Composition of linear transformations and matrix multiplication

Math 130 Linear Algebra

D Joyce, Fall 2015

Throughout this discussion, F refers to a fixed field. In application, F will usually be $\mathbf{R}$. V, W, and X will be vector spaces over F.

Consider two linear transformations $V \xrightarrow{T} W$ and $W \xrightarrow{S} X$ where the codomain of one is the same as the domain of the other. Their composition $V \xrightarrow{S\circ T} X$ is illustrated by the commutative diagram

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, "S\circ T"'] & W \arrow[d, "S"] \\ & X \end{tikzcd}$$

As each of T and S preserve linear combinations, so will the composition, so $S \circ T$ is also a linear transformation.

Coordinates again. When the vector spaces are coordinatized, that is, when we have chosen a basis β for V, γ for W, and δ for X, we have isomorphisms $\phi_\beta : V \xrightarrow{\cong} F^p$, $\phi_\gamma W \xrightarrow{\cong} F^n$, and $\phi_\delta : X \xrightarrow{\cong} F^m$.

Although we could do everything explicitly with these isomorphisms, they really get in the way of understanding. So instead, let's just assume that the vector spaces actually are F^p, F^n, and F^m, and we have two linear transformations $T : F^p \to F^n$ and $S : F^n \to F^m$.

Then $F^p \xrightarrow{T} F^n$ is represented by an $n \times p$ matrix B, $F^n \xrightarrow{S} F^m$ is represented by a $m \times n$ matrix A, and their composition $F^p \xrightarrow{S\circ T} F^m$ is represented by some $m \times p$ matrix. We'll define matrix multiplication so that the product of the two matrices AB represents the composition $S \circ T$.

$$\begin{tikzcd} F^p \arrow[r, "B"] \arrow[dr, "AB"'] & F^n \arrow[d, "A"] \\ & F^m \end{tikzcd}$$

Let's see what the entries in the matrix product AB have to be.

Let $\mathbf{v}$ be a vector in F^p, then $\mathbf{w} = T(\mathbf{v})$ is a vector in F^n, and $\mathbf{x} = S(\mathbf{w}) = (S \circ T)(\mathbf{v})$ is a vector in F^m.

The $n \times p$ matrix B represents T. Its jk^{th} entry is B_{jk}, and it was defined so that for each j,

$$w_j = \sum_k B_{jk} v_k.$$

Likewise, the $m \times n$ matrix A represents S. Its ij^{th} entry is A_{ij}, and it was defined so that for each i,

$$x_i = \sum_j A_{ij} w_j.$$

Therefore

$$x_i = \sum_j A_{ij} \sum_k B_{jk} v_k = \sum_k \left(\sum_j A_{ij} B_{jk} \right) v_k.$$

Definition 1. Given an $m \times n$ matrix A and an $n \times p$ matrix B, we define AB to be an $m \times p$ matrix whose ik^{th} entry is

$$(AB)_{ik} = \sum_j A_{ij} B_{jk}.$$

With this definition, matrix multiplication corresponds to composition of linear transformations.

A mnemonic for multiplying matrices. Although the equation $(AB)_{ik} = \sum_j A_{ij} B_{jk}$ is fine for theoretical work, in practice you need a better way to remember how to multiply matrices.

The entry A_{ij} in a row of the first matrix needs to be multiplied by the corresponding B_{jk} in a column of the second matrix. If you place the matrix A to the left of the product and place the matrix B above the product, it's easier to see what to multiply by what.

Take, for instance, the following two 3 by 3 matrices.

$$A = \begin{bmatrix} 4 & 5 & 6 \\ 3 & -1 & 0 \\ 2 & 0 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 4 & 5 \\ -2 & -3 & 0 \end{bmatrix}$$

Think of A as being made of three row vectors and B as being made of three column vectors.

$$A = \left[\begin{array}{ccc} 4 & 5 & 6 \\ \hline 3 & -1 & 0 \\ \hline 2 & 0 & -2 \end{array}\right], \quad B = \left[\begin{array}{c|c|c} 2 & 1 & 1 \\ 0 & 4 & 5 \\ -2 & -3 & 0 \end{array}\right]$$

$$\begin{array}{cc} & \left[\begin{array}{r|r|r} 2 & 1 & 1 \\ 0 & 4 & 5 \\ -2 & -3 & 0 \end{array}\right] \\ \left[\begin{array}{rrr} 4 & 5 & 6 \\ \hline 3 & -1 & 0 \\ \hline 2 & 0 & -2 \end{array}\right] & \left[\begin{array}{r|r|r} -4 & 6 & 29 \\ \hline 6 & -1 & -2 \\ \hline 8 & 8 & 2 \end{array}\right] \end{array}$$

To get an entry for the product, work with the row in A to the left of it and the column of B above it. For example, the upper left entry of the product, work with the first row of A and the first column of B; you'll get $4 \cdot 2 + 5 \cdot 0 + 6 \cdot (-2) = -4$.

Systems of linear equations are linear matrix equations. We'll have a lot of uses for matrix multiplication as the course progresses, and one of the most important is the interpretation of a system of linear equations as a single matrix equation.

Take, for example, the system of equations

$$\begin{aligned} 5x + 2y &= 12 \\ 3x - y &= 5 \\ x + 3y &= 5 \end{aligned}$$

Let A be the coefficient matrix for this system, so that

$$A = \begin{bmatrix} 5 & 2 \\ 3 & -1 \\ 1 & 3 \end{bmatrix},$$

and let $\mathbf{b}$ be the constant matrix (a column vector) for this system, so that

$$\mathbf{b} = \begin{bmatrix} 12 \\ 5 \\ 5 \end{bmatrix}.$$

Finally, let $\mathbf{x}$ be the variable matrix for this system, that is, a matrix (another column vector) with the variables as its entries, so that

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Then the original system of equations is described by the matrix multiplication $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 5 & 2 \\ 3 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 12 \\ 5 \\ 5 \end{bmatrix}$$

In general, each system of linear equations corresponds to a single matrix equation

$$A\mathbf{x} = \mathbf{b}$$

where A is the matrix of coefficients in the system of equations, $\mathbf{x}$ is a vector of the variables in the equations, and $\mathbf{b}$ is a vector of the constants in the equations. This interpretation allows us to interpret something rather complicated, namely a whole system of equations, as a single equation.

Matrix products in MATLAB**.** If `A` and `B` are two matrices of the right size, that is, `A` has the same number of columns that `B` has rows, then the expression `A*B` gives their product. You can compute powers of square matrices as well. If `A` is a square matrix, then `A^3` computes the same thing as `A*A*A`.

Categories. Categories are higher order algebraic structures. We'll look at a couple of categories. One will be the category of vector spaces and linear transformations over a field, the other the category of matrices over a field F. We'll also consider the category of sets, but primarily just as another example of categories.

Mathematics abounds with categories. There are categories of topological spaces, of differentiable spaces, of groups, of rings, etc.

The purpose of a category is to study the interrelations of its objects, and to do that the category includes 'morphisms' (also called maps or arrows) between the objects. In the case of the category of vector spaces, the morphisms are the linear transformations.

We'll start with the formal definition of categories. Category theory was developed by Eilenberg and Mac Lane in the 1940s.

Definition 2. A *category* $\mathcal{C}$ consists of

1. *objects* often denoted with uppercase letters, and

2. *morphisms* (also called *maps* or *arrows*) often denoted with lowercase letters.

3. Each morphism f has a *domain* which is an object and a codomain which is also an object. If the domain of f is A and the codomain is B, then we write $f : A \to B$ or $A \xrightarrow{f} B$. The set of all morphisms from A to B is denoted $\mathrm{Hom}(A, B)$.

4. For each object A there is a morphism $1_A : A \to A$ called the *identity morphism* on A.

5. Given two morphisms $A \xrightarrow{f} B$ and $B \xrightarrow{g} C$ where the codomain of one is the same as the domain of the other there is another morphism $A \xrightarrow{g\circ f} C$ called the *composition* of the two morphisms. This composition is illustrated by the commutative diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "g\circ f"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

6. for all $A \xrightarrow{f} B$, $f \circ 1_A = f$ and $1_B \circ f = f$. These compositions are illustrated by the two commutative diagrams

$$\begin{tikzcd} A \arrow[d, "1_A"'] \arrow[dr, "f\circ 1_A"] & \\ A \arrow[r, "f"'] & B \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "1_B\circ f"'] & B \arrow[d, "1_B"] \\ & B \end{tikzcd}$$

7. for all $A \xrightarrow{f} B$, $B \xrightarrow{g} C$, and $C \xrightarrow{h} D$, $(h \circ g) \circ f = h \circ (g \circ f)$. In the diagram below, if the two triangles in the diagram each commute, then the parallelogram commutes.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "g\circ f"'] & B \arrow[d, "g"] \arrow[dr, "h\circ g"] & \\ & C \arrow[r, "h"'] & D \end{tikzcd}$$

A diagram of objects and morphisms in a category is said to *commute*, or be a *commutative diagram* if any two paths of morphisms (in the direction of the arrows) between any two objects yield equal compositions.

Isomorphisms in a category $\mathcal{C}$. Although only morphisms are defined in a category, it's easy to determine which ones are isomorphisms. A morphism $f : A \to B$ is an *isomorphism* if there exists another morphism $g : B \to A$, called its *inverse*, such that $f \circ g = 1_B$ and $g \circ f = 1_A$.

Example 3 (The categories of sets $\mathcal{S}$)**.** Although we're more interested in the category of vector spaces right now, the category $\mathcal{S}$ of sets is also relevant. An object in $\mathcal{S}$ is a set, and a morphism in $\mathcal{S}$ is a function. The domain and codomain of a morphism are just the domain and codomain of the function, and composition is composition. If S and T are two sets, then $\mathrm{Hom}(S, T)$ is the set of all functions $S \to T$.

Isomorphisms in the category of sets are bijections.

Example 4 (The category of vector spaces $\mathcal{V}_F$)**.** Fix a field F. The objects in the category $\mathcal{V}_F$ are vector spaces over a F and the morphisms are linear transformations. Different fields have different categories of vector spaces. $\mathrm{Hom}(V, W)$ is the vector space of linear transformations $V \to W$. Since it's a vector space over F itself, it's actually an object in the category.

Isomorphisms in the category of vector spaces are what we've been calling isomorphisms.

Example 5 (The category of matrices $\mathcal{M}_F$)**.** We'd like the matrices over a fixed field F to be the morphisms in this category. Composition will then be multiplication of matrices. But then, what are the objects?

The objects in $\mathcal{M}_F$ are the vector spaces F^n for $n = 0, 1, 2, \ldots$. A morphism $F^n \to F^m$ is an $m \times n$ matrix A. The composition of two matrices $F^p \xrightarrow{B} F^n$ and $F^n \xrightarrow{A} F^m$ is the matrix product $F^p \xrightarrow{AB} F^m$ as we defined it above.

The identity morphism $F^n \to F^n$ is the $n \times n$ identity matrix I with 1's down the diagonal and 0's elsewhere.

$\mathrm{Hom}(F^n, F^m)$ is the set of matrices we've denoted by M_{mn}.

The category $\mathcal{M}_F$ of matrices is can be interpreted as a subcategory of the category of vector spaces $\mathcal{V}_F$. It doesn't include all the vector spaces, as infinite dimensional vector spaces aren't objects of $\mathcal{M}_F$. Furthermore, $\mathcal{M}_F$ doesn't have any finite dimensional vector spaces except those of the form F^n. We know, however, that every vector space V of finite dimension n is isomorphic F^n.

Note that the only isomorphisms $F^n \to F^m$ in $\mathcal{M}_F$ occur when $n = m$.

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