We are going to answer the questions raised at the end of the previous note.

Recall that to every $m \times n$ matrix A we associate a linear transformation $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and any linear transformation L is associated to some (unique) matrix A. We showed that the j-th column of A is $L(e_j)$ (when considered as a vector).

Suppose now that A and B are two $m \times n$ matrices. We have seen that the sum of two linear transformations is again a linear transformation. Thus $L_A + L_B = L_C$ for some $m \times n$ matrix C. We want to know how C can be obtained form A and B. Well, the j-th column of C is $(L_A + L_B)(e_j) = L_A(e_j) + L_B(e_j)$, i.e. the j-th column of C is the sum of the j-th column of A and the j-th column of B. In othe words, the i,j-entry of C is the sum of the i,j-entries of A and B. Therefore we define addition of matrices as follows: if $A = [a_{i,j}]$ and $B = [b_{i,j}]$ then $A + B = [c_{i,j}]$, where $c_{i,j} = a_{i,j} + b_{i,j}$. It follows that $L_A + L_B = L_{A+B}$.

Example. Let

$$A = \begin{bmatrix} 3 & -1 & 2 & 3 & -1 \\ 1 & -1 & 2 & 3 & 5 \\ 2 & -3 & 6 & 9 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 & 0 & 2 & -3 \\ 4 & 2 & -2 & 1 & -1 \\ 7 & -2 & 4 & 6 & -3 \end{bmatrix}.$$

Then

$$A + B = \begin{bmatrix} 3+(-1) & -1+2 & 2+0 & 3+2 & -1+(-3) \\ 1+4 & -1+2 & 2+(-2) & 3+1 & 5+(-1) \\ 2+7 & -3+(-2) & 6+4 & 9+6 & 4+(-3) \end{bmatrix} = \begin{bmatrix} 2 & 1 & 2 & 5 & -4 \\ 5 & 1 & 0 & 4 & 4 \\ 9 & -5 & 10 & 15 & 1 \end{bmatrix}.$$

It should be clear that the addition of matrices is commutative and associative.

Suppose now that $c \in \mathbb{R}$ is a number. We know that cL_A is a linear transformation, so $cL_A = L_C$ for some matrix C. We leave it as an exercise to show that the i,j-entry of C is equal to c times the i,j-entry of A: $c_{i,j} = c \cdot a_{i,j}$. Thus we define $c \cdot A$ to be the matrix obtained from A by multiplying each of its entries by c. We then have $cL_A = L_{cA}$.

Example. Let

$$A = \begin{bmatrix} 3 & -1 & 2 & 3 & -1 \\ 1 & -1 & 2 & 3 & 5 \\ 2 & -3 & 6 & 9 & 4 \end{bmatrix}.$$

Then

$$3A = \begin{bmatrix} 9 & -3 & 6 & 9 & -3 \\ 3 & -3 & 6 & 9 & 15 \\ 6 & -9 & 18 & 27 & 12 \end{bmatrix}.$$

Thus we can add matrices of the same size and multiply them by numbers.

Suppose now that A is an $k \times m$ matrix and B is an $m \times n$ matrix. Then the composition $L_A \circ L_B$ is defined and it is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^k$. Thus $L_A \circ L_B = L_C$ for some $k \times n$ matrix C and our next goal is to see how C can be obtained from A and B. The j-th column of C is $(L_A \circ L_B)(e_j) = L_A(L_B(e_j))$. Thus the i-th entry in the j-th column of C, i.e. $c_{i,j}$, is equal to the i-th coordinate of $L_A(L_B(e_j))$, which is the dot product of the i-th row of A and $L_B(e_j)$. Since $L_B(e_j)$ is the j-th column of B, we see that $c_{i,j}$ is the dot product of the i-th row of A and the j-th column of B. This motivates the following very important definition.

Definition. Let $A = [a_{i,j}]$ be an $k \times m$ matrix and $B = [b_{i,j}]$ be an $m \times n$ matrix. The product AB of A and B is the $k \times n$ matrix $C = [c_{i,j}]$ whose i,j-entry is the dot product of the i-th row of A and the j-th column of B, for any $1 \le i \le k$ and $1 \le j \le n$:

$$c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \ldots + a_{i,m}b_{m,j}.$$

Directly from our discussion above we have $L_A \circ L_B = L_{AB}$.

Example. Let

$$A = \begin{bmatrix} 3 & -1 & 2 & 3 & -1 \\ 1 & -1 & 2 & 3 & 5 \\ 2 & -3 & 6 & 9 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 \\ 4 & 2 \\ 7 & -2 \\ -1 & 3 \\ 4 & -2 \end{bmatrix}.$$

Then the product AB is defined (since the number of columns of A is the same as the number of rows of B) and it is a 3×2 matrix $[c_{i,j}]$, where

$c_{1,1} =$ dot product of the first row of A and the first column of B $= (3, -1, 2, 3, -1) \bullet (-1, 4, 7, -1, 4) =$

$$= -3 - 4 + 14 - 3 - 4 = 0;$$

$c_{1,2} =$ dot product of the first row of A and the second column of B $= (3, -1, 2, 3, -1) \bullet (2, 2, -2, 3, -2) =$

$$= 6 - 2 - 4 + 9 + 2 = 11;$$

$c_{2,1} =$ dot product of the second row of A and the first column of B $= (1, -1, 2, 3, 5) \bullet (-1, 4, 7, -1, 4) =$

$$= -1 - 4 + 14 - 3 + 20 = 26;$$

$c_{1,2} =$ dot product of the second row of A and the second column of B $= (1, -1, 2, 3, 5) \bullet (2, 2, -2, 3, -2) =$

$$= 2 - 2 - 4 + 9 - 10 = -5;$$

$c_{3,1} =$ dot product of the third row of A and the first column of B $= (2, -3, 6, 9, 4) \bullet (-1, 4, 7, -1, 4) =$

$$= -2 - 12 + 42 - 9 + 16 = 35;$$

$c_{3,2} =$ dot product of the third row of A and the second column of B $= (2, -3, 6, 9, 4) \bullet (2, 2, -2, 3, -2) =$

$$= 4 - 6 - 12 + 27 - 8 = 5.$$

Thus

$$AB = \begin{bmatrix} 0 & 11 \\ 26 & -5 \\ 35 & 5 \end{bmatrix}.$$

Note that if A is an $k \times m$ matrix, B is an $m \times n$ matrix and C is an $n \times l$ matrix, then

$$L_{(AB)C} = L_{AB} \circ L_C = (L_A \circ L_B) \circ L_C = L_A \circ (L_B \circ L_C) = L_A \circ L_{BC} = L_{A(BC)}$$

(we used the fact that composition of functions is associative). It follows that $(AB)C = A(BC)$, i.e. matrix multiplication is associative.

In the previous note we have seen that composition of linear transformations is distributive over addition of linear transformations. We can now translate this property to matrix multiplication and addition.

Let A, B be two $k \times m$ matrices and C, D two $m \times n$ matrices. Then $(L_A + L_B) \circ L_C = L_A \circ L_C + L_B \circ L_C$. This means that $(A+B)C = AC + BC$. Similarly, from $L_A \circ (L_C + L_D) = L_A \circ L_C + L_A \circ L_D$ we conclude that $A(C+D) = AC + AD$. In other words, multiplication of matrices is distributive over addition.

Let us repeat one more time that product AB of two matrices is only defines if the number of columns of A is the same as the number of rows of B. This condition is always satisfied if A, B are square matrices of the same size. The set of all square matrices of size n will be denoted by $\mathrm{M}_n(\mathbb{R})$:

> $\mathrm{M}_n(\mathbb{R})$ is the set of all $n \times n$ matrices.

The sum and product of any two $n \times n$ matrices is again an $n \times n$ matrix. The set $M_n(\mathbb{R})$ equipped with addition and multiplication is one of the most fundamental and most important objects in mathematics. Let us list the basic properties of these two operations on $M_n(\mathbb{R})$:

1. $(A+B)+C = A+(B+C)$ for any A, B, C in $M_n(\mathbb{R})$ (addition is associative).
2. $A+B = B+A$ for any A, B in $M_n(\mathbb{R})$ (addition is commutative).
3. $A+0 = A$ for any A in $M_n(\mathbb{R})$ (here 0 is the zero matrix).
4. for any $A \in M_n(\mathbb{R})$ there is $B \in M_n(\mathbb{R})$ such that $A+B=0$ (i.e we can subtract; $B=-A$).
5. $(AB)C = A(BC)$ for any A, B, C in $M_n(\mathbb{R})$ (multiplication is associative).
6. $AI = IA = A$ for any A in $M_n(\mathbb{R})$ (here $I = I_n$ is the identity matrix).
7. $(A+B)C = AC+BC$ and $C(A+B) = CA+CB$ for any A, B, C in $M_n(\mathbb{R})$ (multiplication distributes over addition).
8. if $n > 1$ then **multiplication is not commutative i.e. there exists matrices A, B such that** $AB \neq BA$.

Let us justify the last property: take

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The above example also shows that the product of two non-zero matrices can be zero!

We have seen in the previous note that if a linear transformation is a bijection then the inverse function is also a linear transformation. We know that L_A is a bijection if and only if A is a square matrix of size n such that $\mathrm{rank}(A) = n$. We will call such A an **invertible matrix**. The inverse function to L_A, being linear, is of the form L_B for some matrix B. From $L_A \circ L_B = I$ and $L_B \circ L_A = I$ we get that $AB = I = BA$. We will denote B by A^{-1} and call it the **inverse of** A. Thus $L_A^{-1} = L_{A^{-1}}$. In the next note we will learn how to find A^{-1} for a given invertible matrix A.

Note that if both A and B are invertible then $(L_A \circ L_B)^{-1} = L_B^{-1} \circ L_A^{-1}$. In other words, $L_{(AB)^{-1}} = L_{B^{-1}A^{-1}}$ and consequently $(AB)^{-1} = B^{-1}A^{-1}$. This is a very important property (which only manifests itself when multiplication is not commutative).

If A and B are invertible $n \times n$ matrices then AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

This property can be easily justified directly:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$$

In general, if $A_1, A_2, \ldots, A_k$ are invertible then

$$(A_1A_2\ldots A_k)^{-1} = A_k^{-1}A_{k-1}^{-1}\ldots A_1^{-1}.$$

We end this note by introducing a few names for some spacial types of square matrices.

The main diagonal of a square $n \times n$ matrix $A = [a_{i,j}]$ consists of entries $a_{1,1}, a_{2,2}, \ldots, a_{n,n}$. Thus we can say that the identity matrix is the matrix with 1's on the main diagonal and zero everywhere else.

A **diagonal matrix** is a square matrix with all the entries outside the main diagonal equal to 0.
An **upper triangular** matrix is a square matrix with all entries below the main diagonal equal to 0.
A **lower triangular** matrix is a square matrix with all entries above the main diagonal equal to 0.

It is easy to see that diagonal matrices are exactly the matrices which are both lower triangular and upper triangular.