Linear Maps

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling

(February 5, 2007)

As we have discussed in the lecture on "What is Linear Algebra?" one of the main goals of linear algebra is the characterization of the solutions to the set of m linear equations in n unknowns $x_1, \ldots, x_n$

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad \vdots \qquad &\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

where all coefficients a_{ij} and b_i are in $\mathbb{F}$. Linear maps and their properties that we are about to discuss give us a lot of insight into the characteristics of the solutions.

1 Definition and elementary properties

Throughout this chapter V, W are vector spaces over $\mathbb{F}$. We are going to study maps from V to W that have special properties.

Definition 1. A function $T : V \to W$ is called **linear** if

$$T(u+v) = T(u) + T(v) \qquad \text{for all } u, v \in V, \tag{1}$$

$$T(av) = aT(v) \qquad \text{for all } a \in \mathbb{F} \text{ and } v \in V. \tag{2}$$

The set of all linear maps from V to W is denoted by $\mathcal{L}(V, W)$. We also write Tv for $T(v)$.

Example 1.

1. The **zero** map $0 : V \to W$ mapping every element $v \in V$ to $0 \in W$ is linear.

2. The **identity** map $I : V \to V$ defined as $Iv = v$ is linear.

3. Let $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ be the **differentiation** map defined as $Tp(z) = p'(z)$. Then for two polynomials $p(z), q(z) \in \mathcal{P}(\mathbb{F})$ we have

$$T(p(z) + q(z)) = (p(z) + q(z))' = p'(z) + q'(z) = T(p(z)) + T(q(z)).$$

Similarly for a polynomial $p(z) \in \mathcal{P}(\mathbb{F})$ and a scalar $a \in \mathbb{F}$ we have

$$T(ap(z)) = (ap(z))' = ap'(z) = aT(p(z)).$$

Hence T is linear.

4. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the map given by $T(x, y) = (x - 2y, 3x + y)$. Then for $(x, y), (x', y') \in \mathbb{R}^2$ we have

$$\begin{aligned} T((x, y) + (x', y')) &= T(x + x', y + y') = (x + x' - 2(y + y'), 3(x + x') + y + y') \\ &= (x - 2y, 3x + y) + (x' - 2y', 3x' + y') = T(x, y) + T(x', y'). \end{aligned}$$

Similarly, for $(x, y) \in \mathbb{R}^2$ and $a \in \mathbb{F}$ we have

$$T(a(x, y)) = T(ax, ay) = (ax - 2ay, 3ax + ay) = a(x - 2y, 3x + y) = aT(x, y).$$

Hence T is linear. More generally, any map $T : \mathbb{F}^n \to \mathbb{F}^m$ defined by

$$T(x_1, \dots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \dots, a_{m1}x_1 + \cdots + a_{mn}x_n)$$

with $a_{ij} \in \mathbb{F}$ is linear.

5. Not all functions are linear! For example the exponential function $f(x) = e^x$ is not linear since $e^{2x} \neq 2e^x$. Also the function $f : \mathbb{F} \to \mathbb{F}$ given by $f(x) = x - 1$ is not linear since $f(x + y) = (x + y) - 1 \neq (x - 1) + (y - 1) = f(x) + f(y)$.

An important result is that linear maps are already completely determined if their values on basis vectors are specified.

Theorem 1. *Let $(v_1, \dots, v_n)$ be a basis of V and $(w_1, \dots, w_n)$ an arbitrary list of vectors in W. Then there exists a unique linear map*

$$T : V \to W \quad \textit{such that } T(v_i) = w_i.$$

Proof. First we verify that there is at most one linear map T with $T(v_i) = w_i$. Take any $v \in V$. Since $(v_1, \dots, v_n)$ is a basis of V there are unique scalars $a_1, \dots, a_n \in \mathbb{F}$ such that $v = a_1v_1 + \cdots + a_nv_n$. By linearity we must have

$$T(v) = T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n) = a_1w_1 + \cdots + a_nw_n, \tag{3}$$

and hence $T(v)$ is completely determined. To show existence, use (3) to define T. It remains to show that this T is linear and that $T(v_i) = w_i$. These two conditions are not hard to show and are left to the reader. $\square$

The set of linear maps $\mathcal{L}(V, W)$ is itself a vector space. For $S, T \in \mathcal{L}(V, W)$ addition is defined as

$$(S + T)v = Sv + Tv \quad \text{for all } v \in V.$$

For $a \in \mathbb{F}$ and $T \in \mathcal{L}(V, W)$ scalar multiplication is defined as

$$(aT)(v) = a(Tv) \quad \text{for all } v \in V.$$

You should verify that $S + T$ and aT are indeed linear maps again and that all properties of a vector space are satisfied.

In addition to addition and scalar multiplication we can defined the **composition of linear maps**. Let V, U, W be vector spaces over $\mathbb{F}$. Then for $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$, we define $T \circ S \in \mathcal{L}(U, W)$ as

$$(T \circ S)(u) = T(S(u)) \quad \text{for all } u \in U.$$

The map $T \circ S$ is often also called the **product** of T and S denoted by TS. It has the following properties:

1. **Associativity:** $(T_1T_2)T_3 = T_1(T_2T_3)$ for all $T_1 \in \mathcal{L}(V_1, V_0)$, $T_2 \in \mathcal{L}(V_2, V_1)$ and $T_3 \in \mathcal{L}(V_3, V_2)$.

2. **Identity:** $TI = IT = T$ where $T \in \mathcal{L}(V, W)$ and I in TI is the identity map in $\mathcal{L}(V, V)$ and I in IT is the identity map in $\mathcal{L}(W, W)$.

3. **Distributive property:** $(T_1 + T_2)S = T_1S + T_2S$ and $T(S_1 + S_2) = TS_1 + TS_2$ where $S, S_1, S_2 \in \mathcal{L}(U, V)$ and $T, T_1, T_2 \in \mathcal{L}(V, W)$.

Note that the product of linear maps is not always commutative. For example if $T \in \mathcal{L}(\mathcal{P}(\mathbb{F}), \mathcal{P}(\mathbb{F}))$ is differentiation $Tp(z) = p'(z)$ and $S \in \mathcal{L}(\mathcal{P}(\mathbb{F}), \mathcal{P}(\mathbb{F}))$ is multiplication by z^2 given by $Sp(z) = z^2p(z)$, then

$$(ST)p(z) = z^2p'(z) \quad \text{but} \quad (TS)p(z) = z^2p'(z) + 2zp(z).$$

2 Null spaces

Definition 2. Let $T : V \to W$ be a linear map. Then the **null space** or **kernel** of T is the set of all vectors in V that map to zero:

$$\operatorname{null} T = \{v \in V \mid Tv = 0\}.$$

Example 2. Let $T \in \mathcal{L}(\mathcal{P}(\mathbb{F}), \mathcal{P}(\mathbb{F}))$ be differentiation $Tp(z) = p'(z)$. Then

$$\operatorname{null} T = \{p \in \mathcal{P}(\mathbb{F}) \mid p(z) \text{ is constant}\}.$$

Proposition 2. *Let $T : V \to W$ be a linear map. Then $\operatorname{null} T$ is a subspace of V.*

Proof. We need to show that $0 \in \operatorname{null} T$ and that $\operatorname{null} T$ is closed under addition and scalar multiplication. By linearity we have

$$T(0) = T(0 + 0) = T(0) + T(0)$$

so that $T(0) = 0$. Hence $0 \in \operatorname{null} T$. For closure under addition let $u, v \in \operatorname{null} T$. Then

$$T(u + v) = T(u) + T(v) = 0 + 0 = 0,$$

and hence $u + v \in \operatorname{null} T$. Similarly for closure under scalar multiplication, let $u \in \operatorname{null} T$ and $a \in \mathbb{F}$. Then

$$T(au) = aT(u) = a0 = 0,$$

so that $au \in \operatorname{null} T$. □

Definition 3. The linear map $T : V \to W$ is called **injective** if for all $u, v \in V$, the condition $Tu = Tv$ implies that $u = v$. In other words, different vectors in V are mapped to different vector in W.

Proposition 3. *Let $T : V \to W$ be a linear map. Then T is injective if and only if $\operatorname{null} T = \{0\}$.*

Proof.
"$\Longrightarrow$" Suppose that T is injective. Since $\operatorname{null} T$ is a subspace of V, we know that $0 \in \operatorname{null} T$. Assume that there is another vector $v \in V$ that is in the kernel. Then $T(v) = 0 = T(0)$. Since T is injective this implies that $v = 0$, proving that $\operatorname{null} T = \{0\}$.
"$\Longleftarrow$" Assume that $\operatorname{null} T = \{0\}$. Let $u, v \in V$ such that $Tu = Tv$. Then $0 = Tu - Tv = T(u - v)$, so that $u - v \in \operatorname{null} T$. Hence $u - v = 0$ or equivalently $u = v$ showing that T is indeed injective. □

Example 3.

1. The differentiation map $p(z) \mapsto p'(z)$ is not injective since $p'(z) = q'(z)$ implies that $p(z) = q(z) + c$ where $c \in \mathbb{F}$ is a constant.

2. The identity map $I : V \to V$ is injective.

3. The linear map $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ given by $T(p(z)) = z^2 p(z)$ is injective since $\operatorname{null} T = \{0\}$.

3 Ranges

Definition 4. Let $T : V \to W$ be a linear map. The **range** of T, denoted by $\operatorname{range} T$, is the subset of vectors of W that are in the image of T

$$\operatorname{range} T = \{Tv \mid v \in V\} = \{w \in W \mid \text{ there exists } v \in V \text{ such that } Tv = w\}.$$

Example 4. The range of the differentiation map $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ is $\operatorname{range} T = \mathcal{P}(\mathbb{F})$ since for every polynomial $q \in \mathcal{P}(\mathbb{F})$ there is a $p \in \mathcal{P}(\mathbb{F})$ such that $p' = q$.

Proposition 4. *Let $T : V \to W$ be a linear map. Then $\operatorname{range} T$ is a subspace of W.*

Proof. We need to show that $0 \in \operatorname{range} T$ and that $\operatorname{range} T$ is closed under addition and scalar multiplication. We already showed that $T0 = 0$ so that $0 \in \operatorname{range} T$.

For closure under addition let $w_1, w_2 \in \operatorname{range} T$. Then there exist $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$. Hence

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2$$

so that $w_1 + w_2 \in \operatorname{range} T$.

For closure under scalar multiplication, let $w \in \operatorname{range} T$ and $a \in \mathbb{F}$. Then there exists a $v \in V$ such that $Tv = w$. Thus

$$T(av) = aTv = aw$$

so that $aw \in \operatorname{range} T$. □

Definition 5. A linear map $T : V \to W$ is called **surjective** if $\operatorname{range} T = W$. A linear map $T : V \to W$ is called **bijective** if T is injective and surjective.

Example 5. The differentiation map $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ is surjective since $\operatorname{range} T = \mathcal{P}(\mathbb{F})$. However, if we restrict ourselves to polynomials of degree at most m, then the differentiation map $T : \mathcal{P}_m(\mathbb{F}) \to \mathcal{P}_m(\mathbb{F})$ is not surjective since polynomials of degree m are not in the range of T.

4 Homomorphisms

It should be mentioned that linear maps between vector spaces are also called **vector space homomorphisms**. Instead of the notation $\mathcal{L}(V, W)$ one often sees the convention

$$\operatorname{Hom}_{\mathbb{F}}(V, W) = \{T : V \to W \mid T \text{ is linear}\}.$$

A homomorhpism $T : V \to W$ is called

- **Monomorphism** iff T is injective;
- **Epimorphism** iff T is surjective;
- **Isomorphism** iff T is bijective;
- **Endomorphism** iff $V = W$;
- **Automorphism** iff $V = W$ and T is bijective.

5 The dimension formula

The next theorem is the key result of this chapter. It relates the dimension of the kernel and range of a linear map.

Theorem 5. *Let V be a finite-dimensional vector space and $T : V \to W$ a linear map. Then* $\operatorname{range} T$ *is a finite-dimensional subspace of W and*

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T. \tag{4}$$

Proof. Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V, W)$. Since $\operatorname{null} T$ is a subspace of V, we know that $\operatorname{null} T$ has a basis $(u_1, \ldots, u_m)$. This implies that $\dim \operatorname{null} T = m$. By the Basis Extension Theorem it follows that $(u_1, \ldots, u_m)$ can be extended to a basis of V, say $(u_1, \ldots, u_m, v_1, \ldots, v_n)$, so that $\dim V = m + n$.

The theorem will follow by showing that $(Tv_1, \ldots, Tv_n)$ is a basis of $\operatorname{range} T$ since this would imply that $\operatorname{range} T$ is finite-dimensional and $\dim \operatorname{range} T = n$ proving (4).

Since $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ spans V, every $v \in V$ can be written as a linear combination of these vectors

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n$$

where $a_i, b_j \in \mathbb{F}$. Applying T to v we obtain

$$Tv = b_1 Tv_1 + \cdots + b_n Tv_n,$$

where the terms Tu_i disappeared since $u_i \in \operatorname{null} T$. This shows that $(Tv_1, \ldots, Tv_n)$ indeed spans $\operatorname{range} T$.

To show that $(Tv_1, \ldots, Tv_n)$ is a basis of $\operatorname{range} T$ it remains to show that this list is linearly independent. Assume that $c_1, \ldots, c_n \in \mathbb{F}$ are such that

$$c_1 Tv_1 + \cdots + c_n Tv_n = 0.$$

By linearity of T this implies that

$$T(c_1 v_1 + \cdots + c_n v_n) = 0,$$

so that $c_1v_1 + \cdots + c_nv_n \in \operatorname{null} T$. Since $(u_1, \ldots, u_m)$ is a basis of $\operatorname{null} T$ there must exist scalars $d_1, \ldots, d_m \in \mathbb{F}$ such that

$$c_1v_1 + \cdots + c_nv_n = d_1u_1 + \cdots + d_mu_m.$$

However by the linear independence of $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ this implies that all coefficients $c_1 = \cdots = c_n = d_1 = \cdots = d_m = 0$. Thus $(Tv_1, \ldots, Tv_n)$ is linearly independent and we are done. □

Corollary 6. *Let $T \in \mathcal{L}(V, W)$.*

1. *If $\dim V > \dim W$, then T is not injective.*
2. *If $\dim V < \dim W$, then T is not surjective.*

Proof. By Theorem 5 we have

$$\begin{aligned} \dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \\ &\geq \dim V - \dim W > 0. \end{aligned}$$

Since T is injective if and only if $\dim \operatorname{null} T = 0$, T cannot be injective.

Similarly,

$$\begin{aligned} \dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T \\ &\leq \dim V < \dim W, \end{aligned}$$

so that $\operatorname{range} T$ cannot be equal to W. Hence T cannot be surjective. □

6 The matrix of a linear map

Now we will see that every linear map can be encoded by a matrix, and vice versa every matrix defines a linear map.

Let V, W be finite-dimensional vector spaces, and let $T : V \to W$ be a linear map. Suppose that $(v_1, \ldots, v_n)$ is a basis of V and $(w_1, \ldots, w_m)$ is a basis for W. We have seen in Theorem 1 that T is uniquely determined by specifying the vectors $Tv_1, \ldots, Tv_n \in W$. Since $(w_1, \ldots, w_m)$ is a basis of W there exist unique scalars $a_{ij} \in \mathbb{F}$ such that

$$Tv_j = a_{1j}w_1 + \cdots + a_{mj}w_m \quad \text{for } 1 \leq j \leq n. \tag{5}$$

We can arrange these scalars in an $m \times n$ matrix as follows

$$M(T) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

with m rows and n columns. Often this is also written as $A = (a_{ij})_{1 \le i \le m, 1 \le j \le n}$. The set of all $m \times n$ matrices with entries in $\mathbb{F}$ is denoted by $\mathbb{F}^{m \times n}$.

Remark 7. It is important to remember that $M(T)$ not only depends on the linear map T, but also on the choice of the basis $(v_1, \dots, v_n)$ for V and $(w_1, \dots, w_m)$ for W. The j-th column of $M(T)$ contains the coefficients of the j-th basis vector v_j expanded in terms of the basis $(w_1, \dots, w_m)$ as in (5).

Example 6. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map given by $T(x, y) = (ax + by, cx + dy)$ for some $a, b, c, d \in \mathbb{R}$. Then with respect to the canonical basis of $\mathbb{R}^2$ given by $((1, 0), (0, 1))$ the corresponding matrix is

$$M(T) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

since $T(1, 0) = (a, c)$ gives the first column and $T(0, 1) = (b, d)$ gives the second column.

More generally, if $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$ with the standard basis $(e_1, \dots, e_n)$ for V and $(f_1, \dots, f_m)$ for W, where e_i (resp. f_i) is the n-tuple (resp. m-tuple) with a one in position i and zeroes everywhere else, then the matrix $M(T) = (a_{ij})$ is given by

$$a_{ij} = (Te_j)_i.$$

Example 7. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear map defined as $T(x, y) = (y, x + 2y, x + y)$. Then with respect to the standard basis we have $T(1, 0) = (0, 1, 1)$ and $T(0, 1) = (1, 2, 1)$ so that

$$M(T) = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

However, if alternatively we take the basis $((1, 2), (0, 1))$ of $\mathbb{R}^2$ and $((1, 0, 0), (0, 1, 0), (0, 0, 1))$ of $\mathbb{R}^3$, then $T(1, 2) = (2, 5, 3)$ and $T(0, 1) = (1, 2, 1)$ so that

$$M(T) = \begin{bmatrix} 2 & 1 \\ 5 & 2 \\ 3 & 1 \end{bmatrix}.$$

Example 8. Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map $S(x, y) = (y, x)$. With respect to the basis

$((1,2),(0,1))$ of $\mathbb{R}^2$ we have

$$S(1,2) = (2,1) = 2(1,2) - 3(0,1) \quad \text{and} \quad S(0,1) = (1,0) = 1(1,2) - 2(0,1),$$

so that

$$M(S) = \begin{bmatrix} 2 & 1 \\ -3 & -2 \end{bmatrix}.$$

Note that given the vector spaces V and W of dimensions n and m, respectively, and a fixed choice of bases, there is a one-to-one correspondence between linear maps in $\mathcal{L}(V,W)$ and matrices in $\mathbb{F}^{m\times n}$. Given the linear map T, the matrix $M(T) = A = (a_{ij})$ is defined via (5). Conversely, given the matrix $A = (a_{ij}) \in \mathbb{F}^{m\times n}$ we can define a linear map $T : V \to W$ by setting

$$Tv_j = \sum_{i=1}^{m} a_{ij} w_i.$$

Recall that we saw that the set of linear maps $\mathcal{L}(V,W)$ is a vector space. Since we have a one-to-one correspondence between linear maps and matrices we can also make the set of matrices $\mathbb{F}^{m\times n}$ into a vector space. Given two matrices $A = (a_{ij})$ and $B = (b_{ij})$ in $\mathbb{F}^{m\times n}$ and a scalar $\alpha \in \mathbb{F}$ we define the **sum** of matrices and **scalar multiplication** componentwise:

$$\begin{aligned} A + B &= (a_{ij} + b_{ij}) \\ \alpha A &= (\alpha a_{ij}). \end{aligned}$$

Next we show that the **composition** of linear maps imposes a product on matrices, also called **matrix multiplication**. Suppose U, V, W are vector spaces over $\mathbb{F}$ with bases $(u_1, \dots, u_p)$, $(v_1, \dots, v_n)$ and $(w_1, \dots, w_m)$, respectively. Let $S : U \to V$ and $T : V \to W$ be linear maps. Then the product is a linear map $T \circ S : U \to W$.

Each linear map has its corresponding matrix $M(T) = A, M(S) = B$ and $M(TS) = C$. The question is whether C is determined by A and B. We have for $j \in \{1, 2, \dots p\}$

$$\begin{aligned} (T \circ S)u_j &= T(b_{1j}v_1 + \cdots + b_{nj}v_n) = b_{1j}Tv_1 + \cdots + b_{nj}Tv_n \\ &= \sum_{k=1}^{n} b_{kj} T v_k = \sum_{k=1}^{n} b_{kj} \big(\sum_{i=1}^{m} a_{ik} w_i \big) \\ &= \sum_{i=1}^{m} \big(\sum_{k=1}^{n} a_{ik} b_{kj} \big) w_i. \end{aligned}$$

Hence the matrix $C = (c_{ij})$ is given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{6}$$

Equation (6) can be used as the definition of the $m \times p$ matrix C defined as the product of a $m \times n$ matrix A and a $n \times p$ matrix B

$$C = AB. \tag{7}$$

Our derivation implies that the correspondence between linear maps and matrices respects the product structure.

Proposition 8. *Let $S : U \to V$ and $T : V \to W$ be linear maps. Then*

$$M(TS) = M(T)M(S).$$

Example 9. Take the matrices of the linear maps $T : \mathbb{R}^2 \to \mathbb{R}^3$ with bases $((1,2),(0,1))$ for $\mathbb{R}^2$ and the standard basis for $\mathbb{R}^3$ and $S : \mathbb{R}^2 \to \mathbb{R}^2$ with basis $((1,2),(0,1))$ for $\mathbb{R}^2$ of Examples 7 and 8. Then

$$M(TS) = M(T)M(S) = \begin{bmatrix} 2 & 1 \\ 5 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -3 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \\ 3 & 1 \end{bmatrix}.$$

Given a vector $v \in V$ we can also associate a matrix $M(v)$ to v as follows. Let $(v_1, \ldots, v_n)$ be a basis of V. Then there are unique scalars $b_1, \ldots, b_n$ such that

$$v = b_1 v_1 + \cdots b_n v_n.$$

The matrix of v is the $n \times 1$ matrix defined as

$$M(v) = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Example 10. The matrix of a vector $x = (x_1, \ldots, x_n) \in \mathbb{F}^n$ in the standard basis $(e_1, \ldots, e_n)$ is the column vector or $n \times 1$ matrix

$$M(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

since $x = (x_1, \ldots, x_n) = x_1 e_1 + \cdots + x_n e_n$.

The next result shows how the notion of a matrix of a linear map $T : V \to W$ and the matrix of a vector $v \in V$ fit together.

Proposition 9. *Let $T : V \to W$ be a linear map. Then for every $v \in V$*

$$M(Tv) = M(T)M(v).$$

Proof. Let $(v_1, \dots, v_n)$ be a basis of V and $(w_1, \dots, w_m)$ a basis for W. Suppose that with respect to these bases the matrix of T is $M(T) = (a_{ij})_{1 \le i \le m, 1 \le j \le n}$. Recall that this means that for all $j \in \{1, 2, \dots, n\}$

$$Tv_j = \sum_{k=1}^{m} a_{kj} w_k.$$

The vector $v \in V$ can be written uniquely as a linear combination of the basis vectors as

$$v = b_1 v_1 + \cdots + b_n v_n.$$

Hence

$$\begin{aligned} Tv &= b_1 Tv_1 + \cdots + b_n Tv_n \\ &= b_1 \sum_{k=1}^{m} a_{k1} w_k + \cdots + b_n \sum_{k=1}^{m} a_{kn} w_k \\ &= \sum_{k=1}^{m} (a_{k1} b_1 + \cdots + a_{kn} b_n) w_k. \end{aligned}$$

This shows that $M(Tv)$ is the $m \times 1$ matrix

$$M(Tv) = \begin{bmatrix} a_{11} b_1 + \cdots + a_{1n} b_n \\ \vdots \\ a_{m1} b_1 + \cdots + a_{mn} b_n \end{bmatrix}.$$

It is not hard to check using the formula for matrix multiplication that $M(T)M(v)$ gives the same result. $\square$

Example 11. Take the linear map S from Example 8 with basis $((1,2),(0,1))$ of $\mathbb{R}^2$. To determine the action on the vector $v = (1,4) \in \mathbb{R}^2$ note that $v = (1,4) = 1(1,2) + 2(0,1)$. Hence

$$M(Sv) = M(S)M(v) = \begin{bmatrix} 2 & 1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}.$$

This means

$$Sv = 4(1,2) - 7(0,1) = (4,1)$$

which is indeed true.

7 Invertibility

Definition 6. A linear map $T : V \to W$ is called **invertible** if there exists a linear map $S : W \to V$ such that
$$TS = I_W \quad \text{and} \quad ST = I_V,$$
where $I_V : V \to V$ is the identity map on V and $I_W : W \to W$ is the identity map on W. We say that S is an **inverse** of T.

Note that if the linear map T is invertible, then the inverse is unique. Suppose S and R are inverses of T. Then
$$ST = I_V = RT$$
$$TS = I_W = TR.$$

Hence
$$S = S(TR) = (ST)R = R.$$
We denote the unique inverse of an invertible linear map T by T^{-1}.

Proposition 10. *A linear map $T \in \mathcal{L}(V, W)$ is invertible if and only if T is injective and surjective.*

Proof.
"$\Longrightarrow$" Suppose T is invertible.

To show that T is injective, suppose that $u, v \in V$ are such that $Tu = Tv$. Apply the inverse T^{-1} of T to obtain $T^{-1}Tu = T^{-1}Tv$ so that $u = v$. Hence T is injective.

To show that T is surjective, we need to show that for every $w \in W$ there is a $v \in V$ such that $Tv = w$. Take $v = T^{-1}w \in V$. Then $T(T^{-1}w) = w$. Hence T is surjective.
"$\Longleftarrow$" Suppose that T is injective and surjective. We need to show that T is invertible. We define a map $S \in \mathcal{L}(W, V)$ as follows. Since T is surjective, we know that for every $w \in W$ there exists a $v \in V$ such that $Tv = w$. Moreover, since T is injective, this v is uniquely determined. Hence define $Sw = v$.

We claim that S is the inverse of T. Note that for all $w \in W$ we have $TSw = Tv = w$ so that $TS = I_W$. Similarly for all $v \in V$ we have $STv = Sw = v$ so that $ST = I_V$.

It remains to show that S is a linear map. For all $w_1, w_2 \in W$ we have
$$T(Sw_1 + Sw_2) = TSw_1 + TSw_2 = w_1 + w_2,$$
so that $Sw_1 + Sw_2$ is the unique vector v in V such that $Tv = w_1 + w_2 = w$. Hence
$$Sw_1 + Sw_2 = v = Sw = S(w_1 + w_2).$$

The proof that $S(aw) = aSw$ is similar. For $w \in W$ and $a \in \mathbb{F}$ we have

$$T(aSw) = aT(Sw) = aw$$

so that aSw is the unique vector in V that maps to aw. Hence $S(aw) = aSw$. □

Definition 7. Two vector spaces V and W are called **isomorphic** if there exists an invertible linear map $T \in \mathcal{L}(V, W)$.

Theorem 11. *Two finite-dimensional vector spaces V and W over $\mathbb{F}$ are isomorphic if and only if* $\dim V = \dim W$.

Proof.
"$\Longrightarrow$" Suppose V and W are isomorphic. Then there exists an invertible linear map $T \in \mathcal{L}(V, W)$. Since T is invertible, it is injective and surjective, so that $\operatorname{null} T = \{0\}$ and $\operatorname{range} T = W$. From the dimension formula this implies that $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim W$.
"$\Longleftarrow$" Suppose that $\dim V = \dim W$. Let $(v_1, \ldots, v_n)$ be a basis of V and $(w_1, \ldots, w_n)$ a basis of W. Define the linear map $T : V \to W$ as

$$T(a_1 v_1 + \cdots + a_n v_n) = a_1 w_1 + \cdots + a_n w_n.$$

Since the scalars $a_1, \ldots, a_n \in \mathbb{F}$ are arbitrary and $(w_1, \ldots, w_n)$ spans W, this means that $\operatorname{range} T = W$ and T is surjective. Also, since $(w_1, \ldots, w_n)$ is linearly independent, T is injective (since $a_1 w_1 + \cdots + a_n w_n = 0$ implies that all $a_1 = \cdots = a_n = 0$ and hence only the zero vector is mapped to zero). Hence T is injective and surjective and by Proposition 10 invertible. Therefore, V and W are isomorphic. □

Next we consider the case of linear maps from a vector space V to itself. These linear maps are also called **operators**. The next remarkable theorem shows that the notion of injectivity, surjectivity and invertibility of T are the same, as long as V is finite-dimensional. For infinite-dimensional vector spaces this is not true. For example the set of all polynomials $\mathcal{P}(\mathbb{F})$ is an infinite-dimensional vector space and we saw that the differentiation map is surjective, but not injective!

Theorem 12. *Let V be a finite-dimensional vector space and $T : V \to V$ a linear map. Then the following are equivalent:*

1. *T is invertible.*
2. *T is injective.*
3. *T is surjective.*

Proof. By Proposition 10 1 implies 2.

Next we show that 2 implies 3. If T is injective, then we know that $\operatorname{null} T = \{0\}$. Hence by the dimension formula we have

$$\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T = \dim V.$$

Since $\operatorname{range} T \subset V$ is a subspace of V, this implies that $\operatorname{range} T = V$ and hence T is surjective.

Finally we show that 3 implies 1. Since by assumption T is surjective, we have $\operatorname{range} T = V$. Hence again by the dimension formula

$$\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T = 0,$$

so that $\operatorname{null} T = \{0\}$, and hence T is injective. By Proposition 10 an injective and surjective linear map is invertible. $\square$