ELEMENTARY NOTIONS OF LINEAR TRANSFORMATIONS

Marco Lenci

May 2006

Definition. A linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that, for all $a, b \in \mathbb{R}$ and $X, Y \in \mathbb{R}^n$,

$$T(aX + bY) = aT(X) + bT(Y). \tag{1}$$

Linear transformations are also called *linear functions*. An equivalent way of characterizing linear transformations is the following.

Proposition 1. A function $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear if, and only if, for every $a \in \mathbb{R}$ and $X, Y \in \mathbb{R}^n$,

$$T(aX) = aT(X) \qquad \text{and} \tag{2}$$

$$T(X + Y) = T(X) + T(Y). \tag{3}$$

Proof. Let us prove the "only if" part, that is, (1) $\Longrightarrow$ (2) & (3). Take $b = 0$ in (1) and (2) is proved. Take $a = b = 1$ in (1) and (3) is proved.

Now for the the "if" part, that is, (2) & (3) $\Longrightarrow$ (1). In (3), plug aX in place of X and bY in place of Y. You will read

$$T(aX + bY) = T(aX) + T(bY). \tag{4}$$

By virtue of (2), $T(aX) = aT(X)$ and, analogously, $T(bY) = bT(Y)$. Using these equalities in the right hand side of (4) will give us (1). Q.E.D.

Exercise 1. Show that, if T is a linear transformation, $T(0, 0, \ldots, 0) = (0, \ldots, 0)$ (the first vector being the origin of $\mathbb{R}^n$ and the second the origin of $\mathbb{R}^m$).

Exercise 2. Show that, if $c \in \mathbb{R}$, the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined as $f(x) = cx$ is linear.

Exercise 3. Show that, if $V \in \mathbb{R}^n$, the scalar field $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by $g(X) := V \cdot X$ is a linear transformation.

Exercise 4. Fix a number $c \in \mathbb{R}$ and define the scalar field $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ via $h(X) := c$. For what values of c, if any, is g a linear tranformation?

Property (1) is called the *linearity property*. It involves the sum of two vectors. There is nothing special about the number two, and the property can be extended to any (finite) sum of vectors.

Proposition 2. For k a positive integer, let $a_1, a_2, \ldots, a_k$ be scalars and X_1, $X_2, \ldots, X_k$ be n-dimensional vectors. If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, then

$$T(a_1X_1 + \cdots + a_kX_k) = a_1T(X_1) + \cdots + a_kT(X_k). \tag{5}$$

Proof. We'll prove (5) by induction on $k \geq 1$. When $k = 1$, the sum is on one vector only, and (5) reduces to (2), which holds true for linear transformations.

For the second step of the induction, suppose that (5) has been proved for *any* sum of k vectors. We want to prove it for any sum of $k+1$ vectors (involving the scalars $a_1, a_2, \ldots, a_{k+1}$ and the vectors X_1, $X_2, \ldots, X_{k+1}$). We have

$$\begin{aligned} &T\left(a_1X_1 + \cdots + a_kX_k + a_{k+1}X_{k+1}\right) = \\ &= T\left(a_1X_1 + \cdots + a_kX_k\right) + a_{k+1}T(X_{k+1}) = \\ &= a_1T(X_1) + \cdots + a_kT(X_k) + a_{k+1}T(X_{k+1}), \end{aligned} \tag{6}$$

where in the first equality we have used (1) with $a = 1$, $X = a_1X_1 + \cdots + a_kX_k$, $b = a_{k+1}$, and $Y = X_{k+1}$; while in the second equality we have used the induction hypothesis on the sum of k vectors. Q.E.D.

Linear transformations are among the easiest functions to deal with, second only to the constant functions. In fact, while for a general function $F : D \longrightarrow \mathbb{R}^m$, one gets complete knowledge of F only when they know the values $F(X)$ for *all* $X \in D$, for a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, it is enough to know the value of T on n inputs, namely, the canonical basis vectors $E_1, E_2, \ldots, E_n$. (Remember that $E_i \in \mathbb{R}^n$ is defined as the vector that contains zeroes in all components, except for the i^{th} component, where it contains 1.)

In other words, if one knows $T(E_1), T(E_2), \ldots, T(E_n)$, they can (more or less easily) calculate $T(X)$, $\forall X \in \mathbb{R}^n$. That is, they know the function T *in its entirety!* The next proposition precisely addresses this point:

Proposition 3. If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, then, $\forall X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,

$$T(X) = T(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_iT(E_i). \tag{7}$$

Proof. Remember that any vector $X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ can be rewritten as

$$X = \sum_{i=1}^{n} x_iE_i. \tag{8}$$

(This is indeed the most notable property of the canonical vectors E_i.) If we evaluate T on both sides of (8), we get, by the extended linearity property (5).

$$T(X) = T\left(\sum_{i=1}^{n} x_iE_i\right) = \sum_{i=1}^{n} x_iT(E_i), \tag{9}$$

which is what we wanted to prove. Q.E.D.

A particularly easy type of linear transformation are the *scalar linear transformations*, that is, linear transformations $t : \mathbb{R}^n \longrightarrow \mathbb{R}$. Each such function can be expressed as a dot product. In fact,

Proposition 4. Given a scalar linear transformation $t : \mathbb{R}^n \longrightarrow \mathbb{R}$, there exists a vector $V \in \mathbb{R}^n$ such that, $\forall X \in \mathbb{R}^n$,

$$t(X) = V \cdot X. \tag{10}$$

Proof. Set $V := (t(E_1), t(E_2), \dots, t(E_n))$. (Each $t(E_i)$ is a number so it makes sense to define V as an n-dimensional vector.) We apply (7) to our linear function t. We obtain

$$t(X) = \sum_{i=1}^{n} x_i t(E_i), \tag{11}$$

which is precisely $V \cdot X$, whence (10). Q.E.D.

We already know (from Exercise 3) that any function $g(X) = V \cdot X$ is linear. From this consideration and Proposition 4 we conclude that the scalar linear transformations are those, and only those, that can be expressed as a scalar (dot) product of the type $V \cdot X$.