We continue our discussion of functions associated to matrices. Recall that to an $m \times n$ matrix $A = [a_{i,j}]$ we associate a function $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by

$$L_A(u) = (r_1 \bullet u, \ldots, r_m \bullet u),$$

where r_i stands for the i-th row of A. Functions of the form L_A have the following very nice properties:

1. $L_A(u + w) = L_A(u) + L_A(w)$ for any two vectors u, w in $\mathbb{R}^n$.
2. $L_A(cu) = cL_A(u)$ for any vector $u \in \mathbb{R}^n$ and any number $c \in \mathbb{R}$.

Indeed, let us justify the first property using the basic properties of dot product:

$$L_A(u + w) = (r_1 \bullet (u + w), \ldots, r_m \bullet (u + w)) = (r_1 \bullet u + r_1 \bullet w, \ldots, r_m \bullet u + r_m \bullet w) =$$
$$= (r_1 \bullet u, \ldots, r_m \bullet u) + (r_1 \bullet w, \ldots, r_m \bullet w) = L_A(u) + L_A(w)$$

The second property can be verified in a similar way.

Functions having the above properties will play an important role in this course so we make the following definition:

Definition. A function $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is called a **linear transformation** if it has the following properties:

1. $L(u + w) = L(u) + L(w)$ for any two vectors u, w in $\mathbb{R}^n$.
2. $L(cu) = cL(u)$ for any vector $u \in \mathbb{R}^n$ and any number $c \in \mathbb{R}$.

Thus, the functions associated to matrices are linear transformations. We will soon see that the converse is also true.

It is not hard to see that the defining properties of linear transformations are equivalent to the following property, which we will often use in this course:

L is a linear transformation if and only if, for any vectors $u_1, \ldots, u_k$ and any numbers $c_1, \ldots, c_k$ in $\mathbb{R}$, we have

$$L(c_1u_1 + c_2u_2 + \ldots + c_ku_k) = c_1L(u_1) + c_2L(u_2) + \ldots + c_kL(u_k).$$

It will be convenient to define now some distinguished vectors in $\mathbb{R}^n$. The **zero vector** is the vector $0 = (0, 0, \ldots, 0)$. For every i between 1 and n we define e_i to be the vector whose i-th coordinate is 1 and all other coordinates are 0. Thus

$$0 = (0, 0, \ldots, 0),\ e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0),\ e_3 = (0, 0, 1, 0, \ldots, 0),\ \ldots,\ e_n = (0, \ldots, 0, 1).$$

We will use over and over the following two straightforward properties of the vectors e_i:

- $w \bullet e_j = w_j$ for any vector $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$ and any $1 \le j \le n$.
- $w = w_1e_1 + w_2e_2 + \ldots + w_ne_n$ for any vector $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$.

Example. In $\mathbb{R}^4$ take $w = (1, -1, 2, 5)$. Then

$$w \bullet e_3 = (1, -1, 2, 5) \bullet (0, 0, 1, 0) = 2 = \text{ the third coordinate of } w$$
$$w = e_1 - e_2 + 2e_3 + 5e_4 = (1, 0, 0, 0) - (0, 1, 0, 0) + 2(0, 0, 1, 0) + 5(0, 0, 0, 1).$$

Using the above properties we make the following crucial observation, which recovers the matrix A from the linear transformation L_A:

Let $A = [a_{i,j}]$ be an $m \times n$ matrix. Then $L_A(e_j)$ is the j-th column of A (written as a row).

Indeed, the i-th coordinate of the vector $L_A(e_j)$ is, by definition, $r_i \bullet e_j$, where r_i is the i-th row of A. Now $r_i \bullet e_j$ is the j-th coordinate of the i-th row r_i, i.e. it is $a_{i,j}$. Thus

$$L_A(e_j) = (a_{1,j}, a_{2,j}, \ldots, a_{m,j})$$

which indeed is the j-th column of A written as a row.

Corollary. If A, B are $m \times n$ matrices such that $L_A = L_B$ then $A = B$.

Indeed, as $L_A(e_j) = L_B(e_j)$, we see that the j-th columns of A and B coincided for every j, so $A = B$.

We are now ready to prove that every linear transformations corresponds to some matrix.

Theorem. Let $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then $L = L_A$ for some $m \times n$ matrix A.

In order to justify this result note first that if A exists (i.e. $L = L_A$), then A must be the matrix whose j-th column is $L(e_j)$ for every j. So let A be the matrix with columns $L(e_1), L(e_2), \ldots, L(e_n)$. Then $L_A(e_j) =$ j-th column of $A = L(e_j)$ for $j = 1, 2, \ldots, n$. Now let $w = (w_1, \ldots, w_n)$ be an arbitrary vector in $\mathbb{R}^n$. Recall that $w = w_1e_1 + \ldots + w_ne_n$. Using the fact that both L and L_A are linear transformations we get

$$L(w) = L(w_1e_1 + w_2e_2 + \ldots + w_ne_n) = w_1L(e_1) + w_2L(e_2) + \ldots + w_nL(e_n) =$$

$$= w_1L_A(e_1) + w_2L_A(e_2) + \ldots + w_nL_A(e_n) = L_A(w_1e_1 + w_2e_2 + \ldots + w_ne_n) = L_A(w).$$

Thus indeed $L = L_A$.

We may summarize the above discussion as follows: the association $A \mapsto L_A$ is a bijection between the set of all $m \times n$ matrices and the set of all linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

The following corollary is a simple consequence of our discussion above:

Corollary. 1. If L_1, L_2 are two linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $L_1(e_i) = L_2(e_i)$ for $i = 1, 2, \ldots, n$ then $L_1 = L_2$.

2. For arbitrary vectors $w_1, \ldots, w_n$ in $\mathbb{R}^m$ there exists unique linear transformation $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that $L(e_1) = w_1, L(e_2) = w_2, \ldots, L(e_n) = w_n$.

Example. Is there a linear transformation $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ such that $L(e_1) = (1, 2)$ and $L(e_3) = (-1, 3)$?

Solution. We know that any linear transformation $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ is of the form L_A for some 2×3 matrix A. We also know that the i-th column of A is $L_A(e_i)$. Thus, A must be a matrix of the form

$$A = \begin{bmatrix} 1 & s & -1 \\ 2 & t & 3 \end{bmatrix}$$

for some numbers s, t. Conversely, any choice of s, t will yield a matrix A for which L_A has the required properties.

Example. The constant function from $\mathbb{R}^n$ to $\mathbb{R}^m$ which sends every vector of $\mathbb{R}^n$ to the zero vector in $\mathbb{R}^m$ is clearly a linear transformation. We will call it the **zero transformation**. It corresponds to the $m \times n$ matrix whose all entries are zero. We call it the **zero matrix** and denote by 0.

Example. The identity function $I : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $I(u) = u$ for all u, is clearly a linear transformation. The $n \times n$ matrix corresponding to I has i-th column equal to $I(e_i) = e_i$ for $i = 1, 2, \ldots, n$. We denote this matrix by I_n (or just I if there is no need to indicate the size n) and call it the **identity matrix**.

Recall that in calculus, which studies functions form $\mathbb{R}$ to $\mathbb{R}$ we produce new functions form the ones we already have by various procedure, for example by adding or composing them. We can do similar constructions with functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. Let F, G be two such functions. We define $F+G$ to be the function which sends any $u \in \mathbb{R}^n$ to $F(u)+G(u)$:

$$(F+G)(u) = F(u) + G(u).$$

For any number $c \in \mathbb{R}$ we define a function cF as follows:

$$(cF)(u) = c \cdot F(u) \text{ for any } u \in \mathbb{R}^n.$$

Thus we can add functions and multiply functions by numbers (as we do with vectors). Of course, we can also compose functions, provided the codomain of the first function is equal to the domain of the second. Our next goal is to see how these operations behave on linear transformations. The answer is given by the following, rather simple, result.

Theorem. Let $F, G : \mathbb{R}^n \longrightarrow R^m$ and $S, T : \mathbb{R}^m \longrightarrow R^k$ be linear transformations. Then

1. $F+G$ is a linear transformation.
2. cF is a linear transformation for any $c \in \mathbb{R}$.
3. $S \circ F$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^k$.
4. if $n = m$ and F is a bijection then F^{-1} is a linear transformation.
5. $(S+T) \circ F = S \circ F + T \circ F$ and $(cS) \circ F = c(S \circ F)$ for any $c \in \mathbb{R}$.
6. $S \circ (F+G) = S \circ F + S \circ G$ and $S \circ (cF) = c(S \circ F)$ for any $c \in \mathbb{R}$.

The first 4 properties state that certain function is a linear transformation. In order to justify this we need to verify that this function has the two properties in the definition of linear transformation. We will verify the first property. The second property can be verified in a similar way which we leave to the reader as an exercise.

For 1. we have

$$(F+G)(u+w) = F(u+w) + G(u+w) = F(u) + F(w) + G(u) + G(w) =$$
$$= F(u) + G(u) + F(w) + G(w) = (F+G)(u) + (F+G)(w).$$

For 2. we have

$$(cF)(u+w) = c \cdot F(u+w) = c \cdot (F(u) + F(w)) = c \cdot F(u) + c \cdot F(w) = (cF)(u) + (cF)(w).$$

For 3. we have

$$(S \circ F)(u+w) = S(F(u+w)) = S(F(u) + F(w)) = S(F(u)) + S(F(w)) = (S \circ F)(u) + (S \circ F)(w).$$

For 4. we want to show that $F^{-1}(u+w) = F^{-1}(u) + F^{-1}(w)$. Since F is one-to-one, it suffices to show that when we apply F to each side of the last equation, we get equal results:

$$F(F^{-1}(u) + F^{-1}(w)) = F(F^{-1}(u)) + F(F^{-1}(w)) = u + w = F(F^{-1}(u+w))$$

(we use the fact that $F(F^{-1}(x)) = x$ for any x).

The last two properties establish equality of certain functions. We justify the first equality of 5., the remaining equalities are handled in a similar way. To prove that two functions are equal one needs to show that they have the same value at every element of the domain. Let $u \in \mathbb{R}^n$ be an arbitrary vector. Then

$$((S+T) \circ F)(u) = (S+T)(F(u)) = S(F(u)) + T(F(u)) = (S \circ F)(u) + (T \circ F)(u) = (S \circ F + T \circ F)(u).$$

We proved that every linear transformation is of the form L_A for some matrix A. Suppose that $F = L_B$ for an $m \times n$ matrix B and $S = L_A$ for an $k \times m$ matrix A. Since $S \circ F$ is a linear transformation, we have $L_A \circ L_B = L_C$ for some $k \times n$ matrix C. How can we compute C knowing A and B? Similarly, if $G = L_D$ then, since $F+G$ is a linear transformation, we must have $L_A + L_D = L_K$ for some $m \times n$ matrix K. How can K be obtained from A and D? We will answer these questions in the next note. The answers will lead to us to a way of adding and multiplying matrices.