Math 217: §2.3 Composition of Linear Transformations

Professor Karen Smith[1]

Inquiry: IS THE COMPOSITION OF LINEAR TRANSFORMATIONS A LINEAR TRANSFORMATION? IF SO, WHAT IS ITS MATRIX?

A. Let $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^3$ and $\mathbb{R}^3 \xrightarrow{S} \mathbb{R}^2$ be two linear transformations.

1. Prove that the composition $S \circ T$ is a linear transformation (*using the definition!*). What is its source vector space? What is its target vector space?

> *Solution note:* The source of $S \circ T$ is $\mathbb{R}^2$ and the target is also $\mathbb{R}^2$. The proof that $S \circ T$ is linear: We need to check that $S \circ T$ respect addition and also scalar multiplication. First, note for any $\vec{x}, \vec{y} \in \mathbb{R}^2$, we have
>
> $$S \circ T(\vec{x} + \vec{y}) = S(T(\vec{x} + \vec{y})) = S(T(\vec{x})) + S(T(\vec{y})) = S \circ T(\vec{x}) + S \circ T(\vec{y}).$$
>
> Here, the second and third equal signs come from the linearity of T and S, respectively. Next, note that for any $\vec{x} \in \mathbb{R}^2$ and any scalar k, we have
>
> $$S \circ T(k\vec{x}) = S(T(k\vec{x})) = S(kT(\vec{x})) = kS(T(\vec{x})) = kS \circ T(\vec{x}),$$
>
> so $S \circ T$ also respects scalar multiplication. The second and third equal signs again are justified by the linearity of T and S, respectively. So $S \circ T$ respects both addition and scalar multiplication, so it is linear.

2. Suppose that the matrix of T is $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ -1 & 3 \end{bmatrix}$ and the matrix of S is $B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$.

   Compute explicitly a formula for $S(T(\begin{bmatrix} x \\ y \end{bmatrix}))$.

> *Solution note:* $S \circ T(\begin{bmatrix} x \\ y \end{bmatrix}) = S(\begin{bmatrix} x + 2y \\ -y \\ -x + 3y \end{bmatrix}) = \begin{bmatrix} x + 2y + (-x + 3y) \\ y \end{bmatrix} = \begin{bmatrix} 5y \\ y \end{bmatrix}$

3. What is $S \circ T(\vec{e}_1)$? $S \circ T(\vec{e}_2)$?

> *Solution note:* $S \circ T(\vec{e}_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $S \circ T(\vec{e}_2) = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

4. Find the matrix of the composition $S \circ T$. Compare to (2).

> *Solution note:* $\begin{bmatrix} 0 & 5 \\ 0 & 1 \end{bmatrix}$. Note that the columns are the vectors we computed in (2).

[1]Thanks to Anthony Zheng, Section 5 Winter 2016, for finding numerous errors in the solutions.

5. Compute the matrix product BA. Compare to (4). What do you notice?

Solution note: $BA = \begin{bmatrix} 0 & 5 \\ 0 & 1 \end{bmatrix}$. The same!

6. What about $T \circ S$? What is its matrix in terms of A and B.

Solution note: This is AB.

7. What is the general principle here? Say we have a composition of linear transformations

$$\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^m \xrightarrow{T_B} \mathbb{R}^p$$

given by matrix multiplication by matrices A and B respectively. State and prove a precise theorem about the matrix of the composition. **Be very careful about the order of multiplication!**

Solution note: **Theorem:** If $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^m \xrightarrow{T_B} \mathbb{R}^p$ are linear transformations given by matrix multiplication by matrices A and B (on the left) respectively, then the composition $T_B \circ T_A$ has matrix BA.

Proof: For any $\vec{x} \in \mathbb{R}^n$, we have

$$T_B \circ T_A(\vec{x}) = T_B(T_A(\vec{x})) = T_B(A\vec{x}) = BA\vec{x} = (BA)\vec{x}.$$

Here, every equality uses a definition or basic property of matrix multiplication (the first is definition of composition, the second is definition of T_A, the third is definition of T_B, the fourth is the association property of matrix multiplication).

B. Let $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 1 & c \\ d & 1 \end{bmatrix}$.

1. Compute AC. Compute CA. What do you notice?
2. Does matrix multiplication satisfy the commutative law?
3. TRUE or FALSE: If we have two linear transformations, S and T, both from $\mathbb{R}^n \to \mathbb{R}^n$, then $S \circ T = T \circ S$.

Solution note: $AC = \begin{bmatrix} ad+1 & a+c \\ d & 1 \end{bmatrix}, CA = \begin{bmatrix} 1 & a+c \\ d & ad+1 \end{bmatrix}$. These are not equal in general, so matrix multiplication does not satisfy the commutative law! In particular, linear transformations do not satisfy the commutative law either, so (3) is FALSE. An explicit countexample is to let S be left multiplication by $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, and T be multiplication by $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Then TS is multiplication by $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, but ST is multiplication by $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. These are not the same maps, since for example, they take different values on $\vec{e}_1$.

C. The **identity transformation** is the map $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ doing *nothing:* it sends every vector $\vec{x}$ to $\vec{x}$. A **linear transformation** T is invertible if there exists a linear transformation S such that $T \circ S$ is the identity map (on the source of S) and $S \circ T$ is the identity map (on the source of T).

1. What is the matrix of the identity transformation? Prove it!
2. If $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ is invertible, what can we say about its matrix?

Solution note: The matrix of the identity transformation is I_n. To prove it, note that the identity transformation takes $\vec{e}_i$ to $\vec{e}_i$, and that these are the columns of the identity matrix. So the identity matrix is the unique matrix of the identity map. If T is invertible, then the matrix of T is invertible.

Math 217: §2.3 Block Multiplication

Professor Karen Smith

D. In the book's Theorem 2.3.9, we see that we can think about matrices in "blocks" (for example, a 4×4 matrix may be thought of as being composed of four 2×2 blocks), and then we can multiply as though the blocks were scalars using Theorem 2.3.4. This is a surprisingly useful result!

1. Consider the matrix $C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. If we write this as a block matrix, $C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$, where all the blocks are the same size, what are the blocks C_{ij}?

   Solution note: One way is: $C_{11} = C_{12} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $C_{21} = 0_{2\times 2}$ (the 2×2 zero matrix), and $C_{22} = I_2$ (the 2×2 identity matrix).

2. Suppose we want to calculate the product CD, where D is the block matrix $D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$, with D_1 and D_2 each being a 2×2 block. Write the product in terms of C_{ij} and D_k by multiplying the blocks as if they were scalars, as suggested by Theorem 2.3.9.

   Solution note: We have
   $$CD = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} C_{11}D_1 + C_{12}D_2 \\ C_{21}D_1 + C_{21}D_2 \end{bmatrix}.$$
   Calculate

3. Compute the product AB where
$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} a & p & x & 0 & 0 & 0 & 1 & 0 & 0 \\ b & q & y & 0 & 0 & 0 & 0 & 1 & 0 \\ c & r & z & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 3 & 0 & 0 & 1 & 0 & 0 & 1 \\ -a & -p & -x & 0 & 0 & 0 & 1 & 0 & 0 \\ -b & -q & -y & 0 & 0 & 0 & 0 & 1 & 0 \\ -c & -r & -z & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$
   [Hint: Be clever, take advantage of lurking identity matrices and block multiplication.]

> *Solution note:* Break this up, sudoku-like, into nine 3×3 blocks. Note, in A, almost all these are identity matrices or zero matrices, so the multiplication is especially easy.

4. With C as in (1), find another way to break C up as $= \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$, but where the blocks are *not the same size.*

> *Solution note:* You could take C_{11} to be 1×1 and C_{22} to be 3×3. So $C_{11} = 1$, $C_{12} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, C_{21} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, and $C_{22} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. It's easiest to see what I mean by drawing in two lines cutting the matrix up into blocks.

F. Consider a matrix $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ where A is 3×5 and D is 7×1 (so M is in block form).

1. What are the sizes of B and C? What is the size of M?

> *Solution note:* M is 10×6, B is 3×1 and C is 7×5.

2. Suppose $N = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}$? What are the possible dimensions of N so that the product MN is defined? In this case, what are the dimensions of the smaller blocks A', B', C' and D' so that the product can be computed using a block-product?

> *Solution note:* The only restrictions are as follows: A' and B' must have 5 rows, and C' and D' must have one row. Also, A' and C' must have the same number of columns, as must B' and D'. So A' is $5 \times n$, B' is $5 \times m$, C' is $1 \times n$ and D' is $1 \times m$.

3. What are the possible dimensions of N so that the product NM is defined? In this case, what are the dimensions of the smaller blocks A', B', C' and D' that would allow us to compute this as a block product?

> *Solution note:* For this, the number of columns of N must equal the number of rows of M, so N must by $p \times 10$. For the block multiplication to work, we must have A' is $a \times 3$, B' is $a \times 7$, C' is $b \times 3$ and D' is $b \times 7$. Here $a + b = p$.

4. If A is $m \times d$ and B is $d \times n$, how can we think of the product AB as a column (of row vectors) times a row (of column vectors)?

Solution note: Write A as a column of row vectors: $A = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix}$ where each R_i is a $1 \times m$ matrix (row vector). Write B as a row of columns vectors: $B = \begin{bmatrix} C_1 & C_2 \\ \dots \\ C_m \end{bmatrix}$ where each C_j is a $n \times 1$ matrix (columns vector). Then

$$AB = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix} \begin{bmatrix} C_1 & C_2 & \dots & C_m \end{bmatrix},$$

which is the $n \times m$ matrix with $R_i \cdot C_j$ in the ij-spot.