Math 331
Notes on Matrix Multiplication

Since matrix multiplication is complicated, we should expect that there are several ways to view it. We will examine three approaches: by entry, by row, and by column.

0.1 Thinking about entries of AB

Recall that we defined the product AB of an $m \times n$ matrix A with an $n \times p$ matrix B by specifying how to compute each entry. Specifically, for each choice of i and j,

$$\begin{aligned}
(AB)_{ij} &= \sum_{k=1}^{n} a_{ik}b_{kj} \\
&= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \\
&= [a_{i1}, a_{i2}, \ldots, a_{in}] \bullet [b_{1j}, b_{2j}, \ldots, b_{nj}] \\
&= (\text{row } i \text{ of } A)(\text{column } j \text{ of } B)
\end{aligned}$$

That is, the $i, j - \mathit{entry}$ of AB is just the dot product of the vectors forming the i^{th} row of A and the j^{th} column of B. Thus in forming AB, we work across the rows of A and down the column of B

If A and B are the matrices

$$A = \begin{pmatrix} 4 & -2 & 5 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & 0 \\ -4 & 8 \\ 1 & 1 \end{pmatrix}$$

then AB is the 2×2 matrix

$$\begin{aligned}
AB &= \begin{pmatrix} 4 & -2 & 5 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ -4 & 8 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \begin{bmatrix} 4 & -2 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} & \begin{bmatrix} 4 & -2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 8 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 8 \\ 1 \end{bmatrix} \end{pmatrix} \\
&= \begin{pmatrix} 25 & -11 \\ 5 & 2 \end{pmatrix}
\end{aligned}$$

0.2 Thinking about rows of AB

What parts of A and B contribute to the the i^{th} row of AB? The i^{th} row of AB is $[(AB)_{i1}, (AB)_{i2}, \ldots, (AB)_{ip}]$. Notice that the entries in the i^{th} row of

AB always have row index i, and thus only use row i of A. Notice that there are p entries in the i^{th} row of AB,which means that a term $(AB)_{ij}$ occurs for each value of j with $1 \leq j \leq p$. When $j = 1$,the first column of B is used to produce $(AB)_{i1}$, the first entry in row i of AB. When $j = 2$, the second column of B is used to produce $(AB)_{i2}$, the second entry in row i of AB. Continuing to the end of the row, when $j = p$, column p of B is used to produce $(AB)_{ip}$, the last entry in row i of AB. Thus, row i of AB uses only row i of A but all of matrix B.

$$(\text{ row } i \text{ of } AB\ ) = (\text{ row } i \text{ of } A\ )\,(\text{matrix } B)$$

Suppose that A has three rows, so that AB has 3 rows. Then

$$\begin{aligned} AB &= \begin{pmatrix} \text{row 1 of } AB \\ \text{row 2 of } AB \\ \text{row 3 of } AB \end{pmatrix} \\ &= \begin{pmatrix} (\text{row 1 of } A)\,(\text{matrix } B) \\ (\text{row 2 of } A)\,(\text{matrix } B) \\ (\text{row 3 of } A)\,(\text{matrix } B) \end{pmatrix} \end{aligned}$$

0.3 Thinking about columns of AB

What parts of A and B contribute to the the j^{th} column of AB? Notice that the entries in the j^{th} column of AB always have column index j, and thus only use column j of B..The j^{th} column of AB has m rows, so all values of i are used. When $i = 1$, the first row of A is used to produce $(AB)_{1j}$, the first entry in column j of AB. When $i = 2$, the second row of A is used to produce $(AB)_{2j}$, the second entry in column j of AB. Continuing to the end of the column, when $i = m$, the m^{th} row of A is used to produce $(AB)_{mj}$, the last entry in column j of AB.. Thus, column j of AB uses all of matrix A but only column j of B.

$$\begin{pmatrix} \text{column } j \\ \text{of} \\ AB \end{pmatrix} = (\text{matrix } A) \begin{pmatrix} \text{column } j \\ \text{of} \\ B \end{pmatrix}$$

Suppose that B has 2 columns, so AB has 2 columns

$$\begin{aligned} AB &= \left(\begin{bmatrix} \text{column 1} \\ \text{of} \\ AB \end{bmatrix}, \begin{bmatrix} \text{column 2} \\ \text{of} \\ AB \end{bmatrix} \right) \\ &= \left((\text{matrix } A) \begin{bmatrix} \text{column 1} \\ \text{of} \\ B \end{bmatrix}, (\text{matrix } A) \begin{bmatrix} \text{column 2} \\ \text{of} \\ B \end{bmatrix} \right) \end{aligned}$$

0.4 Multiplying a vector and a matrix

In the preceding two sections, we saw that the product AB can be built up using products of the form

$$(\text{row of first matrix})\,(\text{second matrix})$$

or by using products of the form

$$(\text{first matrix})\begin{pmatrix}\text{column}\\ \text{of}\\ \text{second matrix}\end{pmatrix}.$$

Thus we should examine these constructions carefully. Both constructions are useful, the former for understanding row operations, the later, for understanding solution sets for linear systems.

0.5 When a row vector multiplies a matrix

For convenience, suppose that our row vector u is 1×3, and that B is 3×2. Thus,

$$\begin{aligned}
uB &= [u_1, u_2, u_3]\begin{pmatrix} b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{32}\end{pmatrix} = \left([u_1, u_2, u_3]\begin{bmatrix} b_{11}\\ b_{21}\\ b_{31}\end{bmatrix}, [u_1, u_2, u_3]\begin{bmatrix} b_{12}\\ b_{22}\\ b_{32}\end{bmatrix}\right)\\
&= (u_1b_{11} + u_2b_{21} + u_3b_{31}, u_1b_{12} + u_2b_{22} + u_3b_{32})\\
&= (u_1b_{11}, u_1b_{12}) + (u_2b_{21}, u_2b_{22}) + (u_3b_{31}, u_3b_{32})\\
&= u_1(b_{11}, b_{12}) + u_2(b_{21}, b_{22}) + u_3(b_{31}, b_{32})\\
&= u_1(\text{row 1 of } B) + u_2(\text{row 2 of } B) + u_3(\text{row 3 of } B)
\end{aligned}$$

Note that uB is a 1×2 row vector. The first entry of uB is the dot product of the vector u and the first column of B. The second entry of uB is the dot product of the vector u and the second column of B. Notice that when we compute these dot products, we always multiply entries in the first row of B by u_1, entries in the second row of B by u_2, and entries in the third row of B by u_3 We could view these products as u_1 scaling the first row of B, u_2 as scaling the second row of B, and u_3 scaling the third row of B. Also note that since we are adding down the columns, we are adding the product of u_1 with the first row of B to the product of u_2 with the second row of B, and then adding that to the product of u_3 with the third row of B.. That is, we are summing three scaled vectors, namely the three rows of B scaled by the three entries of u. Thus the row vector uB is actually a linear combination of the row of B using the scalars in u.

As an example, suppose that $u = [4, -2, 5]$. Then uB is the row vector obtained as a linear combination of four copies of the first row of B minus two copies of the second row of B plus five copies of the third row of B.

0.6 Reprise: Thinking about rows of AB

Since each row in the matrix product AB is obtained as the product of some row of A with all of the matrix B, we can use the linear combination approach to describe matrix multiplication. Suppose that A is the 3×2 matrix given by

$$A = \begin{pmatrix} 4 & -2 & 5\\ 1 & 0 & 2\end{pmatrix}$$

and that B is some $3 \times p$ matrix. Then the first row of AB is built using the first row of A and all of B. Specifically, the first row of AB is the row vector obtained as a linear combination of four copies of the first row of B minus two copies of the second row of B plus five copies of the third row of B. The second row of AB is the row vector obtained as a linear combination of one copy of the first row of B added to two copies of the third row of B. Thus each row of AB is a linear combination of the rows of B.

As an example, if B is

$$B = \begin{pmatrix} 3 & 0 \\ -4 & 8 \\ 1 & 1 \end{pmatrix},$$

then

$$\begin{aligned} AB &= \begin{pmatrix} (\text{row 1 of } A)\, B \\ (\text{row 2 of } A)\, B \end{pmatrix} = \begin{pmatrix} (4,-2,5)\, B \\ (1,0,2)\, B \end{pmatrix} \\ &= \begin{pmatrix} 4\,(\text{row 1 of } B) - 2\,(\text{row 2 of } B) + 5\,(\text{row 3 of } B) \\ 1\,(\text{row 1 of } B) + 0\,(\text{row 2 of } B) + 2\,(\text{row 3 of } B) \end{pmatrix} \\ &= \begin{pmatrix} 4\,(3,0) - 2\,(-4,8) + 5\,(1,1) \\ 1\,(3,0) + 0\,(-4,8) + 2\,(1,1) \end{pmatrix} = \begin{pmatrix} (25,-11) \\ (5,2) \end{pmatrix} = \begin{pmatrix} 25 & -11 \\ 5 & 2 \end{pmatrix} \end{aligned}$$

0.7 When a matrix multiplies a column vector

For convenience, suppose that A is 2×3, and that our column vector v is 3×1. Thus,

$$\begin{aligned} Av &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{pmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \\ \begin{bmatrix} a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}v_1 \\ a_{21}v_1 \end{pmatrix} + \begin{pmatrix} a_{12}v_2 \\ a_{22}v_2 \end{pmatrix} + \begin{pmatrix} a_{13}v_3 \\ a_{23}v_3 \end{pmatrix} \\ &= v_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + v_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} \\ &= v_1 \begin{pmatrix} \text{column 1} \\ \text{of} \\ A \end{pmatrix} + v_2 \begin{pmatrix} \text{column 2} \\ \text{of} \\ A \end{pmatrix} + v_3 \begin{pmatrix} \text{column 3} \\ \text{of} \\ A \end{pmatrix} \end{aligned}$$

Note that Av is a 2×1 column vector. The first entry of Av is the dot product of the first row of A and the vector v. The second entry of Av is the dot product of the second row of A and the vector v. Notice that when we compute these dot products, we always multiply v_1 by entries in the first column of A, v_2 by entries in the second column of A, and v_3 by entries in the third column of A. We could view these products as v_1 scaling the first column of A, v_2 as scaling the second column of A, and v_3 scaling the third column of A. Also note that since we are adding across the rows, we are adding the product of v_1 with the first column of A to the product of v_2 with the second column of A, and then adding that to the product of v_3 with the third column of A.. That is, we are summing three scaled vectors, namely the three columns of A scaled by the three entries of v. Thus the column vector Av is a linear combination of the columns of A using the scalars in v

As an example, suppose that $v = [3, -4, 1]^T$. Then Av is the column vector obtained as a linear combination of three copies of the first column of A minus four copies of the second row of A plus one copy of the third row of A.

0.8 Reprise: Thinking about columns of AB

Since each column in the matrix product AB is obtained as the product of all of the matrix A with some column of the matrix B, we can use the linear combination approach to describe matrix multiplication. Suppose that B is the 3×2 matrix given by

$$B = \begin{pmatrix} 3 & 0 \\ -4 & 8 \\ 1 & 1 \end{pmatrix},$$

then the first column of AB is built using all of A and the first column of B. Specifically, the first column of AB is the column vector obtained as the linear combination of three copies of the first column of A minus four copies of the second column of A plus one copy of the third column of A. The second column of AB is eight copies of the second column of A plus one copy of the third third column of A. Thus each column of AB is a linear combination of the columns in A.

As an example, if A is

$$A = \begin{pmatrix} 4 & -2 & 5 \\ 1 & 0 & 2 \end{pmatrix},$$

then

$$
\begin{aligned}
AB &= \left(A \begin{bmatrix} \text{column 1} \\ \text{of} \\ B \end{bmatrix}, A \begin{bmatrix} \text{column 2} \\ \text{of} \\ B \end{bmatrix} \right) = \left(A \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}, A \begin{bmatrix} 0 \\ 8 \\ 1 \end{bmatrix} \right) \\
&= \left(3 \begin{bmatrix} \text{column 1} \\ \text{of} \\ A \end{bmatrix} - 4 \begin{bmatrix} \text{column 2} \\ \text{of} \\ A \end{bmatrix} + 1 \begin{bmatrix} \text{column 3} \\ \text{of} \\ A \end{bmatrix}, \right. \\
&\qquad \left. 0 \begin{bmatrix} \text{column 1} \\ \text{of} \\ A \end{bmatrix} + 8 \begin{bmatrix} \text{column 2} \\ \text{of} \\ A \end{bmatrix} + 1 \begin{bmatrix} \text{column 3} \\ \text{of} \\ A \end{bmatrix} \right) \\
&= \left(3 \begin{bmatrix} 4 \\ 1 \end{bmatrix} - 4 \begin{bmatrix} -2 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 5 \\ 2 \end{bmatrix}, 0 \begin{bmatrix} 4 \\ 1 \end{bmatrix} + 8 \begin{bmatrix} -2 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 5 \\ 2 \end{bmatrix} \right) \\
&= \left(\begin{bmatrix} 25 \\ 5 \end{bmatrix}, \begin{bmatrix} -11 \\ 2 \end{bmatrix} \right) = \begin{pmatrix} 25 & -11 \\ 5 & 2 \end{pmatrix}
\end{aligned}
$$

0.9 Exercises

1. Let A and B be the matrices
$$A = \begin{pmatrix} 2 & -1 & 1 \\ 3 & 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 5 & 9 & 1 \\ 0 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}.$$
Compute the product AB three ways:
   (a) one entry at a time.
   (b) one row at a time using the fact that each row of AB is a linear combination of the rows of B.
   (c) one column at a time using the fact that each column of AB is a linear combination of the columns of A.

2. Let A be a generic $4 \times n$ matrix.
   (a) Find a matrix E such that the matrix EA is the matrix obtained from A by scaling the third row of A by -9 and leaving all other rows unchanged.
   (b) Find a matrix F such that the matrix FA is the matrix obtained from A by switching rows 2 and 4, and leaving all other rows unchanged.
   (c) Find a matrix G such that the matrix GA is the matrix obtained from A by adding twice the fourth row of A to the second row of A, and leaving all other rows unchanged.

3. Explain the statement "The linear system $Ax = b$ has a solution exactly when b is a linear combination of the columns of A."

4. Suppose that the rows of M are linear combinations of the rows of N. Explain why that means there is a matrix P such that $M = PN$.