Linear Maps and Matrices

S Kumaresan
GITAM
Visakhapatnam
kumaresa.zoom@gmail.com

December 23, 2023

Abstract

In this article, we start with setting up a system of coordinates associated with an ordered basis and then talk about matrix representation of linear maps. We shall give some typical examples as well as some unusual examples. We prove all the standard results of this theme.

1 Ordered Basis and Coordinate Systems

Let V be a finite dimensional (non-zero) vector space over a field $\mathbb{F}$ (which, for concreteness, we may take it to be $\mathbb{R}$). Let $\dim V = n$. Let B be a basis of V. Let $\sigma\colon \{1, 2, \dots, n\} \to B$ a bijection. Then for any $v \in B$, there exists $j \in \{1, \dots, n\}$ such that $v = \sigma(k)$. We then denote v by v_k so that $B = \{v_1, \dots, v_n\}$. We then say B is an *ordered basis* of V.

Let us explain this with a concrete example. Let $V = \mathbb{R}^2$ and consider $B_1 = \{e_1, e_2\}$ and $B_2 = \{e_2, e_1\}$. Then, both B_1 and B_2 are bases of V. As sets, there are the same! But the order in which e_1 and e_2 are listed matters now. e_2 is the "first" basic vector in the ordered basis B_2. You may think we are unnecessarily making a fuss. Let us look at another example. Let us consider the vectors $(3, 2)$ and $(2, 3)$. Then these two vectors form a basis of $\mathbb{R}^2$. Is there any natural order so we may call one of them the first and the other the second? We need to make a choice. One may consider $B_1 = \{(2, 3), (3, 2)\}$ and another $B_2 = \{(3, 2), (2, 3)\}$. As ordered bases they are different.

If you still wonder why it is necessary, assume that you are given instruction as (3,2) and it means you take 3 steps in one direction and 2 steps in the opposite direction, then you will find what you want. Will this make sense? You need to be told whether the first direction is towards right or left. Think along these lines and you will appreciate the reason for precision.

Come back to linear algebra. Let $B = \{v_1, \dots, v_n\}$ be an ordered basis of V. We then associate to each $v \in V$, a system of n-coordinates as follows: If $v = x_1 v_1 + \dots, +x_n v_n$, where $x_i \in \mathbb{R}$

$$\varphi_B : v \mapsto [v]_B := (x_1, \ldots, x_n)^T \in \mathbb{R}^n_c. \tag{1}$$

($\mathbb{R}^n_c$ stands for the n-dimensional vector space of column vectors over $\mathbb{R}$.) The map $\varphi_B : V \to \mathbb{R}^n_c$ is called the *coordinate map* associated with the *ordered basis* B. The x_j is called the j-th coordinate of v (relative to the ordered basis B). Of course, $(x_1, \ldots, x_n)^T$ is called the system of coordinates for the vector v relative to the ordered basis B.

Consider $B' := \{w_1, w_2, \ldots, w_n\}$ where $w_1 = v_2$, $w_2 = v_1$ and $w_j = v_j$ for $3 \leq j \leq n$. Then B' is an ordered basis of V. Let $\varphi_{B'}$ be the associated coordinate map and let us write $v = \sum_i y_i w_i$ so that $\varphi_{B'}(v) = (y_1, \ldots, y_n)$. Then note that $y_1 = x_2$, $y_2 = x_1$ and $y_j = x_j$ for $j > 2$. Thus the set of coordinates of v depend on the ordered basis.

The most important fact about the map $\varphi_B : V \to \mathbb{R}^n_c$ is that it is a linear isomorphism. One should try to prove this on one's own.

> If $v = \sum_i x_i v_i$ and $w = \sum_i y_i v_i$, then $x + y = \sum_i (x_i + y_i) v_i$ and hence $\varphi_B(x + y) = (x_1 + y_1, \ldots, x_n + y_n)^T = (x_1, \ldots, x_n)^T + (y_1, \ldots, y_n)^T = \varphi_B(v) + \varphi_B(w)$. Similarly, one can show that $\varphi_B(tv) = t(\varphi_B(v))$ for $t \in \mathbb{R}$ and $v \in V$. Hence, φ_B is linear.

It is one-one. (Why?) It is onto since if $(a_1, \ldots, a_n)^T \in \mathbb{R}^n_c$, then $v := a_1 v_1 + \ldots + a_n v_n \in V$ is such that $\varphi_B(v) = (a_1, \ldots, a_n)^T$. Hence we conclude that the coordinate map φ_B is a linear isomorphism of V onto $\mathbb{R}^n_c$.

Let $B' := \{w_1, \ldots, w_n\}$ be another ordered basis of V. Let $(y_1, \ldots, y_n)$ be the system of coordinates associated with B'. That is, if $\varphi_{B'}(v) = (y_1, \ldots, y_n)^T$, then $v = \sum_i y_i w_i$. The question is: Is there a way of finding the y coordinates of v if we know the x-coordinates of v?

Let $w_j := \sum_i a_{ij} v_i$. Thus, $[w_j]_B = (a_{1j}, a_{2j}, \ldots, a_{nj})^T$. Let $v \in V$ be such that $v = \sum_{j=1}^n y_j w_j$. Then

$$v = \sum_j y_j w_j = \sum_j y_j \left(\sum_i a_{ij} v_i \right) = \sum_i \left(\sum_j a_{ij} y_j \right) v_i = \sum_i x_i v_i. \tag{2}$$

> I find that the beginners often accept the interchange of the sums in (2) without proper verification. Let us verify it.
>
> $$\begin{aligned} y_1 w_1 &= y_1 a_{11} v_1 + y_1 a_{21} v_2 + \cdots + y_1 a_{n1} v_n \\ y_2 w_2 &= y_2 a_{12} v_1 + y_2 a_{22} v_2 + \cdots + y_2 a_{n2} v_n \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ y_n w_n &= y_n a_{1n} v_1 + y_n a_{2n} v_2 + \cdots + y_n a_{nn} v_n. \end{aligned}$$
>
> What is the coefficients of v_1 on the right side if we add the terms "vertically column-wise"? We find that the coefficient of v_1 in $y_1 w_1 + y_2 w_2 + \cdots + y_n w_n$ is $a_{11} y_1 + a_{12} y_2 + \cdots + a_{1n} y_n = \sum_j a_{1j} y_j$. Similarly, the coefficient of v_i is $\sum_j a_{ij} y_j$. Hence $y_1 w_1 + \cdots + y_n w_n = (\sum_j a_{1j} y_j) v_1 + (\sum_j a_{2j} y_j) v_2 + \cdots + (\sum_j a_{nj} y_j) v_n = \sum_i \left(\sum_j a_{ij} y_j \right) v_i$.

Needless to say the various properties of the vector space such as the associativity, commutativity of the addition and others are used in grouping terms.

Hence from the last equality in (2) we find that $x_i = \sum_j a_{ij} y_j$. (Why?) Let us understand this better. Let A be the matrix whose j-th column is $(a_{1j}, \dots, a_{nj})^T$. Then we find that we have arrived at the following.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}. \tag{3}$$

The matrix A is known as the *transition matrix* of going from the basis B' to B. Perhaps, if the context demands, we may denote it by $A_{(B'B)}$.

Note that the entries of the j-th column of $A_{(B'B)}$ are the coordinates of j-th basic vector w_j relative to the basis B, that is, the j-the column is the column vector $\varphi_B(w_j)$.

Why from B' to B? That is because given the y-coordinates associated with B', we obtain the x-coordinates associated with B.

Of course, we can reverse the process. Starting with $v_j := \sum_i b_{ij} w_i$ we arrive at the transition matrix $A_{(B,B')}$.

Note that we can write (3) as $[v]_B = A[v]_{B'}$ and its counterpart $[v]_{B'} = A_{(B,B')}[v]_B$. In less intimidating notation, we observe that

$$x = A_{(B',B)} y \quad \text{and} \quad y = A_{(B,B')} x. \tag{4}$$

Make sure you understand this, as we shall have occasions to employ this notation.

Let us verify (4) in some easy examples.

Example 1. Let $V = \mathbb{R}^2_c$. Let $B := \{v_1 = e_1, v_2 := e_2\}$ and $B' := \{w_1 := e_1 + e_2, w_2 := e_1 - e_2\}$ be two ordered bases of V. We then have

$$A_{(B',B)} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \text{ and } A_{(B,B')} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Then a vector $v = (x, y)^T \in \mathbb{R}^2_c$ can be expressed as

$$v = xv_1 + yv_2 = \frac{x+y}{2} w_1 + \frac{x-y}{2} w_2.$$

That is,

$$[v]_B = (x, y)^T \text{ and } [v]_{B'} = \left(\frac{x+y}{2}, \frac{x-y}{2}\right).$$

One can easily verify (4) now.

Example 2. Let $V = P_n$ be the vector space of (real) polynomials of degree at most n., $n \in \mathbb{N}$. Let $B := \{1, x, x^2, \cdots, x^n\}$ be the standard ordered basis of V. Let $a \in \mathbb{R}$ be nonzero. Then $B' := \{1, (x-a), (x-a)^2, \ldots, (x-a)^n\}$ is also an ordered basis of V. (This will follow from the binomial expansion below.) Using the binomial theorem we have

$$x^k = ((x-a)+a)^k = \sum_{r=0}^{k} \binom{k}{r} (x-a)^r a^{k-r}. \tag{5}$$

Thus the vectors in B are written in terms of B'. One can do it the other way around also.

For simplicity, let us look at $n = 2$. We then have

$$[1]_{B'} = (1, 0, 0)^T, \quad [x]_{B'} = (a, 1, 0)^T \text{ and } [x^2]_{B'} = (a^2, 2a, 1)^T.$$

Similarly, we find that

$$[1]_B = (1, 0, 0)^T, \quad , [x-a]_B = (-a, 1, 0)^T \text{ and } [(x-a)^2]_B = (a^2, -2a, 1)^T.$$

We obtain

$$A_{(B',B')} = \begin{pmatrix} 1 & a & a^2 \\ 0 & 1 & 2a \\ 0 & 0 & 1 \end{pmatrix} \text{ and } A_{(B',B')} = \begin{pmatrix} 1 & -a & -a^2 \\ 0 & 1 & -2a \\ 0 & 0 & 1 \end{pmatrix}.$$

We invite the reader to verify (4).

Now comes the interesting observation. Let $v \in V$ be arbitrary.

$$[v]_B = A_{(B',B)}[v]_{B'} = A_{(B',B)} \left(A_{(B,B')}[v]_B\right) = \left(A_{(B',B)}A_{(B,B')}\right)[v]_B.$$

(We are dealing with the product of matrices.) In other words for any $(x_1, \ldots, x_n)^T \in \mathbb{R}^n_c$ we obtain

$$x = A_{(B',B)}A_{(B,B')}x.$$

It follows that $A_{(B',B)}A_{(B,B')} = I$, the identity $n \times n$-matrix.

> How does it follow? Hint: Take $x = e_i$, the i-th standard basic vector in the above equation.

So, we have found that the transition matrices are invertible!

Example 3. One easily verifies that the product of transition matrices in Examples 1–2 are inverses of each other.

2 Matrix Representation of a Linear Map

Let V and W be finite dimensional vector spaces. In this section, our goal is to associate a matrix to any linear map $f\colon V \to W$. Let $B := \{v_1, \ldots, v_m\}$ and $B := \{w_1, \ldots, w_n\}$ be ordered bases of V and W respectively. We then write

$$f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{nj}w_n, \qquad 1 \leq j \leq m.$$

We now construct a matrix whose j-th column is $(a_{1j}, a_{2j}, \ldots, a_{nj})^T$. Note that we have

$$(a_{1j}, a_{2j}, \ldots, a_{nj})^T = [f(v_j)]_{B'}.$$

The matrix is

$$([f(v_1)]_{B'}, [f(v_2)]_{B'}, \ldots, [f(v_n)]_{B'}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n1} & \ldots & a_{nm} \end{pmatrix}.$$

This matrix is called the matrix associated with the linear map f relative to the ordered bases B and B'. It is denoted by $[f]_{(B,B')}$. Note that the matrix lies in $M_{n\times m}(\mathbb{R})$.

For emphasis sake, let us observe the following fact. If $f\colon V \to W$ is linear then the matrix of f relative to any choice of ordered bases will have $\dim W$-rows and $\dim V$ columns. (Can you see why? Do you visualize it?)

Let us look at some simple examples.

Example 4. Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$ Let $f\colon V \to W$ be defined by $f((x,y)^T) = (x+y, x-y, 2x+3y)^T$. (Why is it linear?) Let $B := \{e_1, e_2\}$ be th the standard ordered basis of V and let $B' = \{w_1, w_2, w_3\}$ be the standard ordered basis of $\mathbb{R}^3$. (Note that $w_1 = (1,0,0)^T$ etc.) What is the matrix $[f]_{(B,B')}$? We need to find $f(e_1)$ and $f(e_2)$. We have $f(e_1) = (1,1,2)^T$ and $f(e_2) = (1,-1,3)^T$. Hence the matrix is $[f]_{(B,B')=} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 3 \end{pmatrix}$.

In place of B, if we take $B_1 := \{e_2, e_1\}$, the matrix $[f]_{(B_1,B')} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 3 & 2 \end{pmatrix}$. That is, the columns of $[f]_{(B,B')}$ are interchanged.

In place of B', let us consider $B_2 = \{u_1 = w_2, u_2 = w_1, u_3 = w_3\}$. Then $f(e_1) = (1,1,2) = 1\cdot w_1 + 1\cdot w_2 + 2w_3 = 1\cdot u_1 + 1\cdot u_2 + 2\cdot u_3$. And, $f(e_2) = (1,-1,3) = 1\cdot w_1 - 1\cdot w_2 + 3\cdot$

$w_3 = -1 \cdot u_1 + 1 \cdot u_2 + 3 \cdot u_3$. Hence the matrix $[f]_{(B,B_2)} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 2 & 3 \end{pmatrix}$. That is, the first and the second rows of $[f]_{(B,B')}$ are interchanged.

These show the importance of the order in the basis.

Example 5. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be linear. Such maps are called *linear functionals.* Let B be the standard ordered basis of $\mathbb{R}^n_c$ and $B' := \{1\}$ the basis for $\mathbb{R}$. Then the matrix of f relative to these bases is $(f(e_1), \dots, f(e_n))$. Note that it is a matrix of type $1 \times n$ and hence may be considered as an element of $\mathbb{R}^n_r$.

Example 6. Let $V = \mathbb{C}$ considered as a real vector space. Let $B := \{1, i\}$ be the standard ordered basis of V. Let $\alpha := a + ib \in \mathbb{C}$. Consider the linear map $f\colon \mathbb{C} \to \mathbb{C}$ defined by $f(z) := \alpha z$. We consider the same ordered basis for both the domain and the co-domain. Then $f(1) = a \cdot 1 + b \cdot i$ and $f(i) = (a + ib)i = -b \cdot 1 + a \cdot i$. Hence the matrix of f is $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Ex. 7. What is the matrix of the conjugation $z \mapsto \overline{z}$ from $\mathbb{C}$ to itself with respect to the above ordered basis both in the domain and in the co-domain?

Example 8. Let $V = P_n$ be the vector space of (real) polynomials of degree at most $n \geq 2$. Let $W := P_{n-1}$. Consider the derivative map $D\colon V \to W$ defined by $Dp(x) = p'(X)$. That is, if $p(X) = c_1 + c_1 X + \cdots + c_n X^n$, then $Dp(X) = p'(X) = c_1 + 2c_2 X + 3c_3 X^2 + \cdots + nc_n X^{n-1}$. We take the standard ordered bases $B := \{1, X, \dots, X^n\}$ and $B' := \{1, X, \dots, X^{n-1}\}$. Then the matrix
$$[D]_{(B,B')} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 0 & 3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & n \end{pmatrix}_{n \times (n+1)}.$$

Ex. 9. This is almost the same example as the last but with a twist. Let $W = V = P_n$ and $B' = B$. Then the matrix of D is the $(n+1) \times (n+1)$ -matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 0 & 3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & n \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Ex. 10. Let $V = \mathbb{R}^m$ and $W = \mathbb{R}^n$. Let $B := \{v_i : 1 \leq i \leq m\}$ and $B' = \{w_j : 1 \leq j \leq n\}$ be the standard ordered bases of V and W respectively. (Do you understand what v_1 and w_n

are?) Let $T_A\colon \mathbb{R}^m_c \to \mathbb{R}^n_c$ be defined as usual: $T_A x := Ax$. Can you guess the matrix $[T_A]_{(B,B')}$? It is A. We invite you to convince yourself.

Example 11. Let us look at an interesting example. Let $V = W$ be a finite dimensional vector space. Let $B := \{v_1, \ldots, v_n\} = B'$ be an ordered basis of V and W. Let $f\colon V \to V$ be the identity map. Then what is the matrix $[f]_{(B,B)}$? It is the identity matrix $I_{n\times n}$. Adapt the argument of a general case in the next paragraph to prove this.

Suppose now, we decide to use $B' := \{w_1, \ldots, w_n\}$ as the ordered basis of $W = V$, the co-domain while keeping B for the domain V. Then what is $[I]_{(B,B')}$? It is the transition matrix $A_{(B'B)}$. For, $I(v_i) = \sum_j a_{ji} w_j$ and hence the i-th column in $[I]_{(B,B')}$ is $(a_{1i}, a_{2i}, \ldots, a_{ni})^T$, which is also the i-th column of the transition matrix $A_{(B',B)}$. To put in other words, we simply observe that $[v_i]_{B'} = (a_{1i}, a_{2i}, \ldots, a_{ni})^T$.

Let us record this for future reference:

$$[I]_{(B,B')} = A_{(B',B)}. \tag{6}$$

Example 12. Let $f\colon V \to W$ be a linear isomorphism. Then $\dim V = \dim W$. Let $B := \{v_i : 1 \le i \le n\}$ be an ordered basis of V. Let $w_i := f(v_i)$, $1 \le i \le n$. Since f is one-one, we know that $B' := \{w_i : 1 \le i \le n\}$ is a linearly independent set in W. (Can you quickly go through a proof?) Since it has the same number of elements as the dimension of W, B' is a basis. (Can you give a direct proof of the fact that B' is a spanning set?) What is the matrix $[f]_{(B,B')}$?

> It is the identity matrix $I_{n\times n}$. For, $f(v_i) = \sum_j c_j w_j$ where $c_j = 1$ if $j = i$ and $c_j = 0$ if $j \ne i$. Hence the i-th column o the matrix is the standard basic vector $e_i \in \mathbb{R}^n_c$.

Ex. 13. Let $f\colon V \to W$ be 1-1 linear map. Then we know that $m := \dim V \le \dim W =: n$. (Why?) Let $B := \{v_i : 1 \le i \le m\}$ be an ordered basis of V. Extend the linearly independent set $\{f(v_i) : 1 \le i \le m\}$ to an ordered basis B' of W. (We assume that $f(v_i)$'s retain their order.) Show that the matrix $[f]_{(B,B')}$ is $\begin{pmatrix} I_{m\times m} \\ --- \\ 0_{n-m\times n} \end{pmatrix}$.

Ex. 14. Let $f\colon V \to W$ be linear and onto. It follows from the rank-nullity theorem that $\dim W \le \dim V$ Let $B' := \{w_j : 1 \le j \le n\}$ be given. Let $v_j \in V$ be any vector such that $f(v_j) = w_j$, $1 \le j \le n$. Let $\{u_1, \ldots, u_k\}$ be a basis of $\ker f$ where $m = n + k$. From the proof of the rank-nullity theorem (or by a direct verification), we see that $\{v_1, \ldots, v_n, v_{n+1} := u_1, \ldots, v_{n+k} := u_k\}$ is a basis of V. We consider the order in which v_i $1 \le i \le m$ appears in B. Then $f(v_i) = w_i$ for $1 \le i \le n$ and $f(v_i) = 0$ for $i > n$. Then show that the matrix $[f]$ is of the form $(I_{n\times n} | 0_{k\times n})$.

Example 12, Ex. 13 and Ex. 14 teach us an important trick. If we want to have a nice-looking or a simpler matrix to represent the linear map, we need to choose the ordered bases in a smart way.

3 Theoretical Results on Matrix Representations

We now return to some theoretical aspect of the matrix representation of a linear map.

If the context is clear, it would be easier for the eye and simpler for me to type to let $[v]$ stand for $[v]_B$, $[w]$ for $[w]_{B'}$ and $[f]$ for $[f]_{(B,B')}$.

One more fact to keep in mind is the following.

Observation 15. Keep the notation above. Let $A := [f]_{(B,B')} = (a_{ij})$ be the matrix. Then we have

$$[f(v)] = A[v], \qquad \text{for any } v \in C. \tag{7}$$

It is the unique matrix in $M_{n\times m}(\mathbb{R})$ with this property.

Proof. Let $v := x_1v_1 + \cdots + x_mv_m$. Let $f(v) = y_1w_1 + \cdots + y_nw_n$. What we are required to show is that $(y_1, \ldots, y_n)^T = A(x_1, \ldots, x_m)^T$.

We have $f(v) := x_1f(v_1) + \cdots + x_mf(v_m)$. Hence $[f(v)] = x_1[f(v_1)] + \cdots + x_m[f(v_m)]$.

Let $f(v) = y_1w_1 + \cdots + y_nw_n$. Let $f(v_i) = a_{1i}w_1 + \cdots + a_{ni}w_n$, $1 \leq i \leq m$. Then $f(v) = \sum_i x_if(v_i)$ can be expanded as follows:

$$\begin{aligned}
f(v) &= x_1f(v_1) + x_2f(v_2) + \cdots + x_mf(v_m) \\
&= x_1(a_{11}w_1 + a_{21}w_2 + \cdots + a_{n1}w_n) \\
&\quad + x_2(a_{12}w_1 + a_{22}w_2 + \cdots + a_{n2}w_n) \\
&\qquad\qquad \vdots \\
&\quad + x_m(a_{1m}w_1 + a_{2m}w_2 + \cdots + a_{nm}w_n) \\
&= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m)w_1 \\
&\quad + (a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m)w_2 \\
&\qquad\qquad \vdots \\
&\quad + (a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m)w_n.
\end{aligned}$$

Due to the uniqueness of expression in terms of basis, we see that $y_i = a_{i1}x_1 + a_{i1}x_2 + \cdots + a_{im}x_m$, $1 \leq i \leq n$. Thus, $(y_1, \ldots, y_n)^T = A(x_1, \ldots, x_m)^T$.

The uniqueness is easy. If $B \in M_{n\times m}(\mathbb{R})$ is another such matrix, then the i-th columns of A and B are same, since $[f(v_i)] = A[v_i]$ and $[f(v_i)] = B[v_i]$, $1 \leq i \leq m$. □

We have a map $\varphi\colon L(V,W) \to M_{n\times m}(\mathbb{R})$ defined as $\varphi(f) := [f]_{(B,B')}$. What kind of map is this? Why do we ask this question? Since both the domain and co-domain are vector spaces, we would like to know whether it is a linear map. Since they have the same dimension, we are curious to know whether the map is a linear isomorphism.

Theorem 16. *The map $\varphi\colon L(V,W) \to M_{n\times m}(\mathbb{R})$ defined as $\varphi(f) := [f]_{(B,B')}$ is a linear isomorphism.*

Proof. Observation 15 will make work easy.

Let $f, g \in L(V, W)$. We show that $\varphi(f+g) = \varphi(f) + \varphi(g)$, that is, the matrix of $f+g$ namely $[f+g]$ is the sum of the matrices $[f]$ and $[g]$. The j-th column of $[f+g]$ is $[(f+g)(v_j)]$. Thanks to the way $f+g$ is defined, $(f+g)(v_j) = f(v_j) + g(v_j)$. Hence we see that $[(f+g)(v_j)] = [f(v_j) + g(v_j)] = [f(v_j)] + [g(v_j)]$ (Why? Recall that the coordinate map is a linear map!). Therefore, we conclude that the j-th columns of $[f+g]$ and $[f]+[g]$ are the same, that is, $[f+g] = [f]+[g]$.

In a similar way we can show that $[tf] = t[f]$. The reader is asked to prove this.

The map is one-one. Let $f, g \in L(V, W)$ be distinct. Then there exists $v_j \in B$ such that $f(v_j) \neq g(v_j)$. (Why? Recall that linear maps are completely determined by their action on a basis.) Since the coordinate map is a linear isomorphism, in particular, it is one-one, $[f(v_j)] \neq [g(v_j)]$. But these are the j-th columns of $f]$ and $[g]$ respectively. Hence $[f] \neq [g]$.

The map is onto. Let $A \in M_{n\times m}(\mathbb{R})$. Define $f \colon V \to W$ by setting $f(v_i) := \sum_{j=1}^{n} a_{ji} w_j$. We extend this linearly to an $f \in L(V, W)$. We see that $[f] = A$ by the very construction. $\square$

Remark 17. There is a better way of looking at what went behind the scene in the last result. A beginner may skip this remark and return to it later when he is confident of the coordinate maps etc.

Recall that if V is a vector space, W is any set and $f \colon V \to W$ is a bijection, we have made W into a vector space using f and we also have seen that $f \colon V \to W$ is a linear isomorphism. We wish to push this kind of construction a bit further.

Let V, W, V', W' be vector spaces. Let $f \colon V \to W$ be a linear map. Let $\varphi \colon V \to V'$ be a linear isomorphism and $\psi \colon W \to W'$ be a linear isomorphism. Is there a natural way to define a linear map $g \colon V' \to W'$?

Inspired by the construction of the first paragraph, we look at the following map:

$$g \colon V' \ni v' \mapsto \varphi^{-1}(v') \mapsto f(\varphi^{-1}v') \mapsto \psi(f(\varphi^{-1}v')) \in W'.$$

(Draw a picture to keep track of the movements.) The map $g \colon V' \to W'$ is linear, being the composition of linear maps. We can easily see that the map $f \mapsto g$ is also a bijection. (Do you see why?)

Let $V' = \mathbb{R}_c^m$ and the $\varphi \colon V \to V'$ be the coordinate map $v \mapsto [v]$. Similarly, let $\psi \colon W \to W' := R_c^n$ be the coordinate map. Note that φ and ψ are linear isomorphisms. Hence we obtain a map $\psi \circ f \circ \varphi^{-1} \colon \mathbb{R}_c^m \to \mathbb{R}_c^n$. The matrix of this map with standard ordered bases of $\mathbb{R}_c^m$ and $\mathbb{R}_c^n$ is the matrix $[f]$.

Remark 18. Since we know $\dim M_{n\times m}(\mathbb{R}) = mn$, the last theorem gives another proof of the fact that $\dim L(V, W) = \dim V \times \dim W$.

Example 19. Let $E_{ij} \in M_{n\times m}(\mathbb{R})$ stand for the matrix whose (ij)-th entry is 1 and others are zero. Then $\{E_{ij} : 1 \leq i \leq n, 1 \leq j \leq m\}$ is a basis for $M_{n\times m}(\mathbb{R})$. What are the corresponding linear maps in $L(V,W)$?

If you guessed it to be f_{ij}, you are partially correct. Note that $f_{ij} \in L(V,W)$ is defined by $f_{ij}(v_r) = 0$ if $r \neq i$ and $f_{ij}(v_i) = w_j$. Hence the i-th column of $[f_{ij}]$ the standard j-th basic vector of $\mathbb{R}^n_c$ and all other columns are zero. Thus the (j,i)-th entry of $[f_{ij}]$ is 1 and all other entries are zero. Hence $[f_{ij}] = E_{ji}$.

Let U, V and W be finite dimensional vector spaces. Let $f\colon U \to V$ and $g\colon V \to W$ be linear maps. Let $B_1 := \{u_i : 1 \leq i \leq m\}$ (respectively, $B_2 := \{v_j : 1 \leq j \leq n\}$, $B_3 := \{w_k : 1 \leq k \leq p\}$) be an ordered basis of U (ordered basis of V, W respectively). Let $A = [f]_{(B_1,B_2)}$ and $B := [g]_{(B_2,B_3)}$. We would like to find the matrix of $g \circ f$ relative to the ordered bases (B_1, B_3).

The answer follows from an easy computation.

$$\begin{aligned}
(g \circ f)(u_i) = g(f(u_i)) &= g(\sum_{j=1}^{n} a_{ji} v_j) \\
&= \sum_{j=1}^{n} a_{ji} g(v_j) \\
&= \sum_{j=1}^{n} a_{ji} \left(\sum_{k=1}^{p} b_{kj} w_k \right) \\
&= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} b_{kj} a_{ji} \right) w_k.
\end{aligned}$$

Hence the (k,i)-th entry of $[g \circ f]_{(B_1,B_3)}$ is $\sum_{j=1}^{n} b_{kj} a_{ji}$, which is nothing other than the (k,i)-th entry of the product BA. Hence we conclude that $[g \circ f]_{(B_1,B_3)} = [g]_{(B_2,B_3)} \times [f]_{(B_1.B_2)}$.

Thus we have proved the following theorem.

Theorem 20. *Let U, V and W be finite dimensional vector spaces. Let $f\colon U \to V$ and $g\colon V \to W$ be linear maps. Let $B_1 := \{u_i : 1 \leq i \leq m\}$ (respectively, $B_2 := \{v_j : 1 \leq j \leq n\}$, $B_3 := \{w_k : 1 \leq k \leq p\}$) be an ordered basis of U (ordered basis of V, W respectively). Let $A = [f]_{(B_1,B_2)}$ and $B := [g]_{(B_2,B_3)}$. Then we have the matrix of $g \circ f$ relative to the bases B_1 on the domain and B_3 on the co-domain is BA.* □

Example 21. Let us verify the result in a simple case. Let $U = \mathbb{R}^2$, $V = \mathbb{R}^4$ and $W = \mathbb{R}^3$, all with the standard ordered bases. Let $f\colon U \to V$ be defined by $f(x_1, x_2) := (x_1, x_2, x_1 + x_2, x_1 - x_2)$. Let $g\colon V \to W$ be defined by $g(y_1, y_2, y_3, y_4) := (y_1 + y_2, , y_2 + y_3, y_3 + y_4)$. Then $(g \circ f)(x_1, x_2) = (x_1 + x_2, x_1 + 2x_2, 2x_1)$. The matrices of these maps relative to the standard ordered bases are

$$[f] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad [g] = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad [g \circ f] = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 0 \end{pmatrix}.$$

The reader may carry out the matrix multiplication and verify the result.

Ex. 22. A special case of the last theorem. Let $U = V$ and $W = V$. Let B and B' be ordered bases of V. Consider the identity map $I \colon V \to V$. Then $[I]_{(B,B')} = A_{B',B}$ and $[I]_{(B',B)} = A_{B,B'}$. Now the product $[I]_{(B',B)}[I]_{(B,B')} = [I]_{(B,B)}$ is the identity matrix. Hence conclude that the product $A_{B,B'}A_{B',B}$ is the identity matrix and hence the transition matrices are invertible.

We now come to the last result of this article. We keep the notation above. Let B_1 and B_1' be ordered bases of V and B_2 and B_2' be two ordered bases of W. Let $M := [f]_{(B_1,B_2)}$ and $N := [f]_{(B_1',B_2')}$ be the matrix representations of f relative to the pairs (B_1, B_2) and (B_1', B_2'). Is there a way to recover one, say, N if we know M? Is there any relation between them in terms of the bases involved?

The answer is surprisingly simple, provided that you have understood what was seen above. Let $R := A_{(B_1',B_1)}$ and $S := A_{(B_2',B_2)}$ denote the corresponding transition matrices. By definition, $N = [f]_{(B_1',B_2')}$. Note that for $w \in W$, we have

$$[w]_{B_2} = S[w]_{B_2'} \implies [w]_{B_2'} = S^{-1}[w]_{B_2}.$$

To find the relation between M and N, the strategy is to start with $[f(v)]_{B_2'}$, express it in terms of $[f(v)]_{B_2}$ and use (7). Let us put this to work. Let $v \in V$. Then we have

$$[v]_{B_1} = R[v]_{B_1'} \tag{8}$$
$$[w]_{B_2'} = S^{-1}[w]_{B_2} \tag{9}$$
$$[f(v)]_{B_2} = M[v]_{B_1}. \tag{10}$$

We substitute $w = f(v)$ in (9). We then get $[f(v)]_{B_2'} = S^{-1}[f(v)]_{B_2}$. We therefore obtain

$$\begin{aligned} [f(v)]_{B_2'} &= S^{-1}[f(v)]_{B_2} \\ &= S^{-1}M[v]_{B_1} \\ &= S^{-1}MR[v]_{B_1'}. \end{aligned}$$

Hence, by the uniqueness part of Observation 15, we conclude that

$$N = S^{-1}MR,$$

where S (respectively, R) is the transition matrix from the basis B_2' to B_2 (respectively, B_1' to B_1).

Most important special case is when $V = W$ and M is the matrix of f relative to the ordered basis $B := B_1 = B_2$ on both the domain and the co-domain and N is the matrix of f relative to the ordered basis $B' := B_1' = B_2'$ on both the domain and the co-domain. Using (6), we then get $N = [I]_{(B,B')}^{-1} M [I]_{(B,B')} = A_{(B',B)}^{-1} M A_{(B',B)}$. (Verify this.)

What is the significance of this observation? We have seen above that once we fix ordered bases of V and W, we have a linear isomorphism of $L(V,W)$ and $M_{n\times m}(\mathbb{R})$ (with the notation above.) In particular, if $V = W$, then $L(V) \simeq M(n,\mathbb{R})$. Let $f \in L(V)$. Let B and B' be ordered bases of V. Then $A := [f]_B$ and $A' := [f]_{B'}$ are 'similar' in $M(n,\mathbb{R})$, that is, there exists an invertible $T \in M(n,\mathbb{R})$ such that $A' = T^{-1}AT$.

Let us look at this in a different perspective. Given $A \in M(n,\mathbb{R})$, we know that it gives rise to a linear map $f \in L(V)$ in such way that the matrix representations of f relative to the standard basis of $\mathbb{R}^n$ is A. Let $A' \in M(n,\mathbb{R})$ be such that $A' = T^{-1}AT$ fo some invertible $T \in M(n,\mathbb{R})$. Can we think of a basis B' of $\mathbb{R}^n$ such that $[f]_{B'} = A'$?

What the last two paragraphs say is that A and A' may actually be representing the same linear map in different ordered bases. Think over this.

From a physicist's point of view, choosing a basis gives us a frame of reference, coordinates of the vector, a matrix for a linear map. But vectors and linear maps exist without any reference to basis. Linear maps express the 'phenomenon' or the 'transformation' free of frames of reference! If you think Mathematics is abstract, search for what Einstein wanted when he was ready with his theories of relativity! He was asking for a mathematical theory which can express physical phenomenon independent of the frame of reference.

Revisit the concluding paragraph of Section 2. Look at the equation above $A' = T^{-1}AT$. If you ponder over the connection, you have started your journey to the canonical forms of matrices. All the best.

Acknowledgment: I thank Professor Bhaba Kumar Sarma (IITG) for sending me a list of typos, corrections and suggestions for improvement.