Algebra of linear transformations and matrices
Math 130 Linear Algebra
D Joyce, Fall 2015

We've looked at the operations of addition and scalar multiplication on linear transformations and used them to define addition and scalar multiplication on matrices. For a given basis β on V and another basis γ on W, we have an isomorphism $\phi_\beta^\gamma : \mathrm{Hom}(V,W) \overset{\simeq}{\to} M_{m\times n}$ of vector spaces which assigns to a linear transformation $T : V \to W$ its standard matrix $[T]_\beta^\gamma$.

We also have matrix multiplication which corresponds to composition of linear transformations. If A is the standard matrix for a transformation S, and B is the standard matrix for a transformation T, then we defined multiplication of matrices so that the product AB is be the standard matrix for $S \circ T$.

There are a few more things we should look at for matrix multiplication. It's not commutative. It is associative. It distributes with matrix addition. There are identity matrices I for multiplication. Cancellation doesn't work. You can compute powers of square matrices. And scalar matrices.

Matrix multiplication is not commutative. It shouldn't be. It corresponds to composition of linear transformations, and composition of functions is not commutative.

Example 1. Let's take a 2-dimensional geometric example. Let T be rotation 90° clockwise, and S be reflection across the x-axis. We've looked at those before. The standard matrices A for S and B for T are

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

Then the two compositions are

$$\begin{aligned} BA &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ AB &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \end{aligned}$$

The products aren't the same.

You can perform these on physical objects. Take a book. First rotate it 90° then flip it over. Start again but flip first then rotate 90°. The book ends up in different orientations.

Matrix multiplication is associative. Although it's not commutative, it is associative. That's because it corresponds to composition of functions, and that's associative. Given any three functions f, g, and h, we'll show $(f \circ g) \circ h = f \circ (g \circ h)$ by showing the two sides have the same values for all x.

$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$$

while

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))).$$

They're the same.

Since composition of functions is associative, and linear transformations are special kinds of functions, therefore composition of linear transformations is associative. Since matrix multiplication corresponds to composition of linear transformations, therefore matrix multiplication is associative.

An alternative proof would actually involve computations, probably with summation notation, something like

$$\begin{aligned} & \sum_j a_{ij} \left(\sum_k b_{jk} c_{kl} \right) \\ = & \sum_{j,k} a_{ij} b_{jk} c_{kl} \\ = & \sum_k \left(\sum_j a_{ij} b_{jk} \right) c_{kl}. \end{aligned}$$

Matrix multiplication distributes over matrix addition. When A, B, and C are the right shape matrices so the the operations can be performed, then the the following are always identities:

$$\begin{aligned} A(B+C) &= AB + AC \\ (A+B)C &= AC + BC \end{aligned}$$

Why does it work? It suffices to show that it works for linear transformations. Suppose that R, S, and T are their linear transformations. The corresponding identities are

$$\begin{aligned} R \circ (S+T) &= (R \circ S) + (R \circ T) \\ (R+S) \circ T &= (R \circ T) + (S \circ T) \end{aligned}$$

Simply evaluate them at a vector $\mathbf{v}$ and see that you get the same thing. Here's the first identity. You'll need to use linearity of R at one point.

$$\begin{aligned} (R \circ (S+T))(\mathbf{v}) &= R((S+T)(\mathbf{v})) \\ &= R(S(\mathbf{v}) + T(\mathbf{v})) \\ &= R(S(\mathbf{v})) + R(T(\mathbf{v})) \\ ((R \circ S) + (R \circ T))(\mathbf{v}) &= (R \circ S)(\mathbf{v}) + (R \circ T)(\mathbf{v}) \\ &= R(S(\mathbf{v})) + R(T(\mathbf{v})) \end{aligned}$$

The identity matrices. Just like there are matrices that work as additive identities (we denoted them all 0 as described above), there are matrices that work as multiplicative identities, and we'll denote them all I and all them *identity matrices*. An identity matrix is a square n by n matrix with 1 down the diagonal and 0 elsewhere. You could denote them I_n to emphasize their sizes, but you can always tell by the context what its size is, so we'll leave out the index n. By the way, whenever you've got a square n by n matrix, you can say the *order* of the matrix is n. Anyway, I acts like an identity matrix

$$AI = A = IA.$$

Note that if A is not a square matrix, then the orders of the two identity matrices I in the identity $AI = A = IA$ are different. For example,

$$\begin{aligned} & \begin{bmatrix} 4 & 5 & 6 \\ 3 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ = & \begin{bmatrix} 4 & 5 & 6 \\ 3 & -1 & 0 \end{bmatrix} \\ = & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 \\ 3 & -1 & 0 \end{bmatrix}. \end{aligned}$$

Cancellation doesn't work for matrix multiplication! Not only is matrix multiplication noncommutative, but the cancellation law doesn't hold for it. You're familiar with cancellation for numbers: if $xy = xz$ but $x \neq 0$, then $y = z$. But we can come up with matrices so that $AB = AC$ and $A \neq 0$, but $B \neq C$. For example $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$, and $C = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$.

Powers of matrices. Frequently, we'll multiply square matrices by themselves (you can only multiply square matrices by themselves), and we'll use the standard notation for powers. The expression A^p stands for the product of p copies of A. Since matrix multiplication is associative, this definition works, so long as p is a positive integer. But we can extend the definition to $p = 0$ by making $A^0 = I$, and the usual properties will will still hold. That is, $A^p A^q = A^{p+q}$ and $(A^p)^q = A^{pq}$. Later, we'll extend powers to the case when A is an invertible matrix and the power p is a negative integer.

Warning: because matrix multiplication is not commutative in general, it is usually the case that $(AB)^p \neq A^p B^p$.

Scalar matrices. A scalar matrix is a matrix with the scalar r down the diagonal. That's the same thing as the scalar r times the identity matrix. For instance,

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} = 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 4I.$$

Among other things, that means that we can identify a scalar matrix with the scalar.

Math 130 Home Page at
`http://math.clarku.edu/~ma130/`