbf{v}) = \mathbf{v}$ for any $\mathbf{v}$ in $\mathbb{R}^2$. Reversing the order geometrically gives the same result as well, i.e. $R_{90} \circ R_{-90}(\mathbf{v}) = \mathbf{v}$ as well. Thus these two linear transformations are inverses of each other and we say that any two transformations related in this manner are called **inverse transformations**.

Definition 0.12. Let S and T be linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then S and T are **inverse transformations** if $S \circ T = I_n$ and $T \circ S = I_n$.

In terms of matrices, if S and T are inverse transformations then $[S] = [T]^{-1}$ since $[S][T] = [S \circ T] = I$ where the last matrix is the identity matrix.This show that $[T]$ and $[S]$ are inverse matrices. Furthermore, if a linear transformation T is invertible, then its standard matrix $[T]$ must be invertible as well. As matrix inverses are unique, this means that the inverse of T is also unique, therefore we can use the notation T^{-1} to denote the unique inverse of each invertible linear transformation.

Theorem 0.13. *Let T:$\mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then its standard matrix $[T]$ is an invertible matrix and $[T^{-1}] = [T]^{-1}$.*

Example 0.14. Q: Find the standard matrix of a 60 degree clockwise rotation about the origin in $\mathbb{R}^2$.
A: Putting $\theta = \pi/3$ in the sines and cosines in the matrix $[R_\theta]$ and using basic trig we find that

$$[R_{60}] = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

Using the fact that a 60 degree clockwise rotation is the inverse of R_{60} , and so we may find that

$$R_{-60}] = [R_{60}]^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

by applying the last theorem.

Example 0.15. Q:Determine whether projection onto the x-axis is an invertible transformation, and if it is, find the inverse.
A: We have seen that the standard matrix for this projection transformation P is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, this is not an invertible matrix as its determinant vanishes. We conclude that P is not invertible as well.

References

[1] D. Poole, Linear Algebra: A modern introduction - 3rd Edition, Brooks/Cole (2012).