MTH 35, SPRING 2017

NIKOS APOSTOLAKIS

1. Linear transformations

Definition 1. A function $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is called a *linear transformation* if, for any scalars $\lambda, \mu \in \mathbb{R}$ and any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ we have:

$$T(\lambda\mathbf{u} + \mu\mathbf{v}) = \lambda T(\mathbf{u}) + \mu T(\mathbf{v})$$

The following properties are consequence of the definition:

Proposition 2. *A linear transformation* $T\colon \mathbb{R}^n \to \mathbf{R}^m$

1. *takes* $\mathbf{0}$ *to* $\mathbf{0}$, *i.e.* $T(\mathbf{0}) = \mathbf{0}$
2. *takes sums to sums, i.e.* $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
3. *takes scalar multiples to scalar multiples, i.e.* $T(\lambda\mathbf{v}) = \lambda T(\mathbf{v})$
4. *takes linear combinations to linear combinations, i.e.*

$$T(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k) = \lambda_1 T(\mathbf{v}_1) + \lambda_2 T(\mathbf{v}_2) + \ldots \lambda_k T(\mathbf{v}_k)$$

Furthermore, if properties (2) and (3) hold for a function $T\colon \mathbb{R}^n \to \mathbb{R}^m$, *then* T *is a linear transformation.*

Example 3. The zero transformation, $O\colon \mathbb{R}^n \to \mathbb{R}^m$ is defined as: $T(\mathbf{x}) = \mathbf{0}$ O is clearly linear.

Example 4. The identity transformation, $I_n\colon \mathbb{R}^n \to \mathbb{R}^n$ is defined as: $I_n(\mathbf{x}) = \mathbf{x}$ I_n is clearly linear.

Example 5. Let $T\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $T(x\mathbf{i} + y\mathbf{j}) = 2x - 3y$. Then T is a linear transformation. Indeed, let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$, and $\lambda, \mu \in \mathbb{R}$. Then we have:

$$\begin{aligned} T(\lambda\mathbf{u} + \mu\mathbf{v}) &= T\left(\left(\lambda u_1 + \mu v_1\right)\mathbf{i} + \left(\lambda u_2 + \mu v_2\right)\mathbf{j}\right) \\ &= 2\left(\lambda u_1 + \mu v_1\right) + 3\left(\lambda u_2 + \mu v_2\right) \\ &= 2\left(\lambda u_1 + \mu v_1\right) + 3\left(\lambda u_2 + \mu v_2\right) \\ &= 2\lambda u_1 + 2\mu v_1 + 3\lambda u_2 + 3\mu v_2 \end{aligned}$$

and

$$\begin{aligned} \lambda T(\mathbf{u}) + \mu T(\mathbf{v}) &= \lambda(2u_1 + 3u_2) + \mu(2v_1 + 3v_2) \\ &= 2\lambda u_1 + 3\lambda u_2 + 2\mu v_1 + 3\mu v_2 \end{aligned}$$

Date: February 23, 2017.

Example 6. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T(x\mathbf{i}+y\mathbf{j}) = -x\mathbf{i}+(x+y)\mathbf{j}+2y\mathbf{k}$. This is also a linear transformation.

Indeed, let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$, and $\lambda, \mu \in \mathbb{R}$. Then we have:

$$\begin{aligned}T(\lambda\mathbf{u} + \mu\mathbf{v}) &= T\left(\left(\lambda u_1 + \mu v_1\right)\mathbf{i} + \left(\lambda u_2 + \mu v_2\right)\mathbf{j}\right)\\ &= -\left(\lambda u_1 + \mu v_1\right)\mathbf{i} + \left(\left(\lambda u_1 + \mu v_1\right) + \left(\lambda u_2 + \mu v_2\right)\right)\mathbf{j} + 2\left(\lambda u_2 + \mu v_2\right)\mathbf{k}\\ &= \left(-\lambda u_1 - \mu v_1\right)\mathbf{i} + \left(\lambda u_1 + \mu v_1 + \lambda u_2 + \mu v_2\right)\mathbf{j} + \left(2\lambda u_2 + 2\mu v_2\right)\mathbf{k}\end{aligned}$$

and

$$\begin{aligned}\lambda T(\mathbf{u}) + \mu T(\mathbf{v}) &= \lambda\left(-u_1\mathbf{i} + (u_1 + u_2)\mathbf{j} + 2u_2\mathbf{k}\right) + \mu\left(-v_1\mathbf{i} + (v_1 + v_2)\mathbf{j} + 2v_2\mathbf{k}\right)\\ &= -\lambda u_1\mathbf{i} + \lambda(u_1 + u_2)\mathbf{j} + 2\lambda u_2\mathbf{k} - \mu v_1\mathbf{i} + \mu(v_1 + v_2)\mathbf{j} + 2\mu v_2\mathbf{k}\\ &= \left(-\lambda u_1 - \mu v_1\right)\mathbf{i} + \left(\lambda u_1 + \lambda u_2 + \mu v_1 + \mu v_2\right)\mathbf{j} + \left(2\lambda u_2 + 2\mu v_2\right)\mathbf{k}\end{aligned}$$

We can get many more examples by using *multiplication of a vector by a matrix*[1]. In previous classes we were writing $A\mathbf{x}$ for the variable part of a linear system. We now make this notation official:

Definition 7. Let A be an $m \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$. Then the product $A\mathbf{x}$ is the a vector in $\mathbb{R}^m$ given by

$$A\mathbf{x} = (\mathbf{r}_1 \cdot \mathbf{x}, \dots, \mathbf{r}_m \cdot \mathbf{x})$$

where $\mathbf{r}_j$, $j = 1, \dots, m$ are the row vectors of A and $\cdot$ is the dot product.

In expanded form, if we write $\mathbf{x}$ and $A\mathbf{x}$ as column vectors we have:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{n2} & \dots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

Proposition 8. *If A is an $m \times n$ matrix, then the function $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ defined by*

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

is a linear transformation.

Example 9. The 2×2 matrix

$$A = \begin{pmatrix} 1 & -2 \\ -3 & 4 \end{pmatrix}$$

gives the linear transformation $T_A\colon \mathbb{R}^2 \to \mathbb{R}^2$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -2 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - 2x_2 \\ -3x_1 + 4x_2 \end{pmatrix}$$

Or by using "horizontal" notation for vectors:

$$T_A\left((x_1, x_2)\right) = (x_1 - 2x_2, -3x_1 + 4x_2)$$

As an **exercise**, do verify that this formula indeed defines a linear transformation.

[1]In the next section we see that this is a special case of multiplication of two matrices.

It turns out that *all* linear transformations come from multiplying by a matrix. We explain this below:

If T is a linear transformation, then if we know the values $T(\mathbf{b}_i)$ for all vectors in a basis of $\mathbb{R}^n$ we can determine the value for any vector $\mathbf{v} \in \mathbb{R}^n$. Indeed, $\mathbf{v}$ is a linear combination of the vectors $\mathbf{b}_i$, say $\mathbf{v} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n$, and by property (4) in Proposition 2 we have that

$$T(\mathbf{v}) = x_1 T(b_1) + \ldots + x_n T(\mathbf{b}_n)$$

In particular, we can use the values of T in the standard basis $\mathbf{e}_i$ of $\mathbb{R}^n$. Each $T(\mathbf{e}_i)$ is a vector in $\mathbb{R}^m$ so it has m coordinates in the standard basis of $\mathbb{R}^m$, and we can organize all these values into and $m \times n$ matrix:

Definition 10. The *matrix* of a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is the matrix $[T]$ that has as columns the values $T(\mathbf{e}_i)$, where $\mathbf{e}_i$, $i = 1, \ldots, n$ is the standard basis of $\mathbf{R}^n$, i.e.

$$[T] = \begin{bmatrix} T(\mathbf{e}_1) & \ldots & T(\mathbf{e})_n \end{bmatrix}$$

Example 11. The matrix of the zero transformation (see Example 3) is the zero matrix: all entries are 0.

Example 12. The matrix of the identity transformation I_n is the so-called $n \times n$ *identity matrix*:

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

Example 13. If T is the linear transformation of Example 5 we have:

$$T(\mathbf{e}_1) = 2, \quad T(\mathbf{e}_2) = -3$$

so we get the 1×2 matrix:

$$[T] = \begin{pmatrix} 2 & -3 \end{pmatrix}$$

Notice that

$$[T]\mathbf{x} = \begin{pmatrix} 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - 3y \end{pmatrix}$$

Example 14. For the linear transformation of Example 6 we have:

$$T(\mathbf{e}_1) = (-1, 1, 0), \quad T(\mathbf{e}_2) = (0, 1, 2)$$

so

$$[T] = \begin{pmatrix} -1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}$$

Notice that

$$[T]\mathbf{v} = \begin{pmatrix} -1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ x + y \\ 2y \end{pmatrix}$$

So, as in the previous Example, the linear transformation we get by multiplying with the matrix $[T]$ is T itself.

Proposition 15. *For any linear transformation T, we have:*

$$T(\mathbf{x}) = [T]\mathbf{x}$$

2. Matrix Multiplication

Definition 16. Let A be an $m \times n$ and B and $n \times k$ matrix. Write A as a column of row vectors and B as a a row of column vectors:

$$A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{pmatrix} \qquad B = \begin{pmatrix} c_1 & c_2 & \cdots & c_k \end{pmatrix}$$

Then we define the product $A \cdot B$ to be the $n \times k$ matrix

$$AB = (\mathbf{r}_i \cdot \mathbf{c}_j), \quad i = 1, \ldots, m, \quad j = 1, \ldots, k$$

i.e. the (i, j) entry of AB is the dot product of the ith row of A with the jth column of B. Notice that this makes sense because each row of A and each column of B are in $\mathbb{R}^n$.

Proposition 17. *The multiplication of matrices has the following properties:*

(1) *It is* associative*: If A, B, and C are matrices then*

$$(AB)C = A(BC)$$

(2) *It* distributes over addition *from the left and from the right:*

$$A(B + C) = AB + AC, \qquad (A + B)C = AC + BC$$

(3) *It is* linear*; if $\lambda \in \mathbb{R}$ then:*

$$(\lambda A)B = \lambda(AB) \qquad A(\lambda B) = \lambda(AB)$$

Since multiplication of matrices is not always defined, the above equations should be understood as saying than whenever the left side is defined, the right side is also defined and is equal to the left side.

***Remark* 18.** In general multiplication of matrices is *not commutative*. If A and B are matrices so that AB and BA are both defined, the two products are not necessarily equal.

Example 19. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

Then we can calculate:

$$AB = \begin{pmatrix} 5 & 11 \\ 11 & 25 \end{pmatrix}, \qquad BA = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}$$

Definition 20. The *zero* $m \times n$ matrix is the matrix whose all entries are 0. We usually denote the zero $m \times n$ matrix by O_{mn} or, when there is no risk of confusion, simply by O.

The *identity* $n \times n$ matrix is the matrix that has every diagonal entry equal to 1, and all other entries equal to 0. We denote the $n \times n$ identity matrix by I_n or, when there is no risk of confusion, simply by I.

Proposition 21. *We have:*

- $OA = O, \qquad AO = O$
- $IA = A, \qquad AI = A$

whenever the products are defined.

Why do we multiply matrices that way? The answer is that matrix multiplication corresponds to *composition* linear transformations.

Proposition 22. *Let* $T\colon \mathbb{R}^n \to \mathbb{R}^m$ *and* $S\colon \mathbf{R}^m \to \mathbb{R}^k$ *then*

(1) *The composition* $S \circ T\colon \mathbb{R}^n \to \mathbb{R}^k$ *is also a linear transformation.*
(2) *The matrix of* $S \circ T$ *equals to the product of the matrix of* S *with the matrix of* T*, that is:*

$$[S \circ T] = [S][T]$$

2.1. **Exercises.**