MATH 223: Linear Transformations and $2 \times 2$ matrices.
Richard Anstee

If we write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = [A^{(1)}\, A^{(2)}],$$

then

$$A\mathbf{x} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = xA^{(1)} + yA^{(2)}.$$

We can consider functions

$$f(\mathbf{x}) = A\mathbf{x}, \qquad f : \mathbf{x} \longrightarrow A\mathbf{x}.$$

We note that $A\mathbf{x}$ is a linear combination of the columns of $A$.

Later in the course the notation may be seen to be more sensible, but we write $\mathbf{R}$ to denote the Real numbers and $\mathbf{R}^2$ to denote 2-tuples of Real numbers, two coordinate vectors such as vectors in the plane.

A transformation $T : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ is *linear* if it satisfies

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

$$T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$$

This is sometimes written as a single rule that $T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v})$.

One can verify that the function $T(\mathbf{x}) = A\mathbf{x}$ is a linear transformation by verifying the linearity properties for matrix multiplication:

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}; \qquad A(\alpha\mathbf{v}) = \alpha A\mathbf{v}.$$

The first is seen as a consequence of the distributive laws, the second yields a matrix rule that $A(\alpha B) = \alpha AB$.

Assume we have a linear transformation $T$, we can determine a matrix $A$ as follows. Given a transformation $T$ that can be represented as $T(\mathbf{x}) = A\mathbf{x}$ for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we see that

$$T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} a \\ c \end{bmatrix} = A^{(1)}, \qquad T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} b \\ d \end{bmatrix} = A^{(2)}.$$

Thus the two columns of $A$ are determined as the images of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ under the transformation. We can determine $A$ completely by $T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}), \quad T(\begin{bmatrix} 0 \\ 1 \end{bmatrix})$.

Thus we have shown that linear transformations $T : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ correspond to $2 \times 2$ matrices with each linear transformation $T$ having an associated matrix $A$ to represent it; namely there is a $2 \times 2$ matrix $A$ with $T(\mathbf{x}) = A\mathbf{x}$. Also, the reverse is true; namely if $A$ is a $2 \times 2$ matrix, then we can obtain a linear transformation $T : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ by setting $T(\mathbf{x}) = A\mathbf{x}$.

Some geometric transformations can be represented by matrices (obviously they need to be linear transformations).

Dilations

These are the transformations stretching by various factors in different directions. Let

$$D(d_1, d_2) = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

then the transformation $T(\mathbf{x}) = D(d_1, d_2)\mathbf{x}$ stretches by a factor $d_1$ in the $x$ direction and a factor $d_2$ in the $y$ direction.

Rotations

These are the most beautiful $2 \times 2$ examples. Let $R(\theta)$ be the matrix corresponding to rotation by $\theta$ in the counterclockwise direction. We note that

$$R(\theta)(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \qquad R(\theta)(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

This yields the matrix which represents the transformation (assuming rotation is linear; which you can show)

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Shears

These transformations seem a little more unusual and are less commonly mentioned. Let

$$G_{12}(\gamma) = \begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix}.$$

This is seen to be the shear by a factor $\gamma$ in the $x$ direction.

The following wonderful thing happens as a consequence of our associating functions with matrices. Function composition becomes matrix multiplication.

Let $T_1(\mathbf{x}) = A_1\mathbf{x}$ and $T_2(\mathbf{x}) = A_2\mathbf{x}$ where

$$A_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad A_2 = \begin{bmatrix} e & f \\ g & h \end{bmatrix}.$$

Then we consider the composition $T_1 \circ T_2$. We have

$$T_1(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} a \\ c \end{bmatrix}, \quad T_1(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} b \\ d \end{bmatrix}, \quad T_2(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} e \\ g \end{bmatrix}, \quad T_2(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} f \\ h \end{bmatrix}.$$

Now

$$T_1 \circ T_2(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = T_1(\begin{bmatrix} e \\ g \end{bmatrix}) = eT_1(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) + gT_1(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} ae \\ ce \end{bmatrix} + \begin{bmatrix} bg \\ dg \end{bmatrix} = \begin{bmatrix} ae + bg \\ ce + dg \end{bmatrix}$$

and similarily

$$T_1 \circ T_2(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = T_1(\begin{bmatrix} f \\ h \end{bmatrix}) = fT_1(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) + hT_1(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} af \\ cf \end{bmatrix} + \begin{bmatrix} bh \\ dh \end{bmatrix} = \begin{bmatrix} af + bh \\ cf + dh \end{bmatrix}.$$

Putting this together, we obtain that the matrix for $T_1 \circ T_2$ is

$$\begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

which is the matrix product $A_1A_2$. Thus function composition corresponds to matrix multiplication. You can imagine that the rules for matrix multiplication came from a desire to have this hold.

Function composition is well known to be associative namely

$$T_1 \circ (T_2 \circ T_3) = (T_1 \circ T_2) \circ T_3.$$

This follows from computing that $(T_1 \circ (T_2 \circ T_3))(\mathbf{x}) = T_1(T_2(T_3(\mathbf{x}))) = ((T_1 \circ T_2) \circ T_3)(\mathbf{x})$ and so

$$A_1(A_2A_3) = (A_1A_2)A_3.$$

A beautiful consequence of this is the associativity of matrix multiplication follows from the associativity of function composition.

Please note the order of the operations. The transformation $T_1 \circ T_2 \circ T_3$ acts as first $T_3$, then $T_2$, then $T_1$. Since the order of matrix multiplication is important, you must check this carefully in problems.

In this vein, we see that a matrix inverse is related to the compositional inverse of linear functions. The uniqueness is easily understood in the function context as well as the fact that the inverse matrix commutes with the original matrix; namely $AA^{-1} = A^{-1}A$.