- There are some matrices $\mathbf{A}$ for which $\mathbf{A}^T = \mathbf{A}$. Such matrices are said to be **symmetric**.

1.3 Matrix Multiplication and Systems of Linear Equations

1.3.1 Several interpretations of matrix multiplication

In the previous section we saw what is required (in terms of matrix dimensions) in order to be able to produce the product $\mathbf{AB}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$, and we saw how to produce this product. There are several useful ways to conceptualize this product, and in this first sub-section we will investigate them. We first make a definition.

Definition 3: Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ be matrices all having the same dimensions. For each choice of real numbers $c_1, \ldots, c_k$, we call

$$c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_k\mathbf{A}_k$$

a **linear combination** of the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$. The set of all such linear combinations

$$S := \{c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_k\mathbf{A}_k \,|\, c_1, \ldots, c_k \in \mathbb{R}\}$$

is called the **linear span** (or simply **span**) of the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$. We sometimes write $S = \text{span}(\{\mathbf{A}_1, \ldots, \mathbf{A}_k\})$.

Here, now, are several different ways to think about product $\mathbf{AB}$ of two appropriately sized matrices $\mathbf{A}$ and $\mathbf{B}$.

1. **Block multiplication**. This is the first of four descriptions of matrix multiplication, and it is the most general. In fact, each of the three that follow is a special case of this one.

   Any matrix (table) may be separated into **blocks** (or *submatrices*) via horizontal and vertical lines. We first investigate the meaning of matrix multiplication at the block level when the left-hand factor of the matrix product $\mathbf{AB}$ has been subdivided using only vertical lines, while the right-hand factor has correspondingly been blocked using only horizontal lines.

Example 2:

Suppose

$$\mathbf{A} = \left[\begin{array}{cc|c|cc} 8 & 8 & 3 & -4 & 5 \\ 6 & -6 & 1 & -8 & 6 \\ 5 & 3 & 4 & 2 & 7 \end{array}\right] = \left[\begin{array}{c|c|c} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 \end{array}\right]$$

(Note how we have named the three blocks found in **A**!), and

$$\mathbf{B} = \left[\begin{array}{cccc} -3 & 5 & -5 & -2 \\ 2 & -2 & 2 & -7 \\ \hline -6 & 6 & 0 & 3 \\ \hline -3 & 2 & -5 & 0 \\ 0 & -1 & -1 & 4 \end{array}\right] = \left[\begin{array}{c} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \\ \hline \mathbf{B}_3 \end{array}\right].$$

Then

$$\begin{aligned} \mathbf{AB} &= \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \mathbf{A}_3\mathbf{B}_3 \\ &= \begin{bmatrix} 8 & 8 \\ 6 & -6 \\ 5 & 3 \end{bmatrix}\begin{bmatrix} -3 & 5 & -5 & -2 \\ 2 & -2 & 2 & -7 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}\begin{bmatrix} -6 & 6 & 0 & 3 \end{bmatrix} + \begin{bmatrix} -4 & 5 \\ -8 & 6 \\ 2 & 7 \end{bmatrix}\begin{bmatrix} -3 & 2 & -5 & 0 \\ 0 & -1 & -1 & 4 \end{bmatrix} \\ &= \begin{bmatrix} -8 & 24 & -24 & -72 \\ -30 & 42 & -42 & 30 \\ -9 & 19 & -19 & -31 \end{bmatrix} + \begin{bmatrix} -18 & 18 & 0 & 9 \\ -6 & 6 & 0 & 3 \\ -24 & 24 & 0 & 12 \end{bmatrix} + \begin{bmatrix} 12 & -13 & 15 & 20 \\ 24 & -22 & 34 & 24 \\ -6 & -3 & -17 & 28 \end{bmatrix} \\ &= \begin{bmatrix} -14 & 29 & -9 & -43 \\ -12 & 26 & -8 & 57 \\ -39 & 40 & -36 & 9 \end{bmatrix}. \end{aligned}$$

■

While we were trying to keep things simple in the previous example by drawing only vertical lines in **A**, the number and locations of those vertical lines was somewhat arbitrary. Once we chose how to subdivide **A**, however, the horizontal lines in **B** had to be drawn to create blocks with rows as numerous as the columns in the blocks of **A**.

Now, suppose we subdivide the left factor with *both* horizontal and vertical lines. Say that

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \\ \hline \mathbf{A}_{31} & \mathbf{A}_{32} \end{array}\right].$$

Where the vertical line is drawn in **A** continues to dictate where a horizontal line must be drawn in the right-hand factor **B**. On the other hand, if we draw any vertical lines in to create blocks in the right-hand factor **B**, they can go anywhere, paying no heed to where the horizontal lines appear in **A**. Say that

$$\mathbf{B} = \left[\begin{array}{c|c|c|c} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} & \mathbf{B}_{14} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} & \mathbf{B}_{24} \end{array}\right].$$

Then

$$\begin{aligned} \mathbf{AB} &= \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \\ \hline \mathbf{A}_{31} & \mathbf{A}_{32} \end{array}\right]\left[\begin{array}{c|c|c|c} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} & \mathbf{B}_{14} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} & \mathbf{B}_{24} \end{array}\right] \\ &= \left[\begin{array}{c|c|c|c} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} & \mathbf{A}_{11}\mathbf{B}_{13} + \mathbf{A}_{12}\mathbf{B}_{23} & \mathbf{A}_{11}\mathbf{B}_{14} + \mathbf{A}_{12}\mathbf{B}_{24} \\ \hline \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} & \mathbf{A}_{21}\mathbf{B}_{13} + \mathbf{A}_{22}\mathbf{B}_{23} & \mathbf{A}_{21}\mathbf{B}_{14} + \mathbf{A}_{22}\mathbf{B}_{24} \\ \hline \mathbf{A}_{31}\mathbf{B}_{11} + \mathbf{A}_{32}\mathbf{B}_{21} & \mathbf{A}_{31}\mathbf{B}_{12} + \mathbf{A}_{32}\mathbf{B}_{22} & \mathbf{A}_{31}\mathbf{B}_{13} + \mathbf{A}_{32}\mathbf{B}_{23} & \mathbf{A}_{31}\mathbf{B}_{14} + \mathbf{A}_{32}\mathbf{B}_{24} \end{array}\right]. \end{aligned}$$

Example 3:

Suppose **A**, **B** are the same as in Example 2. Let's subdivide **A** in the following (arbitrarily chosen) fashion:

$$\mathbf{A} = \left[\begin{array}{cccc|c} 8 & 8 & 3 & -4 & 5 \\ \hline 6 & -6 & 1 & -8 & 6 \\ 5 & 3 & 4 & 2 & 7 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right].$$

Given the position of the vertical divider in **A**, we must place a horizontal divider in **B** as shown below. Without any requirements on where vertical dividers appear, we choose (again arbitrarily) not to have any.

$$\mathbf{B} = \left[\begin{array}{cccc} -3 & 5 & -5 & -2 \\ 2 & -2 & 2 & -7 \\ -6 & 6 & 0 & 3 \\ -3 & 2 & -5 & 0 \\ \hline 0 & -1 & -1 & 4 \end{array}\right] = \left[\begin{array}{c} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \end{array}\right].$$

Then

$$\begin{aligned}
\mathbf{AB} &= \left[\begin{array}{c} \mathbf{A}_{11}\mathbf{B}_1 + \mathbf{A}_{12}\mathbf{B}_2 \\ \hline \mathbf{A}_{21}\mathbf{B}_1 + \mathbf{A}_{22}\mathbf{B}_2 \end{array}\right] \\
&= \left[\begin{array}{c} \begin{bmatrix} 8 & 8 & 3 & -4 \end{bmatrix} \begin{bmatrix} -3 & 5 & -5 & -2 \\ 2 & -2 & 2 & -7 \\ -6 & 6 & 0 & 3 \\ -3 & 2 & -5 & 0 \end{bmatrix} + 5 \begin{bmatrix} 0 & -1 & -1 & 4 \end{bmatrix} \\ \hline \begin{bmatrix} 6 & -6 & 1 & -8 \\ 5 & 3 & 4 & 2 \end{bmatrix} \begin{bmatrix} -3 & 5 & -5 & -2 \\ 2 & -2 & 2 & -7 \\ -6 & 6 & 0 & 3 \\ -3 & 2 & -5 & 0 \end{bmatrix} + \begin{bmatrix} 6 \\ 7 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 & 4 \end{bmatrix} \end{array}\right] \\
&= \left[\begin{array}{rrrr} -14 & 29 & -9 & -43 \\ \hline -12 & 26 & -8 & 57 \\ -39 & 40 & -36 & 9 \end{array}\right].
\end{aligned}$$

■

2. **Sums of rank-one matrices**. Now let us suppose that $\mathbf{A}$ has n columns and $\mathbf{B}$ has n rows. Suppose also that we block (as described allowed for in the previous case above) $\mathbf{A}$ by column—one column per block—and correspondingly $\mathbf{B}$ by row:

$$\mathbf{A} = \left[\begin{array}{c|c|c|c} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_n \end{array}\right] \qquad \text{and} \qquad \mathbf{B} = \left[\begin{array}{c} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \\ \hline \vdots \\ \hline \mathbf{B}_n \end{array}\right].$$

Following Example 2, we get

$$\mathbf{AB} = \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{A}_n\mathbf{B}_n = \sum_{j=1}^{n} \mathbf{A}_j\mathbf{B}_j \,. \tag{1.1}$$

The only thing new here to say concerns the individual products $\mathbf{A}_j\mathbf{B}_j$ themselves, in which the first factor $\mathbf{A}_j$ is a vector in $\mathbb{R}^m$ and the 2nd $\mathbf{B}_j$ is the *transpose* of a vector in $\mathbb{R}^p$ (for some m and p).

So, take $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^p$. Since $\mathbf{u}$ is m-by-1 and $\mathbf{v}^T$ is 1-by-p, the product $\mathbf{uv}^T$, called the **outer product** of $\mathbf{u}$ and $\mathbf{v}$, makes sense, yielding an m-by-p matrix.

Example 4:

Given $\mathbf{u} = (-1, 2, 1)$ and $\mathbf{v} = (3, 1, -1, 4)$, their vector outer product is

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} -3 & -1 & 1 & -4 \\ 6 & 2 & -2 & 8 \\ 3 & 1 & -1 & 4 \end{bmatrix}.$$

■

If you look carefully at the resulting outer product in the previous example, you will notice it has relatively simple structure—its 2nd through 4th columns are simply scalar multiples of the first, and the same may be said about the 2nd and 3rd rows in relation to the 1st row. Later in these notes, we will define the concept of the **rank of a matrix**. Vector outer products are always matrices of rank 1 and thus, by (1.1), every matrix product can be broken into the sum of rank-one matrices.

3. **Linear combinations of columns of A**. Suppose $\mathbf{B}$ has p columns, and we partition it in this fashion (Notice that $\mathbf{B}_j$ represents the j^{th} *column* of $\mathbf{B}$ instead of the j^{th} row, as it meant above!):
$$\mathbf{B} = \left[\ \mathbf{B}_1 \mid \mathbf{B}_2 \mid \cdots \mid \mathbf{B}_p\ \right].$$
This partitioning by *vertical* lines of the right-hand factor in the matrix product $\mathbf{AB}$ does not place any constraints on how $\mathbf{A}$ is partitioned, and so we may write
$$\mathbf{AB} = \mathbf{A}\left[\ \mathbf{B}_1 \mid \mathbf{B}_2 \mid \cdots \mid \mathbf{B}_p\ \right] = \left[\ \mathbf{AB}_1 \mid \mathbf{AB}_2 \mid \cdots \mid \mathbf{AB}_p\ \right].$$
That is, for each $j = 1, 2, \ldots, p$, the j^{th} column of $\mathbf{AB}$ is obtained by left-multiplying the j^{th} column of $\mathbf{B}$ by $\mathbf{A}$.

Having made that observation, let us consider more carefully what happens when $\mathbf{A}$—suppose it has n columns $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$—multiplies a vector $\mathbf{v} \in \mathbb{R}^n$. (Note that each $\mathbf{B}_j$ is just such a vector.) Blocking $\mathbf{A}$ by columns, we have
$$\mathbf{Av} = \left[\ \mathbf{A}_1 \mid \mathbf{A}_2 \mid \cdots \mid \mathbf{A}_n\ \right] \begin{bmatrix} v_1 \\ \hline v_2 \\ \hline \vdots \\ \hline v_n \end{bmatrix} = v_1\mathbf{A}_1 + v_2\mathbf{A}_2 + \cdots + v_n\mathbf{A}_n.$$
That is, the matrix-vector product $\mathbf{Av}$ is simply a linear combination of the columns of $\mathbf{A}$, with the scalars multiplying these columns taken (in order, from top to bottom) from $\mathbf{v}$. The implication for the matrix product $\mathbf{AB}$ is that each of its columns $\mathbf{AB}_j$ is a linear combination of the columns of $\mathbf{A}$, with coefficients taken from the j^{th} column of $\mathbf{B}$.

4. **Linear combinations of rows of B**. In the previous interpretation of matrix multiplication, we begin with a partitioning of $\mathbf{B}$ via vertical lines. If, instead, we begin with a partitioning of $\mathbf{A}$, a matrix with m rows, via horizontal lines, we get

$$\mathbf{AB} = \left[\begin{array}{c} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{array}\right] \mathbf{B} \;=\; \left[\begin{array}{c} \mathbf{A}_1\mathbf{B} \\ \hline \mathbf{A}_2\mathbf{B} \\ \hline \vdots \\ \hline \mathbf{A}_m\mathbf{B} \end{array}\right].$$

That is, the j^{th} row of the matrix product $\mathbf{AB}$ is obtained from left-multiplying the entire matrix $\mathbf{B}$ by the j^{th} *row* (considered as a submatrix) of $\mathbf{A}$.

If $\mathbf{A}$ has n columns, then each $\mathbf{A}_j$ is a 1-by-n matrix. The effect of multiplying a 1-by-n matrix $\mathbf{V}$ by an n-by-p matrix $\mathbf{B}$, using a blocking-by-row scheme for $\mathbf{B}$, is

$$\mathbf{VB} \;=\; \left[\begin{array}{c|c|c|c} v_1 & v_2 & \cdots & v_n \end{array}\right] \left[\begin{array}{c} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \\ \hline \cdots \\ \hline \mathbf{B}_n \end{array}\right] \;=\; v_1\mathbf{B}_1 + v_2\mathbf{B}_2 + \cdots + v_n\mathbf{B}_n \;,$$

a linear combination of the rows of $\mathbf{B}$. Thus, for each $j = 1, \ldots, m$, the j^{th} row $\mathbf{A}_j\mathbf{B}$ of the matrix product $\mathbf{AB}$ is a linear combination of the rows of $\mathbf{B}$, with coefficients taken from the j^{th} row of $\mathbf{A}$.

1.3.2 Systems of linear equations

Motivated by **Viewpoint 3** concerning matrix multiplication—in particular, that

$$\mathbf{Ax} \;=\; x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n \;,$$

where $\mathbf{A}_1, \ldots, \mathbf{A}_n$ are the columns of a matrix $\mathbf{A}$ and $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$—we make the following definition.

Definition 4: Suppose $\mathbf{A} = \left[\begin{array}{c|c|c|c} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_n \end{array}\right]$, where each submatrix $\mathbf{A}_j$ consists of a single column (so $\mathbf{A}$ has n columns in all). The set of all possible linear combinations of these columns (also known as span($\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$))

$$\{c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_n\mathbf{A}_n \,|\, c_1, c_2, \ldots, c_n \in \mathbb{R}\} \;,$$

is called the **column space** of $\mathbf{A}$. We use the symbol col($\mathbf{A}$) to denote the column space.