More on Matrix Multiplication

Definition: Let B be an $m \times n$ matrix with rows $\vec{b}_1^T, \dots, \vec{b}_m^T$, and let A be an $n \times p$ matrix with columns $\vec{a}_1, \dots, \vec{a}_p$. Then we define BA to be the matrix whose ij-th entry is

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

Taking a closer look at the calculation $\vec{b}_i \cdot \vec{a}_j$:

$$\vec{b}_i \cdot \vec{a}_j = \begin{bmatrix} b_{i1} \\ b_{i2} \\ \vdots \\ b_{in} \end{bmatrix} \cdot \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} = b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{in}a_{nj}$$

We can rewrite this dot product using summation notation: $b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{in}a_{nj} = \sum_{k=1}^{n} b_{ik}a_{kj}$

This will give us a new definition for BA.

Definition: Let B be an $m \times n$ matrix and let A be an $n \times p$ matrix. Then the ij-th entry of BA is

$$(BA)_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj} = \sum_{k=1}^{n} (B)_{ik}(A)_{kj}$$

More on Matrix Multiplication

Theorem 3.1.3

If A,B, and C are matrices of the correct size so that the required products are defined, and $t \in \mathbb{R}$, then

1. $A(B + C) = AB + AC$
2. $t(AB) = (tA)B = A(tB)$
3. $A(BC) = (AB)C$
4. $(AB)^T = B^T A^T$

Proof

1. Suppose A,B, and C are matrices such that $A(B + C)$ is defined, and let $D = B + C$.
Then $(D)_{ij} = d_{ij} = b_{ij} + c_{ij}$.

$$\begin{aligned} (A(B+C))_{ij} &= (AD)_{ij} \\ &= \sum a_{ik}d_{kj} \\ &= \sum a_{ik}(b_{kj} + c_{kj}) \\ &= \sum (a_{ik}b_{kj} + a_{ik}c_{kj}) \\ &= \sum a_{ik}b_{kj} + \sum a_{ik}c_{kj} \\ &= (AB)_{ij} + (AC)_{ij} = (AB + AC)_{ij} \end{aligned}$$

We see that $(A(B + C))_{ij} = (AB + AC)_{ij}$ for all i, j, and thus $A(B + C) = AB + AC$. □

More on Matrix Multiplication

Theorem 3.1.3

If A,B, and C are matrices of the correct size so that the required products are defined, and $t \in \mathbb{R}$, then

1. $A(B + C) = AB + AC$
2. $t(AB) = (tA)B = A(tB)$
3. $A(BC) = (AB)C$
4. $(AB)^T = B^T A^T$

Proof

4. Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix.
Note that $(B^T)_{ik} = (B)_{ki}$ and $(A^T)_{kj} = (A)_{jk}$.

$$\begin{aligned}(B^T A^T)_{ij} &= \sum_{k=1}^{n} (B^T)_{ik}(A^T)_{kj} \\ &= \sum_{k=1}^{n} (B)_{ki}(A)_{jk} \\ &= \sum_{k=1}^{n} (A)_{jk}(B)_{ki} \\ &= (AB)_{ji}\end{aligned}$$

So, we've seen that $(B^T A^T)_{ij} = (AB)_{ji}$ for all i and j, which means that $B^T A^T = (AB)^T$. □

More on Matrix Multiplication

Notes about Matrix Multiplication Properties:

- Matrix multiplication is not commutative. That is, in general, $AB \neq BA$.
- BA may not even be defined even if AB is. Even if both AB and BA are defined they might not even be the same size.
- Sometimes $AB = BA$, so $AB \neq BA$ is not always true.
- The cancellation law does not hold. That is, if $AB = AC$, this does not necessarily mean that $B = C$.

  Example

  Let $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 8 \\ -1 & 3 \end{bmatrix}$, and $C = \begin{bmatrix} -5 & 5 \\ 8 & 6 \end{bmatrix}$.

  Then:

  $$AB = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 4 & 8 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 11 \\ 6 & 22 \end{bmatrix}$$

  and

  $$AC = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} -5 & 5 \\ 8 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 11 \\ 6 & 22 \end{bmatrix} = AB$$

  Even though B does not equal C, we still have $AB = AC$.
- Matrix division cannot be defined.
- Only the properties listed in Theorem 3.1.4. are the properties you can use regarding matrix multiplication.

More on Matrix Multiplication

Theorem 3.1.4

If A and B are $m \times n$ matrices such that $A\vec{x} = B\vec{x}$ for every $\vec{x} \in \mathbb{R}^n$, then $A = B$.

Proof

The key fact here is that $A\vec{x} = B\vec{x}$ **for all** $\vec{x}$, not just any one in particular.

Recall that $\vec{e}_j$ is the j-th standard basis vector (that is, it is a vector that has a 1 in the j-th component, and all other components are zeros), then $A\vec{e}_j = B\vec{e}_j$ for all $1 \leq j \leq n$. But what is $A\vec{e}_j$?

Well, first of all, note that it is an $n \times 1$ matrix, so we need only look at $(A\vec{e}_j)_{i1}$.

Using our summation definition, we see that $(A\vec{e}_j)_{i1} = \sum_{k=1}^{n} (A)_{ik}(\vec{e}_j)_{k1}$, but since $(\vec{e}_j)_{k1}$ is zero for $k \neq j$ (and equals 1 when $k = j$), we see that $(A\vec{e}_j)_{i1} = (A)_{ij}$ for all i.
This means that $A\vec{e}_j$ is the j-th column of A.
Similarly, $B\vec{e}_j$ is the j-th column of B.

We have $A\vec{e}_j = B\vec{e}_j$, so this means that the j-th column of A is the same as the j-th column of B. Since this is true for every j, we see that $A = B$. □

Note: The fact that $A\vec{e}_j$ is the j-th column of A is actually quite useful. More specifically, the fact that if we line up all the products $A\vec{e}_1\ A\vec{e}_2 \cdots A\vec{e}_n$, then we get another copy of A. We use this fact to define a very special matrix.

More on Matrix Multiplication

Definition: The $n \times n$ matrix $I_n = \text{diag}(1, 1, \dots, 1)$ is called the identity matrix. That is, the identity matrix is a diagonal matrix, with all the diagonal entries equal to 1.

Examples

$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $I_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Note: Often we simply use I to denote an identity matrix, with the expectation that the size of I can be determined from context.

Theorem 3.1.5

If A is any $m \times n$ matrix, then $I_m A = A = A I_n$.