We continue our discussion of matrix multiplication. When one of the matrices in the product is the zero matrix then clearly the product is also the zero matrix: $A \cdot 0 = 0 \cdot A = 0$ (assuming, of course, that the product is defined). Let us consider now the case when A is a square $n \times n$ matrix with only one non-zero entry, i.e. A has some a in the s,t-entry and zero everywhere else. Let $B = [b_{i,j}]$ be an arbitrary $n \times k$ matrix. We want to see what AB is.

Example. Suppose that

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & -2 & 3 \\ 7 & 5 & 2 & 1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 7a & 5a & 2a & a \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

In this example $n = 3$, $k = 4$, $s = 2$, $t = 3$. We see that 2nd row of AB is equal to a times the third row of B and all the other rows are 0.

In general, since all rows of A except the s-th row are zero, all rows of AB except the s-th row will be zero (as dot product of any column of B with a zero row of A will be equal to 0). The i-th entry in the s-th row of AB will be the dot product of the s-th row of A and the i-th column of B. But the s-th row of A has only one non-zero entry, so this dot product will by equal to $ab_{t,i}$. In other words, the i-th entry in the s-th row of AB is equal to a times i-th entry in the t-th row of B. Or, equivalently, the s-th row of AB is equal to a times the t-th row of B and all the other rows of AB are zero.

Suppose not that $s \neq t$. Then $B + AB$ is obtained from B by adding to the s-th row of B the t-th row of B multiplied by a. In our example above, we have

$$B + AB = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & -2 & 3 \\ 7 & 5 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 7a & 5a & 2a & a \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2+7a & 1+5a & -2+2a & 3+a \\ 7 & 5 & 2 & 1 \end{bmatrix}.$$

On the other hand, $B + AB = (I + A)B$. In other words, multiplication by $I + A$ on the left has the same effect on B as performing the elementary row operation $E_{s,t}(a)$. This important observation prompts the following definition.

Definition. If $s \neq t$ are positive integers then $E_{s,t}(a)$ is the square matrix with s,t-entry equal to a, all the entries on the main diagonal equal to 1, and all other entries equal to 0.

Note that we do not specify the size of the square matrix $E_{s,t}(a)$ in the above definition. This size is usually clear from the context (it is determined by the requirement that matrix multiplication is defined). In the example above, $I + A = E_{2,3}(a)$ is the 3×3 matrix

$$E_{2,3}(a) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & a \\ 0 & 0 & 1 \end{bmatrix}$$

but $E_{2,3}(a)$ could also denote the 5×5 matrix

$$E_{2,3}(a) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & a & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Our discussion above can be now summarized as follows:

For any $n \times k$ matrix B, the product $E_{s,t}(a)B$ is obtained from B by performing the elementary row operation $E_{s,t}(a)$ on B.

We know that elementary row operations $E_{s,t}(a)$ and $E_(s,t)(-a)$ are inverse of each other. In particular, it follows that $E_{s,t}(-a)E_{s,t}(a)I = I$. In other words, we have

$$E_{s,t}(-a)E_{s,t}(a) = I = E_{s,t}(a)E_{s,t}(-a), \text{ i.e. } E_{s,t}(a)^{-1} = E_{s,t}(-a).$$

After discovering that elementary row operations of the type $E_{i,j}(a)$ can be given in terms matrix multiplication it is natural to ask whether the operations $D_i(a)$ and $S_{i,j}$ can also be realized as matrix multiplication. It is not hard to see that the answer is "yes". Suppose that there is a matrix $D_i(a)$ such that $D_(a)B$ is obtained from B by multiplying the i-th row of B by a. Taking B to be the identity matrix we see that the matrix $D_i(a) = D_i(a)I$ is obtained from the identity matrix via the elementary row operation $D_i(a)$. Thus $D_i(a)$ is the diagonal matrix with a in the i,i-entry and all other diagonal entries equal to 1:

$D_i(a)$ is the square (diagonal) matrix with a in the i,i-entry, 1 in all other entries on the main diagonal, and zero everywhere else. For any $n \times k$ matrix B, the product $D_i(a)B$ is obtained from B by performing the elementary row operation $D_i(a)$ on B.

Note that, as in the case of $E_{s,t}(a)$, we do not specify the size of the matrix $D_i(a)$. For example, $D_3(a)$ can mean any of the following matrices:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \ldots.$$

As $D_i(1/a)$ is the inverse operation to $D_i(a)$, we have $D_i(a)^{-1} = D_i(1/a)$.

An exercise in the first note states that $D_j(-1)E_{i,j}(1)E_{j,i}(-1)E_{i,j}(1)M = S_{i,j}M$ for any matrix M. Since all the elementary row operations on the left are given by matrix multiplication, the elementary row operation $S_{i,j}$ is given by multiplication by the matrix $D_j(-1)E_{i,j}(1)E_{j,i}(-1)E_{i,j}(1)$, which we will denote also by $S_{i,j}$. Taking $M = I$, we see that $S_{i.j}$ is the matrix obtained from the identity matrix by performing the elementary row operation $S_{i,j}$. As before, we do not specify the size of the matrix $S_{i,j}$. For example, $S_{2,3}$ can be any of the following matrices:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \ldots.$$

Clearly $S_{i,j}S_{i,j} = I$,i.e. $S_{i,j}^{-1} = S_{i,j}$.

$S_{i,j}$ is the square (diagonal) matrix obtained by switching the i-th and j-th rows of the identity matrix. For any $n \times k$ matrix B, the product $S_{i,j}B$ is obtained from B by performing the elementary row operation $S_{i,j}$ on B. We have $S_{i,j}^{-1} = S_{i,j}$.

The matrices of the form $E_{i,j}(a)$, $D_i(a)$, $S_{i,j}$ are called **elementary matrices**.

Suppose now that A is a square matrix such that L_A is a bijection. We know that this means that A is an $n \times n$ matrix with $\text{rank}(A) = n$. Furthermore, we showed that the inverse function L_A^{-1} is also a linear transformation, so it corresponds to a matrix B such that $AB = BA = I$. We are going now to describe a method to compute B for any given invertible matrix A.

Our starting point is a realization that if A is an $n \times n$ matrix of rank n then the matrix in reduced row-echelon form row equivalent to A is the identity matrix. Thus there is a sequence of elementary row operations which transforms A to the identity matrix. We have just learned that performing an elementary row operation is equivalent to multiplication by an elementary matrix on the left. It follows that there exist elementary matrices $E_1, E_2, \ldots, E_m$ such that $E_m E_{m-1} \ldots E_2 E_1 A = I$. In particular,

$$A^{-1} = E_m E_{m-1} \ldots E_2 E_1 = E_m E_{m-1} \ldots E_2 E_1 I.$$

This means that the same sequence of elementary row operations which changes A into the identity matrix, will change the identity matrix into A^{-1}. This observation justifies the following procedure to compute the matrix A^{-1}:

> Let A be an $n \times n$ invertible matrix. Construct an $n \times 2n$ matrix whose first n columns constitute the matrix A and the last n columns form the identity matrix. Perform a sequence of elementary row operations to transform this matrix into reduced row-echelon form. The first n columns of the reduced row echelon form make the identity matrix and the last n columns make the matrix A^{-1}.

Example. We follow the above procedure to compute the inverse of the matrix

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

We start with the matrix

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

and want to transform it by a sequence of elementary row operations to a matrix in reduced row-echelon form (i.e. into a matrix whose left part is the identity matrix). The right part of the resulting matrix will give us A^{-1}.

We perform the operations $E_{2,1}(-1)$, $E_{3,1}(-1)$, and $E_{4,1}(-1)$ to get

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 2 & -2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -2 & 2 & -1 & 0 & 0 & 1 \end{array}\right].$$

Now we perform $E_{3,2}(-1)$ and get

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & -2 & 2 & -1 & 0 & 0 & 1 \end{array}\right].$$

Next do $E_{4,3}(-1)$ and get

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 4 & -1 & 1 & -1 & 1 \end{array}\right].$$

Now do $D_4(1/4)$, $D_3(-1/2)$, $D_2(1/2)$ and get

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & \frac{-1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & \frac{1}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 0 & 1 & \frac{-1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{1}{4} \end{array}\right].$$

Next do $E_{3,4}(-1)$, $E_{2,4}(-1)$, $E_{1,4}(1)$ to get

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 1 & 0 & \frac{3}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & 0 & \frac{-1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} \\ 0 & 0 & 1 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{-1}{4} \\ 0 & 0 & 0 & 1 & \frac{-1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{1}{4} \end{array}\right].$$

Now do $E_{1,3}(-1)$ and get

$$\left[\begin{array}{cccc:cccc} 1 & -1 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{-1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} \\ 0 & 0 & 1 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{-1}{4} \\ 0 & 0 & 0 & 1 & \frac{-1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{1}{4} \end{array}\right].$$

Finally, do $E_{1,2}(1)$ and get

$$\left[\begin{array}{cccc:cccc} 1 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & 0 & \frac{-1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} \\ 0 & 0 & 1 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{-1}{4} \\ 0 & 0 & 0 & 1 & \frac{-1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{1}{4} \end{array}\right].$$

It follows that

$$A^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{-1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{-1}{4} \\ \frac{-1}{4} & \frac{1}{4} & \frac{-1}{4} & \frac{1}{4} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}.$$

Note that recording the elementary row operations allows us to express A^{-1} as a product of elementary matrices: $A^{-1} =$

$$E_{1,2}(1)E_{1,3}(-1)E_{3,4}(-1)E_{2,4}(-1)E_{1,4}(1)D_4\left(\frac{1}{4}\right)D_3\left(\frac{-1}{2}\right)D_2\left(\frac{1}{2}\right)E_{4,3}(-1)E_{3,2}(-1)E_{2,1}(-1)E_{3,1}(-1)E_{4,1}(-1).$$

Recall now that the inverse of a product of invertible matrices is the product of the inverses but **in the reversed order**. Thus $A = (A^{-1})^{-1}$ is equal to the following product of elementary matrices:

$$A = E_{4,1}(1)E_{3,1}(1)E_{2,1}(1)E_{3,2}(1)E_{4,3}(1)D_2(2)D_3(-2)D_4(4)E_{1,4}(-1)E_{2,4}(1)E_{3,4}(1)E_{1,3}(1)E_{1,2}(-1).$$

We used here the fact that $E_{i,j}(a)^{-1} = E_{i,j}(-a)$, $D_i(a)^{-1} = D_i(1/a)$, and $S_{i,j}^{-1} = S_{i,j}$.

In general, if $A^{-1} = E_m E_{m-1} \dots E_2 E_1$ is a product of elementary matrices then $A = E_1^{-1}E_2^{-1}\dots E_m^{-1}$ is also a product of elementary matrices. Thus, as a consequence of our discussion, we get the following important result.

Theorem. Every invertible matrix can be expressed as a product of elementary matrices.

The procedure discussed above not only allows us to find the inverse of an invertible matrix, but it also tells us how to write it as a product of elementary matrices. Note however that there are many ways of writing a given matrix as a product of elementary matrices.

Exercise. a) Prove the following equalities:

1. $E_{s,t}(a)D_i(b) = D_i(b)E_{s,t}(a)$ if $i \neq s$ and $i \neq t$.
2. $E_{s,t}(a)D_t(b) = D_t(b)E_{s,t}(ab)$.
3. $E_{s,t}(a)D_s(b) = D_s(b)E_{s,t}(a/b)$.
4. $D_s(a)E_{s,t}(1)E_{t,s}(-1)E_{s,t}(1-a)E_{t,s}(1/a)E_{s,t}(-a) = D_t(a)$

b) Prove that if A is an invertible matrix then there is $d \neq 0$ such that $D_1(1/d)A$ is a product of elementary matrices of the form $E_{i,j}(a)$.

Remark. It is not obvious at all, but true, that the number d in part b) is unique for a given invertible matrix A. This number is called **the determinant** of A. We will learn about it later in the course.