12. Linear transformations

Definition 12.1. *Let $\phi \colon V \longrightarrow W$ be a function between two vector spaces V and W over the same field F. We say that ϕ* ***respects addition*** *if*

$$\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2),$$

whenever v_1 and v_2 belong to V. We say that ϕ ***respects scalar multiplication*** *if*

$$\phi(rv) = r\phi(v),$$

whenever $v \in V$ and $r \in F$.

We say that ϕ is ***linear*** *if it respects addition and scalar multiplication. We say that ϕ is an* ***linear isomorphism*** *(or isomorphism of vector spaces) if ϕ is linear and ϕ is a bijection.*

We have already seen that any matrix gives rise to a function. In fact these functions are always linear:

Lemma 12.2. *Let $A \in M_{m,n}(F)$.*

If $\phi \colon F^n \longrightarrow F^m$ is the function $\phi(v) = Av$ then ϕ is linear.

Proof. This is easy. Suppose v and $w \in F^n$. Then

$$\begin{aligned} \phi(v + w) &= A(v + w) \\ &= Av + Aw \\ &= \phi(v) + \phi(w). \end{aligned}$$

So ϕ respects addition. Suppose $v \in F^n$ and $\lambda \in F$. Then

$$\begin{aligned} \phi(\lambda v) &= A(\lambda v) \\ &= \lambda(Av) \\ &= \lambda\phi(v). \end{aligned}$$

Hence ϕ respects scalar multiplication. Thus ϕ is linear. □

The interesting thing is that the opposite is true:

Theorem 12.3. *Let $\phi \colon F^n \longrightarrow F^m$ be a linear map.*

Then there is a matrix $M_{m,n}(F)$ such that $\phi(v) = Av$.

Proof. Let A be the matrix whose ith column is $\phi(e_i)$. Suppose that $v \in V$. If $v = (r_1, r_2, \dots, r_n) \in F^n$ then $v = \sum r_i e_i$. In this case

$$\begin{aligned}\phi(v) &= \phi(r_1 e_1 + r_2 e_2 + \cdots + r_n e_n)\\ &= r_1\phi(e_1) + r_2\phi(e_2) + \cdots + r_n\phi(e_n)\\ &= r_1 A e_1 + r_2 A e_2 + \cdots + r_n A e_n\\ &= A(r_1 e_1 + r_2 e_2 + \cdots + r_n e_n)\\ &= Av.\end{aligned}$$

□

These two results make it easy to check if a map $F^n \longrightarrow F^m$ is linear or not:

Example 12.4. *Let $T\colon F^3 \longrightarrow F^2$ be the function*

$$T(x, y, z) = (2x - 3y + 7z, 6x + y - 3z).$$

Then T is the linear map associated to the matrix

$$\begin{pmatrix} 2 & -3 & 7 \\ 6 & 1 & -3 \end{pmatrix}.$$

It is easy to write down examples of linear maps between abstract vector spaces.

Example 12.5. *Let*

$$\phi\colon V \longrightarrow W,$$

be the function which sends everything to 0. *It is easy to check that this map is linear. Similarly the identity map*

$$\phi\colon V \longrightarrow V \qquad \text{which sends} \qquad v \longrightarrow v,$$

is linear. For a slightly more complicated example, consider the map

$$\phi\colon V \longrightarrow V \qquad \text{which sends} \qquad v \longrightarrow \lambda v,$$

where λ is a fixed scalar. Then ϕ is linear. Indeed

$$\begin{aligned}\phi(v + w) &= \lambda(v + w)\\ &= \lambda v + \lambda w\\ &= \phi(v) + \phi(w).\end{aligned}$$

Thus ϕ respects addition. Similarly, ir $r \in F$ then

$$\begin{aligned}\phi(rv) &= \lambda(rv)\\ &= (\lambda r)v\\ &= (r\lambda)v\\ &= r(\lambda v)\\ &= r\phi(v).\end{aligned}$$

Thus ϕ respects scalar multiplication and so ϕ is linear.

Let

$$\phi\colon M_{m,n}(F) \longrightarrow M_{n,m}(F),$$

be the function which sends a matrix $A = (a_{ij})$ to its transpose $A^t = (a_{ji})$. Since the tranpose of a sum of two matrices is the sum of the tranposes, ϕ respects addition. Formally if $A = (a_{ij})$ and $B = (b_{ij}) \in M_{m,n}(F)$ then A^t is the matrix with entries (a_{ji}) and B^t is the matrix with entries (b_{ji}). $A + B$ is the matrix with entries $(a_{ij} + b_{ij})$. The transpose of $A + B$ is the matrix with entries $(a_{ji} + b_{ji})$, which is the same as the entries of the matrix $A^t + B^t$. Thus ϕ respects addition. Similarly ϕ respects scalar multiplication. Thus ϕ is linear.

Let

$$\phi\colon P_d(F) \longrightarrow F,$$

be the function which sends a polynomial $f(x)$ of degree at most d to its value $f(0)$ at 0. It is easy to check that this map is linear. For a slightly more interesting example, consider the function

$$\phi\colon P_d(\mathbb{R}) \longrightarrow P_{d-1}(\mathbb{R}),$$

defined by the rule $\phi(f(x)) = f'(x)$ the derivative of $f(x)$. Basic properties of the derivative ensure that this map is linear.

Definition-Lemma 12.6. *Let V be a finite dimensional vector space over a field F. Suppose we pick a basis $v_1, v_2, \dots, v_n$. Then we define a function*

$$\phi\colon V \longrightarrow F^n$$

by the following rule: Given $v \in V$ there are unique scalars $r_1, r_2, \dots, r_n$ such that $v = \sum r_i v_i$. Let $\phi(v) = (r_1, r_2, \dots, r_n) \in F^n$.

Then ϕ is a linear isomorphism.

Proof. Since every vector v has a unique expression in the form $v = \sum r_i v_i$ it is clear that ϕ is well-defined and that ϕ is a bijection. We check that ϕ is linear. Given v and $w \in V$ we may find scalars $r_1, r_2, \dots, r_n$ and $s_1, s_2, \dots, s_n$ such that

$$v = \sum_i r_i v_i \qquad \text{and} \qquad w = \sum_i s_i v_i.$$

Then

$$v + w = \sum (r_i + s_i) v_i.$$

Thus

$$\begin{aligned}\phi(v+w) &= (r_1+s_1, r_2+s_2, \dots, r_n+s_n)\\ &= (r_1, r_2, \dots, r_n) + (s_1, s_2, \dots, s_n)\\ &= \phi(v)+\phi(w),\end{aligned}$$

and so ϕ respects addition. Suppose that $\lambda \in F$. Then

$$\lambda v = (\lambda r_1)v_1 + (\lambda r_2)v_2 + \dots + (\lambda r_n)v_n.$$

It follows that

$$\begin{aligned}\phi(\lambda v) &= (\lambda r_1, \lambda r_2, \dots, \lambda r_n)\\ &= \lambda(r_1, r_2, \dots, r_n)\\ &= \lambda\phi(v),\end{aligned}$$

and so ϕ respects scalar multiplication. Hence ϕ is linear and so ϕ is a linear isomorphism. □

Definition-Lemma 12.7. *Let $\phi\colon V \longrightarrow W$ and $\psi\colon U \longrightarrow V$ be two linear maps between vector spaces over a field F.*

Then

(1) *The composition $\phi \circ \psi\colon U \longrightarrow W$ is linear.*
(2) $\phi(0) = 0$.
(3)
$$\operatorname{Ker}(\phi) = \{\, v \in V \mid \phi(v) = 0 \,\},$$
is a subspace of V, called the ***kernel*** *(aka nullspace) of ϕ. The dimension of $\operatorname{Ker}(\phi)$ is called the* ***nullity*** *of ϕ and is denoted $\nu(\phi)$.*
(4)
$$\operatorname{Im}(\phi) = \{\, w \in W \mid w = \phi(v), v \in V \,\},$$
is a subspace of W, called the image of ϕ. The dimension of $\operatorname{Im}(\phi)$ is called the ***rank*** *of ϕ and is denoted $\operatorname{rk}(\phi)$.*

Proof. We check that the composition respects addition and leave the fact that the composition respects scalar multiplication to the reader. Suppose that u_1 and $u_2 \in U$. We have

$$\begin{aligned}(\phi\circ\psi)(u_1+u_2) &= \phi(\psi(u_1+u_2))\\ &= \phi(\psi(u_1)+\psi(u_2))\\ &= \phi(\psi(u_1))+\phi(\psi(u_2))\\ &= (\phi\circ\psi)(u_1)) + (\phi\circ\psi)(u_2).\end{aligned}$$

Thus $\phi \circ \psi$ respects addition. Hence (1).

We have

$$\begin{aligned}\phi(0) &= \phi(0+0) \\ &= \phi(0) + \phi(0).\end{aligned}$$

If we subtract $\phi(0)$ from both sides then we get (2).

We check that the kernel is non-empty and closed under addition and scalar multiplication. (2) implies that $0 \in \operatorname{Ker}(\phi)$. Suppose that v_1 and $v_2 \in \operatorname{Ker}(\phi)$. Then

$$\begin{aligned}\phi(v_1 + v_2) &= \phi(v_1) + \phi(v_2) \\ &= 0 + 0 \\ &= 0.\end{aligned}$$

But then $v_1 + v_2 \in \operatorname{Ker}(\phi)$ and so $\operatorname{Ker}(\phi)$ is closed under addition. Similarly if $v \in \operatorname{Ker}(\phi)$ and $\lambda \in F$ then

$$\begin{aligned}\phi(\lambda v) &= \lambda\phi(v) \\ &= \lambda 0 \\ &= 0.\end{aligned}$$

Thus $\lambda v \in \operatorname{Ker}(\phi)$ and so $\operatorname{Ker}(\phi)$ is closed under scalar multiplication. Thus $\operatorname{Ker}(\phi)$ is a linear subspace of V and this is (3).

$0 \in \operatorname{Im}(\phi)$ by (2). Suppose that w_1 and $w_2 \in \operatorname{Im}(\phi)$. Then there are vectors v_1 and v_2 such that $w_1 = \phi(v_1)$ and $w_2 = \phi(v_2)$. We have

$$\begin{aligned}w_1 + w_2 &= \phi(v_1) + \phi(v_2) \\ &= \phi(v_1 + v_2).\end{aligned}$$

Thus $w_1 + w_2 \in \operatorname{Im}(\phi)$. Hence $\operatorname{Im}(\phi)$ is closed under addition. Now suppose $w \in \operatorname{Im}(\phi)$ and $\lambda \in F$. Then

$$\begin{aligned}\lambda w &= \lambda\phi(v) \\ &= \phi(\lambda v).\end{aligned}$$

Thus $\operatorname{Im}(T)$ is closed under scalar multiplication. It follows that $\operatorname{Im}(T)$ is a linear subspace. $\square$

Suppose we are given a function $f\colon X \longrightarrow Y$ between two sets X and Y. If $A \subset X$ is a subset of X then we get a function from $f_1\colon A \longrightarrow Y$ in a natural way. If $a \in A$ then $f_1(a) = f(a) \in Y$. $f_1 = f|_A$ is called the restriction of f to A. In fact there is a natural map $i\colon A \longrightarrow X$ called the inclusion map. It is defined by $i(a) = a$. Then $f_1 = f \circ i$, the composition. Note that the functions f and f_1 are not the same functions, since they have different domains (namely X and A). Now suppose that the image of f lands inside B a subset of Y. Then we get a function $f_2\colon X \longrightarrow B$. $f_2(x) = f(x) \in B$. If $j\colon B \longrightarrow Y$ is the natural inclusion of then f_2 factors f that is $f = j \circ f_2$. In particular if B is the image of f then f_2 is surjective.

Lemma 12.8. *Let $\phi\colon V \longrightarrow W$ be a surjective linear map.*
TFAE

(1) *ϕ is a linear isomorphism.*
(2) *ϕ is injective.*
(3) $\operatorname{Ker}(\phi) = \{0\}$.

Proof. (1) implies (2) is clear. (2) implies (3) is easy.
Suppose that $\operatorname{Ker}(\phi) = \{0\}$. Suppose that $\phi(v_1) = \phi(v_2)$. Then

$$\begin{aligned} 0 &= \phi(v_1) - \phi(v_2) \\ &= \phi(v_1) + \phi(-v_2) \\ &= \phi(v_1 - v_2). \end{aligned}$$

But then $v_1 - v_2 \in \operatorname{Ker}(\phi)$. It follows that $v_1 - v_2 = 0$ so that $v_1 = v_2$. Thus ϕ is injective. As ϕ is surjective by assumption it follows that ϕ is an isomorphism. Thus (3) implies (1). $\square$

Example 12.9. *Let $V \subset P_d(\mathbb{R})$ be the set of polynomials whose constant term is zero. Note that V is a subspace (easy check left to the reader). Let*

$$\phi\colon V \longrightarrow P_{d-1}(F),$$

be the derivative. Note that ϕ is onto. Indeed

$$\phi\left(\frac{x^{i+1}}{i+1}\right) = x^i \qquad \textit{for} \qquad i \leq d-1.$$

Since $1, x, x^2, \dots, x^{d-1}$ span $P_{d-1}(F)$ (they are even a basis) it follows that $\operatorname{Im}(\phi) = P_{d-1}(\mathbb{R})$. On the other hand the kernel is trivial (since we threw out the constant polynomials) and so ϕ is an isomorphism.

It is curious to observe that even though we can define the formal derivative over any field (see the hwk) this result fails in general. For example suppose we work over $\mathbb{F}_2$. The problem is that

$$\frac{x^2}{2},$$

makes no sense ($2 = 0$ in $\mathbb{F}_2$). Put differently

$$\phi(x^2) = 2x = 0 \in P(\mathbb{F}_2).$$

So x is not in the image of ϕ (equivalently one cannot integrate x).

Lemma 12.10. *Let $\phi\colon V \longrightarrow W$ be a linear isomorphism.*
If $\psi\colon W \longrightarrow V$ is the inverse of ϕ then ψ is a linear map.

Proof. Suppose that w_1 and $w_2 \in W$. Let $v_i = \psi(w_i)$ so that $w_i = \phi(v_i)$, $i = 1$ and 2.

$$\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2) = w_1 + w_2,$$

as ϕ respects addition. As ψ is the inverse of ϕ it follows that

$$\psi(w_1 + w_2) = v_1 + v_2 = \psi(w_1) + \psi(w_2).$$

Thus ψ respects addition. Now suppose that $w \in W$ and $r \in F$. Let $v = \psi(v)$ so that $w = \phi(v)$.

$$\phi(rv) = r\phi(v) = rw,$$

as ϕ respects scalar multiplication. As ψ is the inverse of ϕ it follows that

$$\psi(rw) = rv = r\psi(w).$$

Thus ψ respects scalar multiplication. Hence ψ is linear. □

Lemma 12.11. *Let V and W be isomorphic vector spaces.*

If one of V and W is finite dimensional then they are both finite dimensional and they both have the same dimension.

Proof. By assumption there is a bijective linear map $\phi\colon V \longrightarrow W$. By (12.10) the situation is symmetric and we may suppose that V is finite dimensional. Let $v_1, v_2, \dots, v_n$ be a basis of V. It suffices to prove that $w_1, w_2, \dots, w_n$ is a basis of W, where $w_i = \phi(v_i)$.

Suppose that $w \in W$. As ϕ is surjective we may find $v \in V$ such that $\phi(v) = w$. As $v_1, v_2, \dots, v_n$ span V it follows that we may find scalars $r_1, r_2, \dots, r_n$ such that

$$v = r_1v_1 + r_2v_2 + \dots + r_nv_n.$$

But then

$$\begin{aligned} w &= \phi(v) \\ &= \phi(r_1v_1 + r_2v_2 + \dots + r_nv_n) \\ &= r_1\phi(v_1) + r_2\phi(v_2) + \dots + r_n\phi(v_n) \\ &= r_1w_1 + r_2w_2 + \dots + r_nw_n. \end{aligned}$$

It follows that $w_1, w_2, \dots, w_n$ span W. Now suppose that $r_1, r_2, \dots, r_n$ are scalars such that

$$0 = r_1w_1 + r_2w_2 + \dots + r_nw_n.$$

If we apply $\psi\colon W \longrightarrow V$, the inverse of ϕ, to both sides, we get

$$\begin{aligned} 0 &= \psi(r_1 w_1 + r_2 w_2 + \cdots + r_n w_n) \\ &= r_1\psi(w_1) + r_2\psi(w_2) + \cdots + r_n\psi(w_n) \\ &= r_1 v_1 + r_2 v_2 + \cdots + r_n v_n. \end{aligned}$$

As $v_1, v_2, \dots, v_n$ are independent it follows that $r_1, r_2, \dots, r_n = 0$. But then $w_1, w_2, \dots, w_n$ are independent. □

Theorem 12.12 (Rank-nullity, bis)**.** *Let $\phi\colon V \longrightarrow W$ be a linear map between two finite dimensional vector spaces.*

Then $\mathrm{rk}(\phi) + \nu(\phi) = \dim V$.

Proof. Let $v_1, v_2, \dots, v_n$ and $w_1, w_2, \dots, w_m$ be bases of V and W. By (12.6) this gives rise to two linear isomorphisms

$$f\colon V \longrightarrow F^n \qquad \text{and} \qquad g\colon W \longrightarrow F^m.$$

Let $\psi = g \circ \phi \circ f^{-1}\colon F^n \longrightarrow F^m$. As the composition of linear maps is linear and the inverse of a linear map is linear it follows that ψ is a linear map.

Claim 12.13. *The kernel of ψ is equal to the image of the kernel of ϕ, that is*

$$\mathrm{Ker}(\psi) = f(\mathrm{Ker}(\phi)).$$

Proof of (12.13)*.* Pick $u \in \mathrm{Ker}(\psi)$. Then

$$g \circ \phi \circ f^{-1}(u) = 0.$$

Let $v = f^{-1}(u) \in V$. Then $g(\phi(v)) = 0$. As g is injective, $\phi(v) = 0$ and so $v \in \mathrm{Ker}(\phi)$. But then $u = f(v) \in f(\mathrm{Ker}(\phi))$. Thus

$$\mathrm{Ker}(\psi) \subset f(\mathrm{Ker}(\phi)).$$

Now suppose that $u \in f(\mathrm{Ker}(\phi))$. Then $u = f(v)$ where $\phi(v) = 0$. Then

$$\psi(u) = g(\phi(f^{-1})(u)) = g(\phi(v)) = g(0) = 0.$$

This proves the reverse inclusion and establishes the claim. □

Therefore, restricting f to $\mathrm{Ker}(\phi)$, we get a surjective linear map

$$f_1\colon\ \mathrm{Ker}(\phi) \longrightarrow \mathrm{Ker}(\psi).$$

Since f_1 is injective, it is a linear isomorphism.

Claim 12.14. *The image of ψ is equal to the image of the kernel of ϕ, that is*

$$\mathrm{Im}(\psi) = g(\mathrm{Im}(\phi)).$$

Proof of (12.14). Pick $t \in \operatorname{Im}(\psi)$. Then there is a vector $u \in F^n$ such that $\psi(u) = t$. Let $v = f^{-1}(u) \in V$. Then

$$g(\phi(v)) = g(\phi(f^{-1}(u))) = \psi(u) = t.$$

Thus $t \in g(\operatorname{Im}(\phi))$. Hence

$$\operatorname{Im}(\psi) \subset g(\operatorname{Im}(\phi)).$$

Now suppose that $t \in g(\operatorname{Im}(\phi))$. Then there is a vector $v \in V$ such that $t = g(\phi(v))$. Let $u = f(v)$. Then

$$\psi(u) = g(\phi(f^{-1}(u))) = g(\phi(v)).$$

This proves the reverse inclusion and establishes the claim. $\square$

Therefore, restricting g to $\operatorname{Im}(\phi)$, we get a surjective linear map

$$g_1 \colon \operatorname{Im}(\phi) \longrightarrow \operatorname{Im}(\psi).$$

Since g_1 is injective, it is also a linear isomorphism.

(12.11) implies that

$$\nu(\phi) = \nu(\psi) \qquad \text{and} \qquad \operatorname{rk}(\phi) = \operatorname{rk}(\psi).$$

On the other hand, we know that ψ is given by a matrix A, that is $\psi(v) = Av$. It is almost immediate that $\operatorname{Ker}(\psi) = \operatorname{Ker}(A)$. Now the image of ψ is the same as the column space. But we know that the dimension of the column space of A is the same as the rank. Putting all of this together we have

$$\nu(\phi) = \nu(\psi) = \nu(A) \qquad \text{and} \qquad \operatorname{rk}(\phi) = \operatorname{rk}(\psi) = \operatorname{rk}(A).$$

Thus this result reduces to rank-nullity for A. $\square$