MATRIX OPERATIONS: MATRIX ADDITION, SCALAR MULTIPLICATION, AND MATRIX MULTIPLICATION

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

In this lecture, we will discuss how to build new matrices from old ones.

1. Addition of matrices

The first, and simplest operation, is that of addition. To aid our description, we will find a bit of notation helpful. We will say that a matrix is an $m \times n$ matrix if it has m rows and n columns. For example, the matrix

$$A = \begin{pmatrix} 1 & 3 & 7 \\ 5 & 0 & -2 \end{pmatrix}$$

is a 2×3 matrix. We will also use the subscripts i, j to denote the entry in the i-th row and j-th column of a matrix as A_{ij}; for example, if A is as above, then

$$A_{23} = -2.$$

Definition. Given two matrices A and B of **the same dimensions** $m \times n$, their sum $A + B$ is the $m \times n$ matrix whose entries are given by the equation

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

In other words, we simply add each of the corresponding components of the two matrices.

Example. *If*

$$A = \begin{pmatrix} 1 & 3 & 7 \\ 5 & 0 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 4 & -1 \\ 0 & 1 & 3 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 3 & 7 \\ 5 & 0 & -2 \\ 1 & 2 & 8 \end{pmatrix},$$

then

$$A + B = \begin{pmatrix} 1+2 & 3+4 & 7-1 \\ 5+0 & 0+1 & -2+3 \end{pmatrix} = \begin{pmatrix} 3 & 7 & 6 \\ 5 & 1 & 1 \end{pmatrix},$$

while $A + C$ **is not defined**, *as* A *is a* 2×3 *matrix, while* C *is a* 3×3 *matrix.*

Example. *Given two vectors, if we think of them as* $n \times 1$ *matrices, instead of tuples of numbers as originally, then matrix addition of the two vectors coincides with our first definition of vector addition.*

Date: October 11, 2016.

Example. *If A is any $m \times n$ matrix, and if 0 denotes the $m \times n$ matrix with all entries equal to zero, then $A + 0 = 0$. We call this the* **zero matrix** *(in dimensions $m \times n$). Clearly, matrix addition is always commutative ($A + B = B + A$) and associative ($A + (B + C) = (A + B) + C$), as addition of real numbers satisfies both of these properties.*

2. Scalar Multiplication

Just as we defined for vectors, we can multiply a matrix by a real number, which simply multiplies each component by that same number. That is, if $c \in \mathbb{R}$ and A is an $m \times n$ matrix, then cA is also an $m \times n$ matrix, defined by

$$(cA)_{ij} = c(A_{ij}).$$

Example. *If*

$$A = \begin{pmatrix} 1 & 3 & 7 \\ 5 & 0 & -2 \end{pmatrix},$$

then

$$2A = \begin{pmatrix} 2 & 6 & 14 \\ 10 & 0 & -4 \end{pmatrix}.$$

3. Matrix Products

We now come to our second matrix operation, which is slightly more subtle.

Definition. Given an $m \times n$ matrix A and an $n \times \ell$ matrix B, the product AB is an $m \times \ell$ matrix defined by

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

That is, the i, j-th entry of AB is the dot product of the i-th row of A and the j-th column of B.

If the number of columns of A doesn't equal the number of rows of B, then the matrix product AB will not be defined. In this situation, when we would try to take the dot product of a row of the first matrix with a column of the second matrix, we would get a dot product between two vectors of different dimensions, which isn't defined as we have seen.

Example. *If*

$$A = \begin{pmatrix} 1 & 3 & 7 \\ 5 & 0 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 & 1 & -1 \\ 4 & 1 & 1 & 5 \\ 3 & 4 & -7 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 0 & 1 & -1 \\ 4 & 1 & 1 & 5, \end{pmatrix}$$

then A is a 2×3 matrix, B is a 3×4 matrix, and C is a 2×4 matrix. Thus, AB is the 2×4 matrix given by

$$AB = \begin{pmatrix} 1\cdot 2+3\cdot 4+7\cdot 3 & 1\cdot 0+3\cdot 1+7\cdot 4 & 1\cdot 1+3\cdot 1+7\cdot(-7) & 1\cdot(-1)+3\cdot 5+7\cdot 0 \\ 5\cdot 2+0\cdot 4-2\cdot 3 & 5\cdot 0+0\cdot 1-2\cdot 4 & 5\cdot 1+0\cdot 1-2\cdot(-7) & 5\cdot(-1)+0\cdot 5-2\cdot 0 \end{pmatrix}$$
$$= \begin{pmatrix} 35 & 31 & -45 & 14 \\ 4 & -8 & 19 & -5 \end{pmatrix},$$

and AC, BA, BC, CA, and CB are all undefined.

Example. *Solving the system of linear equations in the variables $x_1, \dots, x_n$ corresponding to the augmented matrix*

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

is the same as solving the matrix equation $Ax = b$ for an unknown vector x, where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Example. *If*

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix},$$

then

$$AB = \begin{pmatrix} 1\cdot 1+1\cdot(-1) & 1\cdot 1+1\cdot(-1) \\ 2\cdot 1+2\cdot(-1) & 2\cdot 1+2\cdot(-1) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

This shows that the **product of two non-zero matrices can give the zero matrix.** *This is a property that is a little unusual, as for example, multiplication of real numbers doesn't have this property.*

As A and B are **square matrices** *of the same size, we can also consider the product BA, which we compute to be*

$$BA = \begin{pmatrix} 3 & 3 \\ -3 & -3 \end{pmatrix}.$$

Thus, matrix multiplication isn't commutative; that is, **we don't always have $AB = BA$.**

Example. *In addition to the zero matrix, which we saw above plays a distinguished role in matrix addition, we have the important* **identity matrix** I_n*, which is the square* $n\times n$ *matrix with ones along the* **diagonal** *and zeros elsewhere:*

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

The reason we call this the identity matrix is that if A *is any* $m \times n$ *matrix, then*

$$I_m A = A$$

and

$$AI_n = A.$$

In other words, multiplying on either side by I_n *leaves a matrix unchanged whenever the multiplication is defined. To see this, consider the first equation,* $I_m A = A$*. By definition, this is an* $m \times n$ *matrix, and the* i, j*-th entry is*

$$(I_m A)_{ij} = 0 + \cdots + 1 \cdot A_{ij} + \cdots + 0 = A_{ij},$$

as the i*-th row of* I_m *is* $\begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$*, where the* 1 *is in the* i*-th entry.*

Example. *The matrix product has two further simple properties which we will frequently use. These are* **associativity***:*

$$A(BC) = (AB)C,$$

and the **distributive property** *over matrix addition:*

$$A(B + C) = AB + AC.$$

That is, these two identity hold whenever the products on both sides are defined. To check associativity, assume that A *is an* $m \times n$ *matrix,* B *is an* $n \times \ell$ *matrix, and* C *is an* $\ell \times r$ *matrix. Then the* i, j*-th entry of the left hand side of the above is*

$$[A(BC)]_{ij} = \sum_{k=1}^{n} A_{ik}(BC)_{kj} = \sum_{k=1}^{n}\sum_{q=1}^{\ell} A_{ik}B_{kq}C_{qj},$$

and similarly computing the right hand side, we find

$$[(AB)C]_{ij} = \sum_{q=1}^{\ell}(AB)_{iq}C_{qj} = \sum_{q=1}^{\ell}\sum_{k=1}^{n} A_{ik}B_{kq}C_{qj} = [A(BC)]_{ij}.$$

The distributive property follows by a similar type of proof.