Unit 012 Matrix multiplication

Slide 01: In this unit, we will discuss matrix multiplication.

Slide 02: In a previous unit, we have seen how we can add and subtract matrices as well as multiplying a scalar to a matrix. Nothing was surprising as these were done in similar fashion as real numbers.

(#)

For real number multiplication, we know that for real numbers x, y, the product xy is always defined. What about for matrices $\boldsymbol{A}$, $\boldsymbol{B}$?

Slide 03: Let us define matrix multiplication formally. Suppose $\boldsymbol{A}$ is a $m \times p$ matrix with entries a_{ij} and $\boldsymbol{B}$ is a $p \times n$ matrix with entries b_{ij}. Then the matrix $\boldsymbol{AB}$ is a $m \times n$ matrix whose (i, j)-entry is given by the expression $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}$.

(#)

To understand how the expression comes about, remember that to obtain the (i, j)-entry of $\boldsymbol{AB}$, we identify the ith row of $\boldsymbol{A}$ and the jth column of $\boldsymbol{B}$. Notice that the ith row of $\boldsymbol{A}$ contains entries a_{i1}, a_{i2} and so on till a_{ip} while the jth column of $\boldsymbol{B}$ contains entries b_{1j}, b_{2j} and so on till b_{pj}. What we do is to match the corresponding entries from the ith row of $\boldsymbol{A}$ and the jth column of $\boldsymbol{B}$. This would give us the terms in the expression for the (i, j)-entry of $\boldsymbol{AB}$.

(#)

A short hand way of writing this expression is to use the summation sign as shown here. We are summing up the expression $a_{ik}b_{kj}$ from $k = 1$ to p.

Slide 04: Let us see and example. Consider the matrix $\boldsymbol{A}$ and $\boldsymbol{B}$ shown here. Note that $\boldsymbol{A}$ is a 3×3 matrix while $\boldsymbol{B}$ is 3×2. By the definition given in the previous slide, $\boldsymbol{AB}$ would be a 3×2 matrix. But what are the entries?

(#)

To find the $(1, 1)$-entry of $\boldsymbol{AB}$, we identify the first row of $\boldsymbol{A}$ and the first column of $\boldsymbol{B}$, as shown here. The corresponding entries are matched and the added together, so we have 2×0 plus 3×1 plus -1×2. This gives us 1 and thus the $(1, 1)$-entry of $\boldsymbol{AB}$ is 1.

(#)

Moving on to the $(1, 2)$-entry, we identify the first row of $\boldsymbol{A}$ and the second column of $\boldsymbol{B}$. Matching and adding the corresponding entries, we have 2×1 plus 3×-1 plus -1×2 which gives -3. Thus the $(1, 2)$-entry of $\boldsymbol{AB}$ is -3.

(#)

Similarly, the $(2, 1)$-entry of $\boldsymbol{AB}$ is 0×0 plus 1×1 plus -1×2 which results in -1.

(#)

The $(2, 2)$-entry of $\boldsymbol{AB}$ is 0×1 plus 1×-1 plus -1×2 which results in -3.

(#)

The final two entries in $\boldsymbol{AB}$ is shown here. You may wish to verify them yourself.

Slide 05: From the way $\boldsymbol{AB}$ is defined, it is clear that for the matrix $\boldsymbol{AB}$ to be defined, the number of columns of $\boldsymbol{A}$ must be equal to the number of rows of $\boldsymbol{B}$.

(#)

This is to ensure that each time we identify a row from $\boldsymbol{A}$ and a column from $\boldsymbol{B}$, there is an equal number of entries in a row from $\boldsymbol{A}$ and a column from $\boldsymbol{B}$ so that the entries can be matched.

Slide 06: Let us look at some comparisons with real numbers multiplication. For example, if x, y are real numbers, then $xy = yx$. Is that the same for matrices?

(#)

Note that even if the sizes of $\boldsymbol{A}$ and $\boldsymbol{B}$ are such that both $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are defined, most of the time $\boldsymbol{AB}$ does not equal to $\boldsymbol{BA}$.

(#)

The following simple example illustrates this point. Here we see two 2×2 matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. Note that in this case, we know that $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are both defined. However, upon computing the matrices $\boldsymbol{AB}$ and $\boldsymbol{BA}$, we see that $\boldsymbol{AB}$ is not equal to $\boldsymbol{BA}$.

Slide 07: Another comparison with real number is the following. We know that for real numbers x and y, if $xy = 0$, then we say that either $x = 0$ or $y = 0$.

(#)

However, another simple example below shows that it is possible for $\boldsymbol{AB}$ to be the zero matrix but neither $\boldsymbol{A}$ nor $\boldsymbol{B}$ is the zero matrix.

(#)

This is another point where real numbers multiplication and matrix multiplication differ.

Slide 08: Even when the sizes of $\boldsymbol{A}$ and $\boldsymbol{B}$ are such that both $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are defined, we need to be precise when we say we want to multiply the two matrices together. To do this, we say that $\boldsymbol{AB}$ is when $\boldsymbol{A}$ is premultiplied to $\boldsymbol{B}$. Note that pre means 'before', so we have $\boldsymbol{A}$ before $\boldsymbol{B}$ in the product $\boldsymbol{AB}$.

(#)

On the other hand, we say the matrix $\boldsymbol{BA}$ is obtained when $\boldsymbol{A}$ is post-multiplied to $\boldsymbol{B}$. Here, post means 'after', so we have $\boldsymbol{A}$ after $\boldsymbol{B}$ in the prdocut $\boldsymbol{BA}$.

Slide 09: Let us introduce some laws of matrix multiplication. Firstly, this is the Associative law for matrix multiplication. When matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are such that they have the correct size to be multiplied in the following way, we see that pre-multiplying $\boldsymbol{A}$ to $\boldsymbol{BC}$ is the same as the post-multiplication of $\boldsymbol{C}$ to $\boldsymbol{AB}$.

(#)

Next is the Distributive law. The two results here shows that we can pre-multiply a matrix $\boldsymbol{A}$ into a sum of two matrices $\boldsymbol{B_1}$ and $\boldsymbol{B_2}$ or post-multiply $\boldsymbol{A}$ to the sum of two matrices $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$.

Slide 10: This next statement should be obvious if you understand how a scalar is multiplied into a matrix.

(#)

Next, for an appropriately sized zero matrix, we always have $\boldsymbol{A0} = \boldsymbol{0}$ and $\boldsymbol{0A} = \boldsymbol{0}$.

(#)

Lastly, with an appropriately sized $\boldsymbol{I}$, we have both $\boldsymbol{AI}$ and $\boldsymbol{IA}$ equal to $\boldsymbol{A}$. It should be noted that the identity matrix behaves like the number 1 in the real number system.

Slide 11: Now that we have defined matrix multiplication, it is natural to define the powers of a matrix. For a square matrix $\boldsymbol{A}$ and nonnegative integer n, we define $\boldsymbol{A}$ to the power of n as multiplying n copies of $\boldsymbol{A}$ together. Note that in this case, it is not necessary to specify it is pre or post multiplication. If $n = 0$, we define $\boldsymbol{A}$ to the power of 0 as the identity matrix.

(#)

It is now obvious that premultiplying $\boldsymbol{A}^m$ to $\boldsymbol{A}^n$ gives us $\boldsymbol{A}^{m+n}$.

(#)

One should remember that in general, $(\boldsymbol{AB})^m$ is not equal to $\boldsymbol{A}^m\boldsymbol{B}^m$. This is simply because in general $\boldsymbol{AB}$ is not equal to $\boldsymbol{BA}$.

Slide 12: This simple example will illustrate the previous point. Consider matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ as shown. The product $\boldsymbol{AB}$ is easily obtained.

(#)

Squaring $\boldsymbol{AB}$ gives the matrix two, zero, zero, two.

(#)

However, when we compute $\boldsymbol{A}^2\boldsymbol{B}^2$, we have the matrix one, zero, zero, four.

Slide 13: To summarise this unit,

(#)

We defined how matrices can be multiplied, provided their sizes are compatible.

(#)

We saw several matrix multiplication laws that hold. These include the associative law and distributive law.

(#)

Unlike addition and subtraction, matrix multiplication has some interesting differences with real numbers multiplication and we have seen some of them in this unit.