Linear Maps from $\mathbb{R}^m$ to $\mathbb{R}^n$

S Kumaresan
kumaresa@gmail.com

December 8, 2023

Abstract

We classify (find) all linear maps $T\colon \mathbb{R}^m \to \mathbb{R}^n$. The plan is in the following list of steps.

1. Find all linear maps from $\mathbb{R}$ to $\mathbb{R}$. They are of the from $f(x) = ax$ where $a = f(1)$.
2. Find all linear maps from $\mathbb{R}^n$ to $\mathbb{R}$. They are of the from $f(x) = \sum_k a_k x_k$ where $a_k = f(e_k)$.
3. The projection maps $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$, $1 \le i \le m$, defined by $\pi_(x_1, \dots, x_m) = x_i$ are linear. Note that $\pi_i(x) = x \cdot e_i$.
4. Composition of linear maps is a linear map.
5. Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be linear. Let $T_i := \pi_i \circ T$. We "know" what T_i looks like from Item 2. $T_i x = a_{i1}x_1 + \cdots + a_{im}x_m$, $1 \le i \le n$. Note that $a_{ij} = Te_j \cdot e_i$.
6. Let $Tx = (y_1, \dots, y_n)$. Then $y_i := \pi_i(y) = \pi_i(Tx) = (\pi_i \circ T)(x)$.
7. Matrix representation of $T\colon \mathbb{R}^m \to \mathbb{R}^n$. Let
$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}. \quad \text{Note that } Tx = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$
8. What is Te_i? Geometric meaning of the matrix A of T: Ae_i is the i-th column of A.
9. What is the geometric meaning of the rows? The i-th row constitutes the coordinates of $T_i x$.
10. Matrix representation of a linear map $T\colon \mathbb{R}^m \to \mathbb{R}^n$.

We start with the question: What are all the linear maps from $\mathbb{R}$ to $\mathbb{R}$? Let $f\colon \mathbb{R} \to \mathbb{R}$ be linear. Given $x \in \mathbb{R}$, we write it as $x = x \cdot 1$ where we treat x on the right side as a scalar and 1 as a (basic) vector. That is, we consider $\{1\}$ as a basis for $\mathbb{R}$ and express the vector x as a scalar multiple of the basic vector 1. Since f is linear, $f(x) = f(x \cdot 1) = xf(1) = ax$ where $a := f(1)$. Thus, if $f\colon \mathbb{R} \to \mathbb{R}$ is linear, then f is given by $f(x) = ax$ where $a = f(1)$. Conversely, if $a \in \mathbb{R}$ and if we define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = ax$, then f is linear. We also note that $f(1) = a$. Thus we have arrived at the following lemma.

Lemma 1. *A map $f\colon \mathbb{R} \to \mathbb{R}$ is linear iff $f(x) = ax$ for all $x \in \mathbb{R}$, where $a = f(1)$.* $\square$

Let us now look at a linear map $f\colon \mathbb{R}^m \to \mathbb{R}$. We consider $\mathbb{R}^m$ as the space of column vectors. Let $e_i := (0, \dots, 0, 1, 0, \dots, 0)^t$ be the standard basic vectors, $1 \le i \le m$. If $x = (x_1, \dots, x_m)^t \in \mathbb{R}^m$, we note that $x = x_1e_1 + \cdots + x_me_m$. Since f is linear we see that

$$\begin{aligned} f(x) &= f(x_1e_1 + \cdots + x_me_m) \\ &= f(x_1e_1) + \cdots + f(x_me_m) \\ &= x_1f(e_1) + \cdots + x_mf(e_m). \end{aligned}$$

If we let $a_i := f(e_i)$, we see that $f(x) = a_1x_1 + \cdots + a_mx_m$. In conclusion, if $f\colon \mathbb{R}^m \to \mathbb{R}$ is linear then f is of the from $f(x) = a_1x_1 + \cdots + a_mx_m$ where $a_i = f(e_i)$, $1 \le i \le m$. Conversely, if $f\colon \mathbb{R}^m \to \mathbb{R}$ is defined by $f(x) = a_1x_1 + \cdots + a_mx_m$ where $a_i \in \mathbb{R}$ are arbitrary, then f is linear and we have $a_i = f(e_i)$, $1 \le i \le m$. We have thus arrived at the following result.

Lemma 2. *Any linear map $T\colon \mathbb{R}^m \to \mathbb{R}$ is of the form $T(x) = a_1x_1 + \cdots + a_mx_m$ and we have $a_i = T(e_i)$, $1 \le i \le m$. In particular, T is a first degree polynomial in the variables x_i, $1 \le i \le m$, with zero constant term.* $\square$

Recall the standard dot product on $\mathbb{R}^m$: If $x, y \in \mathbb{R}^m$, then $x \cdot y = x_1y_1 + \cdots + x_my_m$. Using this dot product, we can reformulate the last lemma as follows.

Proposition 3. *Let $T\colon \mathbb{R}^m \to \mathbb{R}$ be a linear map. Then there exists a unique vector $a \in \mathbb{R}^n$ such that for all $x \in \mathbb{R}^m$ we have $T(x) = x \cdot a$. The vector $a = (a_1, \dots, a_m)$ is prescribed by $a_i = T(e_i)$, $1 \le i \le m$.* $\square$

Let us look at the projection map $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$ defined by $\pi_i((x_1, \dots, x_m)^t) = x_i$, $1 \le i \le m$. It is easy to check that it is linear. By Lemma 2, we know that it is of the form $\pi_i(x) = \sum_k a_kx_k$. What are a_k's? Note that $a_k = 0$ if $k \ne i$ and $a_i = 1$. Let us formulate this as a lemma.

Lemma 4. *Let the projection map $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$ be defined by $\pi_i((x_1, \dots, x_m)^t)$. Then π_i is linear and we have $\pi_i(x) = \sum_k a_kx_k$ where $a_k = \begin{cases} 1 & k = i \\ 0 & k \ne i. \end{cases}$ In terms of the dot-product, $\pi_i(x) = x \cdot e_i$, $1 \le i \le m$.* $\square$

We claim that the composition of linear maps is linear.

Lemma 5. *Let U. V and W be vector spaces. Let $A\colon U \to V$ and $B\colon V \to W$ be linear maps. Then composition $B \circ A\colon U \to W$ is a linear map.*

Proof. This is an easy verification and is left to the reader. $\square$

Now, let a linear map $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be given. Do you know how to generate n-linear maps $T_i\colon \mathbb{R}^m \to \mathbb{R}$? Lemmas 4-5 give us a way of generating such maps. How about $T_i := \pi \circ T\colon \mathbb{R}^m \to \mathbb{R}$? We know how to express these maps $T_i\colon \mathbb{R}^m \to \mathbb{R}$. They are given by

$$T_i x := T_i(x) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{im}x_m, \quad \text{where } a_{ij} \in \mathbb{R}. \tag{1}$$

But we wanted to know Tx! Observe that if we write $Tx = y$, then to know y is the same as knowing its coordinates y_i, $1 \leq i \leq n$. That is easily achieved. Note that $y_i = \pi_i(y) = \pi_i(Tx) = (\pi_i \circ T)(x)$. Hence

$$\begin{aligned}
y = Tx &= ((\pi_1 \circ T)(x), (\pi_2 \circ T)(x), \ldots, (\pi_n \circ T)(x))^t \\
&= (T_1(x), T_2(x), \ldots, T_n(x))^t \\
&= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + a_{n2}x_2 + \cdots, a_{nm}x_m)^t && (2) \\
&= \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = Ax, \text{ say.} && (3)
\end{aligned}$$

Observation 6. If $T\colon \mathbb{R}^m \to \mathbb{R}^n$ is a linear map, then we found the following:

(i) For $x \in \mathbb{R}^m$, the vector $Tx = (p_1(x), \ldots, p_n(x))^t$ where each p_i is a first degree polynomial with zero constant term.

(ii) Each p_i is given by $p_i(x) = a_{i1}x_1 + \cdots + a_{im}x_m$ where $a_{ij} = T_i(e_j) = \pi_i(Te_j) = (Te_j) \cdot e_i$.

(iii) There is an $(n \times m)$-matrix A such that $Tx = Ax$ where x is considered as a matrix of size $m \times 1$.

Conversely, if A is an $(n \times m)$-matrix and if we define $Tx := Ax$ for $x \in \mathbb{R}^m$, then T is a linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$. This follows from the properties of the matrix multiplication. Thus we have proved the following theorem.

Theorem 7. *A map $T\colon \mathbb{R}^m \to \mathbb{R}^n$ is linear iff there exists $(n \times m)$-matrix A such that $Tx = Ax$.* $\square$

There are a few more observations that bring out the 'geometric perspective' to a matrix. We already saw the geometric meaning of the rows. The entries of the i-th row are the coefficients of the $T_i x \in \mathbb{R}^n$ expressed as a linear combination of the standard basic vectors of $\mathbb{R}^n$. The columns also admit a geometric meaning. Look at $Te_i = Ae_i$. This is nothing other than the i-th column of A.

The matrix A is unique: If $Tx = Ax = Bx$, then $(A - B)(x) = 0$ for all $x \in \mathbb{R}^m$. In particular, if we take $x = e_i$, then the i-th column of the matrix $A - B$ is the zero vector in $\mathbb{R}^n$, $1 \leq i \leq n$. Thus each column of $A - B$ is the zero vector in $\mathbb{R}^n$ and hence $A - B$ is the zero matrix. Hence, $A = B$.

The matrix A in (3) is known as the matrix of the linear map T. Why is the last paragraph relevant here?

An observation: Note that the proofs of Lemmas 1-2 indicate the following result.

Theorem 8. *Let V be a finite dimensional vector space and let W be a vector space. Let $T\colon V \to W$ be linear. Let $\{v_1, \ldots, v_n\}$ be a basis of V. If we "know" Tv_i, say, $Tv_i = w_i$, $1 \leq i \leq n$, then we "know" Tv for any $v \in V$. If $v = \sum_i a_i v_i$, then $Tv = \sum_i a_i w_i$.* □

There is a 'converse' to the last result.

Theorem 9. *Let V be a finite dimensional vector space and let W be a vector space. Fix a basis $\{v_k : 1 \leq k \leq n\}$ of V. Let $w_1, w_2, \ldots, w_n$ be a finite sequence of vectors in W. (Note that w_j's need not be distinct.) Define $T\colon V \to W$ by setting*

$$Tv = a_1 w_1 + a_2 w_2 + \cdots + a_n w_n, \quad \text{where } v = \sum_{k=1}^{n} a_k v_k. \tag{4}$$

Then T is a linear map from V to W.

Proof. If $u, v \in V$ and if $u = \sum_k a_k v_k$ and $v = \sum_k b_k v_k$, note that $v + w = \sum_k (a_k + b_k) v_k$ and for $\lambda \in \mathbb{R}$, $\lambda v = \sum_k (\lambda a_k) v_k$. Hence we have

$$\begin{aligned} T(u+v) &= \sum_k (a_k + b_k) w_k = \sum_k a_k w_k + \sum_k b_k w_k = Tu + Tv \\ T(\lambda v) &= \sum_k (\lambda a_k) w_k = \lambda (\sum_k a_k w_k) = \lambda T v. \end{aligned}$$

Hence T is linear. □

What is significant in the last theorem is that we can "prescribe" arbitrary vectors (from W) to Tv_k and "extend T linearly" by (4) to any vector v. Theorems 8 and 9 substantiate the following important take-way: when dealing with a linear map $T\colon V \to W$, we should not be concerned with the explicit expressions for T in terms of "coordinates", but we should focus on its action on a (*conveniently chosen*) basis.

To appreciate this, carry out the following exercise: Construct a linear map from $\mathbb{R}^2$ to $\mathbb{R}^3$ and another from $\mathbb{R}^3$ to $\mathbb{R}^2$ and express them in terms of coordinates! Use (u,v) as coordinates for $\mathbb{R}^2$ and (x,y,z) as coordinates for $\mathbb{R}^3$.

Observation 6-(i) suggests we may define $A\colon \mathbb{R}^3 \to \mathbb{R}^2$ and $B\colon \mathbb{R}^2 \to \mathbb{R}^3$ as follows:

$$A(x,y,z) = (p_1(x,y,z), p_2(x,y,z)) = (3x - 4y + 5z, -x + 2y + z), \text{ say.} \tag{5}$$

$$B(u,v) = (q_1(u,v), q_2(u,v), q_3(u,v)) = (3u + 2v, u - v, u + v), \text{ say.} \tag{6}$$

Theorems 8 and 9 suggest to define $A\colon \mathbb{R}^3 \to \mathbb{R}^2$, we choose arbitrary vectors $w_1, w_2, w_3 \in \mathbb{R}^2$ and set $Ae_i = w_i$, $1 \leq i \leq 3$. For instance, take $w_1 = (3,-1), w_2 = (-4,2), w_3 = (5,1)$. Then

$$A(x,y,z) = xAe_1 + yAe_2 + zAe_3 = xw_1 + yw_2 + zw_3 = (3x - 4y + 5z, -x + 2y + z).$$

Voilà! We got the expression for A in (5). Can you carry out a similar exercise for B? That is, can you define Be_1, $Be_2 \in \mathbb{R}^3$ suitably so that you get the expression for B in (6)? Here $e_1 = (1,0)$ and $e_2 = (0,1)$ are the standard basis of $\mathbb{R}^2$.

Remark 10. If I remember correctly, this article was written after recording the following video. Though it is not an exact transcription, this article captures the spirit of the lecture! Linear Maps RmtoRn: `https://youtu.be/njIqG4-aQr4`