Matrix multiplication

Recall that if A is an $m \times n$ matrix and $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ an n-vector, where $A = [\vec{a}_1 \ \ldots \ \vec{a}_n]$,

columns of A

Then the product is

$$A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n, \quad \text{an } m\text{-vector.}$$

In this section, we extend this to multiplication of matrices. First we show how it connects to transformations.

Composition of transformations

Suppose $T : \mathbb{R}^k \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^m$ are transformations. If we first apply T and then apply S, we get the composition, or composite, $S \circ T$.

$$\begin{tikzcd} \mathbb{R}^k \arrow[r, "T"] \arrow[rr, bend right, "S \circ T"'] & \mathbb{R}^n \arrow[r, "S"] & \mathbb{R}^m \end{tikzcd}$$

We define $S \circ T$ as follows. If $\vec{x}$ is a k-vector, then

$$S \circ T(\vec{x}) := S(T(\vec{x})).$$

Note that $T(\vec{x})$ is an n-vector, so it makes sense as input for S.

Note that $S \circ T$ takes as input k-vectors, and outputs m-vectors.

Now let $\overset{n\times k}{\downarrow}\; B$ be the matrix associated to T and $\overset{m\times n}{\downarrow}\; A$ the matrix associated to S. That is,

$$T(\vec{x}) = B\vec{x} \quad \text{and} \quad S(\vec{y}) = A\vec{y}.$$

Suppose $B = [\vec{b}_1\ \vec{b}_2\ \cdots\ \vec{b}_k]$, where $\vec{b}_i$ is the i^{th} column, and $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$. Then

$$\begin{aligned} S \circ T(\vec{x}) = S(T(\vec{x})) &= A(B\vec{x}) \\ &= A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_k\vec{b}_k) \\ &= x_1 \underbrace{A\vec{b}_1}_{\text{m-vector}} + x_2 A\vec{b}_2 + \cdots + x_k A\vec{b}_k \\ &= \underbrace{[A\vec{b}_1\ \ A\vec{b}_2\ \cdots\ A\vec{b}_k]}_{\text{matrix!}}\vec{x} \end{aligned}$$

columns

This shows that $S \circ T$ is the transformation induced by the $m \times k$ matrix

$$[A\vec{b}_1\ \ A\vec{b}_2\ \cdots\ A\vec{b}_k].$$

This matrix is the product of A and B:

Definition: Let A be an $m \times n$ matrix and B an $n \times k$ matrix. Write $B = [\vec{b}_1 \cdots \vec{b}_k]$, where $\vec{b}_i$ is the i^{th} column of B. The product AB is the $m \times k$ matrix defined

$$AB = [A\vec{b}_1 \quad A\vec{b}_2 \cdots A\vec{b}_k].$$

i.e. the i^{th} column is the m-vector $A\vec{b}_i$.

Note: We can only take the product AB if

$$\boxed{\#\text{ of columns of } A = \#\text{ of rows of } B}$$

We saw above that if $\vec{x}$ is a k-vector, we have:

$$A(B\vec{x}) = (AB)\vec{x}.$$

So if T_A and T_B are the associated transformations, then

$$(T_A \circ T_B)(\vec{x}) = T_A(T_B(\vec{x})) = A(B\vec{x}) = (AB)\vec{x} = T_{AB}(\vec{x}).$$

Thus $\boxed{T_A \circ T_B = T_{AB}}$

Ex: Let $A = \underbrace{\begin{bmatrix} 1 & 0 & 3 \\ 2 & 2 & -1 \\ 1 & 1 & 0 \end{bmatrix}}_{3}$, $\quad B = \left.\begin{bmatrix} 2 & 3 \\ 4 & 0 \\ -1 & -2 \end{bmatrix}\right\} 3$

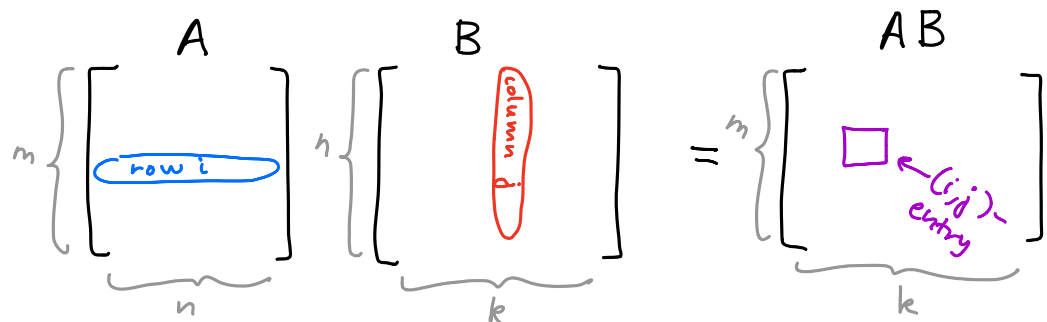
$$A\begin{bmatrix}2\\4\\-1\end{bmatrix} = \begin{bmatrix}2\\4\\2\end{bmatrix} + \begin{bmatrix}0\\8\\4\end{bmatrix} + \begin{bmatrix}-3\\1\\0\end{bmatrix} = \begin{bmatrix}-1\\13\\6\end{bmatrix}$$

$$A\begin{bmatrix}3\\0\\-2\end{bmatrix} = \begin{bmatrix}3\\6\\3\end{bmatrix} + \begin{bmatrix}0\\0\\0\end{bmatrix} + \begin{bmatrix}-6\\2\\0\end{bmatrix} = \begin{bmatrix}-3\\8\\3\end{bmatrix}$$

So $AB = \begin{bmatrix}-1 & -3\\13 & 8\\6 & 3\end{bmatrix}$

There is another way to compute the product of two matrices that computes each enty rather than each column:

Dot Product Rule: Let A be an $m \times n$ matrix and B an $n \times k$ matrix. The (i,j)-entry of AB is the dot product of row i of A with column j of B.



Ex: $A = \begin{bmatrix}1 & 2 & 3\\-1 & 0 & 4\end{bmatrix}$ $\quad B = \begin{bmatrix}4 & 6\\0 & 2\\3 & 5\end{bmatrix}$

$AB = \begin{bmatrix}13 & 25\\8 & 14\end{bmatrix}$, $\quad BA = \begin{bmatrix}-2 & 8 & 36\\-2 & 0 & 8\\-2 & 6 & 29\end{bmatrix}$

This example shows that $AB \neq BA$. That is, matrix multiplication is not commutative. It's not even commutative if A and B are both square matrices of the same size.

Ex: $A = \begin{bmatrix} -2 & 1 \\ 3 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 5 & -3 \\ -4 & -2 \end{bmatrix}$

Then $AB = \begin{bmatrix} -14 & 4 \\ 15 & -9 \end{bmatrix}$ and $BA = \begin{bmatrix} -19 & 5 \\ 2 & -4 \end{bmatrix}$

However, some pairs of matrices do commute with each other:

Ex: $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$

$AB = \begin{bmatrix} 3 & 7 \\ 0 & 3 \end{bmatrix}$, $BA = \begin{bmatrix} 3 & 7 \\ 0 & 3 \end{bmatrix}$

Ex: $A = \begin{bmatrix} 1 & 2 & 0 \\ -3 & 0 & 5 \end{bmatrix}$, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Then $AI_3 = A = I_2 A$.

This is true in general:

If A is an $m \times n$ matrix, and I_m and I_n are the $m \times m$ and $n \times n$ identity matrices, respectively, then

$$AI_n = A = I_m A.$$

Some properties of matrix multiplication:

Let a be a scalar, and A, B, C matrices of sizes such that the given products are defined. Then:

1. $A(BC) = (AB)C$ (associativity)
2. $A(B+C) = AB + AC$ and $(A+B)C = AC + BC$ (distributivity)
3. $a(AB) = (aA)B = A(aB)$.
4. $(AB)^T = B^T A^T$.

Block matrices

Sometimes it is convenient to write matrices whose entries are themselves matrices, called blocks. Such a matrix is partitioned into blocks.

For example, when we write a matrix in terms of its columns, it is a block partition.

$$A = [\vec{a}_1 \ \cdots \ \vec{a}_n] \quad \text{is a block partition of } A.$$

blocks

Ex:

$$A = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 & 1 \\ 3 & -1 & -1 & 1 & 0 \end{array}\right] = \begin{bmatrix} I_2 & O_{23} \\ P & Q \end{bmatrix}$$

$$B = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 7 & 2 \\ -1 & 0 \\ 1 & 0 \end{array}\right] = \begin{bmatrix} I_2 \\ X \end{bmatrix}$$

The product can be computed in block form:

$$AB = \begin{bmatrix} I_2 & O_{23} \\ P & Q \end{bmatrix} \begin{bmatrix} I_2 \\ X \end{bmatrix} = \begin{bmatrix} I^2 + OX \\ PI + QX \end{bmatrix}$$

$$QX = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 7 & 2 \\ -1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 2 \\ -8 & -2 \end{bmatrix}$$

So $AB = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 10 & 4 \\ -5 & -3 \end{array}\right]$.

Theorem: If matrices A and B are partitioned into blocks, AB can be computed using matrix multiplication using blocks as entries, as long as the corresponding blocks being multiplied are compatible.

Practice Problems: 2.3 : 2b, 3a, 6b, 7, 8ac, 9