Lecture 3d

Matrix Multiplication

(pages 121-124)

The transpose is an interesting concept, but what really separates matrices from $\mathbb{R}^n$ is the fact that we define multiplication of matrices. But matrix multiplication is NOT defined entrywise. Instead, the inspiration for matrix multiplication lies back with a system of linear equations. In fact, consider the following system of linear equations:

$$\begin{array}{ccccccc} x_1 & - & 3x_2 & + & 5x_3 & = & 2 \\ 2x_1 & & & + & 2x_3 & = & 0 \\ -x_1 & & & + & x_3 & = & -2 \\ -2x_1 & + & 4x_2 & - & 3x_3 & = & -7 \end{array}$$

The coefficient matrix for this system is

$$\begin{bmatrix} 1 & -3 & 5 \\ 2 & 0 & 2 \\ -1 & 0 & 1 \\ -2 & 4 & -3 \end{bmatrix}$$

and matrix multiplication is defined in a way that lets us rewrite our system of linear equations using a product of matrices:

$$\begin{bmatrix} 1 & -3 & 5 \\ 2 & 0 & 2 \\ -1 & 0 & 1 \\ -2 & 4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \\ -7 \end{bmatrix}$$

This, of course, should not make sense to you yet, but sometimes it is helpful to keep the inspiration in mind when going through the definition. Amongst other things, this gives you a glimpse at the complexity of matrix multiplication. For one thing, you don't multiply matrices of the same size! (Sure, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 0 \\ -2 \\ -7 \end{bmatrix}$ look like vectors, not matrices, but we are treating them like 3×1 and 4×1 matrices at this point. The definition of matrix multiplication will allow for more situations than simply matrix times vector.) So instead of focusing on the size of the various matrices, let's instead see how we get from the matrix product to our system of equations. First, the coefficients in the each ROW of the coefficient matrix need to get matched up with the variables in the COLUMN of the second matrix. If we think of the rows of the coefficient matrix as vectors (so we take their transpose–see, I told you it would come in handy), then we see that

$$\begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 - 3x_2 + 5x_3 \qquad \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x_1 + 2x_3$$

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -x_1 + x_3 \qquad \begin{bmatrix} -2 \\ 4 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -2x_1 + 4x_2 - 3x_3$$

And so we want the matrix product $\begin{bmatrix} 1 & -3 & 5 \\ 2 & 0 & 2 \\ -1 & 0 & 1 \\ -2 & 4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ to be

$$\begin{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ \begin{bmatrix} -2 \\ 4 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 + 5x_3 \\ 2x_1 + 2x_3 \\ -x_1 + x_3 \\ -2x_1 + 4x_2 - 3x_3 \end{bmatrix}$$

Don't get too wrapped up in the actual numbers here–the idea is that we end up looking at the rows of our coefficient matrix as vectors, and then take the dot product with our variables column, and viola, we get our equations. The crucial idea here is that we take the dot product of a row with a column. Not multiply entry by entry, not dot product of row with row. Row-dot-Column. Okay, with that idea firmly planted in your mind, here's how we define matrix multiplication:

<u>Definition:</u> Let B be an $m \times n$ matrix with rows $\vec{b}_1^T, \ldots, \vec{b}_m^T$, and let A be an

$n \times p$ matrix with columns $\vec{a}_1, \ldots, \vec{a}_p$. Then we define BA to be the matrix whose ij-th entry is

$$(BA)_{ij} = \vec{b}_i \cdot \vec{a}_j$$

If the notation is a bit too much for you just now, keep the following idea in mind: when we look at the product BA, we will end up taking the dot product of every row of B with every column of A. Every possible row-column combination will occur, and when we take the dot product of the i-th row of B with the j-th column of A, we end up with the ij-th entry of BA. So if it was the 3rd row of B with the 5th column of A, we would end up with $(BA)_{35}$. And so on. At this point, I think an example is in order:

Example:

$$\begin{bmatrix} 1 & 4 \\ -5 & 7 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 & 7 \\ -5 & 1 & 4 & -1 \end{bmatrix} =$$

$$\begin{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} & \begin{bmatrix} 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} & \begin{bmatrix} 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ -1 \end{bmatrix} \\ \begin{bmatrix} -5 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} & \begin{bmatrix} -5 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} & \begin{bmatrix} -5 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} & \begin{bmatrix} -5 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 0 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} & \begin{bmatrix} 0 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} & \begin{bmatrix} 0 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ -1 \end{bmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} (1)(2)+(4)(-5) & (1)(-1)+(4)(1) & (1)(3)+(4)(4) & (1)(7)+(4)(-1) \\ (-5)(2)+(7)(-5) & (-5)(-1)+(7)(1) & (-5)(3)+(7)(4) & (-5)(7)+(7)(-1) \\ (0)(2)+(-3)(-5) & (0)(-1)+(-3)(1) & (0)(3)+(-3)(4) & (0)(7)+(-3)(-1) \end{bmatrix} =$$

$$\begin{bmatrix} -18 & 3 & 19 & 3 \\ -45 & 12 & 13 & -42 \\ 15 & -3 & -12 & 3 \end{bmatrix}$$

The next thing we need to think about is the fact that we can't always take the product of two matrices. While we don't need matrices A and B to be the same size in order to calculate the product BA, they do need to have *compatible* sizes. If you go through the fine print in the definition of matrix multiplication, you'll notice that the value n appeared in the size of both A and B. This corresponds to the fact that in order to calculate the product BA, the number of columns in B must be the same as the number of rows in A. When you look at any specific matrices, this fact should be immediately obvious, as we can only take the dot product of vectors that are the same size. Just to throw one last fact at you, if B is an $m \times n$ matrix, and A is an $n \times p$ matrix, then BA will be an $m \times p$ matrix. This creates interesting scenarios where you might be taking the product of two seemingly large matrices (say B is 2×8 and A is 8×3), but their product is much smaller (in this case, BA would be 2×3), and that you can start with two seemingly small matrices (say B is 3×1 and A is 1×4), but their product is much larger (in this case, BA is 3×4). Let's look at some more examples:

Example: $\begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -5 & 6 \\ 9 & -8 & 7 & 2 \end{bmatrix} =$

$$\begin{bmatrix} (2)(1)+(3)(9) & (2)(0)+(3)(-8) & (2)(-5)+(3)(7) & (2)(6)+(3)(2) \\ (-1)(1)+(4)(9) & (-1)(0)+(4)(-8) & (-1)(-5)+(4)(7) & (-1)(6)+(4)(2) \end{bmatrix} =$$

$$\begin{bmatrix} 29 & -24 & 11 & 18 \\ 35 & -32 & 33 & 2 \end{bmatrix}$$

But the product $\begin{bmatrix} 1 & 0 & -5 & 6 \\ 9 & -8 & 7 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}$ does not exist, because the number of columns in the matrix on the left is not the same as the number of rows in the matrix on the right.

Example: $\begin{bmatrix} 1 & -1 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ -8 \end{bmatrix} = [2 + 3 - 40] = [-35]$, but

$$\begin{bmatrix} 2 \\ -3 \\ -8 \end{bmatrix} \begin{bmatrix} 1 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 10 \\ -3 & 3 & -15 \\ -8 & 8 & -40 \end{bmatrix}$$

Not only are these great examples of how to perform matrix multiplication (if I do say so myself), they also show off one of the other fascinating properties of matrix multiplication: MATRIX MULTIPLICATION IS NOT COMMUTATIVE. That is, we do not always have that $AB = BA$. In fact, the first example shows that BA may not even be defined, even if AB is defined. And even if both AB and BA are defined (as in the second example) they may not even be the same size. Now, both of those problems can be overcome if A and B are square matrices (of the same size). But even in that perfect world, we can not guarantee that $AB = BA$. Here's one last example for you:

Example: $\begin{bmatrix} 3 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ -1 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 12 \\ 2 & 8 \end{bmatrix}$, but

$$\begin{bmatrix} 1 & 4 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 11 & 0 \\ -11 & 0 \end{bmatrix}.$$