Linear transformations

A. Eremenko

September 5, 2024

Let U and V be two vector spaces. A map $F : U \to V$ is called a *linear transformation* if it sends sums to sums and products with numbers to products with the same numbers. More precisely,

$$F(\mathbf{x} + \mathbf{y}) = F(\mathbf{x}) + F(\mathbf{y}), \tag{1}$$

$$F(c\mathbf{x}) = cF(\mathbf{x}), \tag{2}$$

for all $\mathbf{x}, \mathbf{y}$ in U and all numbers c. V is called the *domain* of F and U the *target* of F.

In the important special case when $U = V$, linear transformations are called *linear operators*.

Examples.

1. Let $F_1(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in U$, and $F_2(\mathbf{x}) = c\mathbf{x}$. These are linear transformations. For the first one, V can be any space, for the second one $V = U$.

2. More generally, let A be an $m \times n$ matrix. Then

$$F(\mathbf{x}) = A\mathbf{x}$$

is a linear transformation from $U = \mathbf{R}^n$ to $V = \mathbf{R}^m$. We will later see that *every linear transformation between these spaces has this form.*

3. Suppose that F rotates each vector in $\mathbf{R}^2$ counterclockwise by some fixed angle θ. This is a linear transformation $\mathbf{R}^2 \to \mathbf{R}^2$. Similarly, rotation of all vectors in $\mathbf{R}^3$ by some fixed angle about a fixed axis is a linear transformation $\mathbf{R}^3 \to \mathbf{R}^3$.

4. Function $f(x) = x + c$ from $\mathbf{R}$ to $\mathbf{R}$, with fixed c is *not* a linear transformation. (Every linear transformation sends the zero vector to the zero vector!)

5. On the space of all polynomials, differentiation is a linear transformation of this space into itself. Some other examples on this space: let us fix a polynomial g. Then $F_1(f) = gf,\ F_2(f) = f \circ g$ are linear transformations. But, for example, $F(f) = f^2$, $F(f) = f + g$ $(g \neq 0)$ are not.

Our next result uses bases to describe arbitrary linear transformations between two vector spaces of finite dimension. Let V and U be vector spaces, $\dim V = n,\ \dim U = m$. Fix some bases

$$\mathbf{v}_1, \ldots, \mathbf{v}_n \quad \text{in} \quad V \quad \text{and} \quad \mathbf{u}_1, \ldots, \mathbf{u}_m \quad \text{in} \quad U.$$

We apply our transformation F to the elements of the basis v and expand the result with respect to the basis u:

$$F(\mathbf{v}_j) = \sum_{i=1}^{m} a_{i,j}\mathbf{u}_i.$$

Notice the order of subscripts in $a_{i,j}$: summation is in the *first* subscript, unlike in the matrix multiplication!

We obtain an $m \times n$ matrix $A = (a_{i,j})$ which is called the *matrix of the linear transformation A with respect to the bases u and v*. In words: columns of A are labeled by the elements of the v-basis (from the domain of F) and rows by the elements of the u-basis (in the range of F). The element $a_{i,j}$ is the coefficient at $\mathbf{u}_i$ in the expansion of $F(\mathbf{v}_j)$.

Now let $\mathbf{x}$ be a vector in V. We expand it in the v-basis

$$\mathbf{x} = \sum_{j=1}^{n} x_j\mathbf{v}_j.$$

Now we apply F and use linearity and the expansion of $F(\mathbf{v}_j)$ above:

$$F(\mathbf{x}) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{i,j}\mathbf{u}_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j}x_j \right) \mathbf{u}_i.$$

The the expressions in parentheses are the coefficients of the expansion of $F(\mathbf{x})$ in the u basis (in the target space). Thus:

The vector $\mathbf{y}$ *of coefficients of* $F(\mathbf{x})$ *in the target basis is obtained from the vector of coefficients of* $\mathbf{x}$ *in the domain basis by multiplication from the left on the matrix* A *of* F.

$$\mathbf{y} = A\mathbf{x}. \tag{3}$$

So when we choose the standard basis, the action of the linear transformation is just the multiplication of column vectors.

When F is a linear operator, that is the spaces U, V are the same, they choose the same basis in the domain and in the range.

Example. Consider the linear operator $\mathbf{R}^2 \to \mathbf{R}^2$ which rotates every vector by 90^o counterclockwise. Choose the standard basis in the plane $e_1 = (1,0)^T$, $e_2 = (0,1)^T$. The matrix of T will be 2×2. We have $F(e_1) = e_2$ and $F(e_2) = -e_1$. (Make a picture!). So the matrix of this operator is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Now what is the result of this rotation of the vector $\mathbf{x} = (3,5)^T$? The answer is

$$F(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -5 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \end{pmatrix}.$$

So the matrix of an operator F permits to compute coordinates of $F(\mathbf{x})$ in terms of coordinates of $\mathbf{x}$.

This matrix depends not only on the operator F but also on the choice of the basis! In the case of a general linear transformation it depends on the choice of two bases: one in the domain another in the target space. So we have an important question: suppose we have a linear transformation written as a matrix in some bases. What happens to this matrix if we choose another pair of bases? The question is answered using the coordinate transformation formula (2) at the end of the lecture "Vector spaces". Let $\mathbf{x}$ and $\mathbf{x}'$ be columns of coordinates of some vector in two bases of V. For convenience, we call them "old" and "new" basis, so that $\mathbf{x}$ is the column of coordinates in the old basis, and $\mathbf{x}'$ is the column of coordinates of the same vector in the new basis. Similarly in the target space U, we introduce the columns of coordinates $\mathbf{y}$ and $\mathbf{y}'$ in some two bases, old and new. Then in the old basis $\mathbf{y} = A\mathbf{x}$, and suppose that in the new basis the matrix of our operator is A', so that $\mathbf{y}' = A'\mathbf{y}$. Now the coordinate transformation formula gives

$$\mathbf{x}' = C\mathbf{x} \quad \text{and} \quad \mathbf{y}' = D\mathbf{y},$$

where C and D are matrices of expansions of the old bases in terms of the new bases. Notice that matrices C and D are invertible. So we have

$$\mathbf{y}' = D\mathbf{y} = DA\mathbf{x} = DAC^{-1}\mathbf{x}',$$

in other words

$$A' = DAC^{-1}.$$

Especially important case is the case of linear operators, when F acts from a space to itself, and we use only one basis to represent it. Then $D = C$, and the formula becomes

$$A' = CAC^{-1}, \quad \text{or} \quad A = C^{-1}A'C.$$

Digression. What is Linear Algebra? From the modern, logical point of view it is the study of vector spaces and linear transformations. Matrices are introduced as a way to describe and and compute with linear transformations, and especially linear operators. Historically, this was not so. People first studied determinants (which we introduce later), then matrices, under the name "Theory of matrices" and only in 20th century the notions of vector space and linear transformation took their central place.

One reason for this comes from physics. The most basic and fundamental theory about the real world that we have is quantum mechanics. It is phrased in terms of vector spaces and linear operators. When Werner Heisenberg proposed his version of quantum mechanics in 1925, he did not know anything about matrices, which shows that at that time they were not taught to undergraduate students as they are nowadays. Heisenberg had to invent matrix multiplication himself! Perhaps this is the main real reason why linear algebra is taught nowadays to every student: it is the mathematical tool of the most basic physical theory that we currently have. Of course matrices have many other applications in almost every branch of science. And in 20th century Linear Algebra is recognized as a fundamental area of mathematical sciences, even more fundamental than Calculus.