III. Linear transformations

III.A. Linear transformations without matrices

Given any algebraic structure ubiquitous in mathematics, we have a choice: either we can prove the same results repeatedly in different contexts; or we can abstractify the structure, name it, and prove the results once. For instance, applying this process to symmetries leads to *groups* in abstract algebra, which can then be classified.

Let's finally name the abstract object at the heart of linear algebra:

III.A.1. DEFINITION. A **linear transformation** is a map (function) between two vector spaces

$$T : V \to W$$

obeying linearity:

$$T(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha T\vec{v}_1 + \beta T\vec{v}_2$$

for all vectors $\vec{v}_1, \vec{v}_2 \in V$ and scalars α, β. We call V the *domain* and W the *codomain* or *range space*.[1] If $V = W$ then T is called an **endomorphism** of V.

III.A.2. EXAMPLE. Reflections, rotations, and projections associated with subspaces of V give a source of endomorphisms. Here are a few with $V = \mathbb{R}^2$:

- $T(a, b) = (-a, b)$ [reflection about the y-axis]
- $T(a, b) = (a\cos\theta - b\sin\theta, a\sin\theta + b\cos\theta)$ [rotation about $(0, 0)$]
- $T(a, b) = (0, b)$ [projection onto the y-axis]

III.A.3. EXAMPLE. Let V be the vector space of all continuous functions $f : \mathbb{R} \to \mathbb{R}$. (Note that this space is not finitely generated,

[1] Some texts (not this one) use "range" for what we shall call the *image* of T.

hence not finite-dimensional.) Then

$$f(x) \mapsto \int_0^x f(t)dt =: (Tf)(x)$$

is an endomorphism of V since $\int$ is linear: $\int(\alpha f_1 + \beta f_2) = \alpha \int f_1 + \beta \int f_2$. Alternatively, fixing $a, b \in \mathbb{R}$ we get a linear transformation $I : V \to \mathbb{R}$ by sending $f(x) \mapsto \int_a^b f(x)dx$.

III.A.4. THEOREM. *T is determined* uniquely *by where it sends an* (*ordered*) *basis. More precisely, let V and W be vector spaces, V finite-dimensional with basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$, and $\{\vec{w}_1, \ldots, \vec{w}_n\}$ any vectors in W. Then there is exactly one $T : V \to W$ with $T\vec{v}_i = \vec{w}_i$ $(\forall i)$.*

PROOF. Given $\vec{v} \in V$, we know from §II.E that there is a unique way of writing $\vec{v} = \sum \alpha_i \vec{v}_i$. So there's exactly one option for $T\vec{v}$, if we insist that T be linear:

$$T\vec{v} = T(\textstyle\sum_i \alpha_i \vec{v}_i) = \sum_i \alpha_i T\vec{v}_i = \sum_i \alpha_i \vec{w}_i.$$

That is, we "linearly extend" the assignment $\vec{v}_i \mapsto \vec{w}_i$ to all of V. $\square$

III.A.5. DEFINITION. The **image** of T consists of all the vectors in W that are "hit" by the transformation:

$$\mathrm{im}(T) = \{\, \vec{w} \in W \mid \vec{w} = T\vec{v} \text{ for some } \vec{v} \in V \}.$$

The **kernel** of T (or "null space") is the set of vectors in V that are "killed" by T:

$$\ker(T) = \{\, \vec{v} \in V \mid T\vec{v} = 0 \}.$$

As the reader will easily check, both $\mathrm{im}(T) \subseteq W$ and $\ker(T) \subseteq V$ are vector (sub)spaces. We shall call the dimension of $\mathrm{im}(T)$ the **rank** of T, and the dimension of $\ker(T)$ the **nullity** of T.

III.A.6. EXAMPLE. Consider $T : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R})$ (= polynomials of degree n over $\mathbb{R}$) defined by taking the derivative: $f = \sum a_j x^j \longmapsto \sum j a_j x^{j-1} = f'$. The kernel and image have dimensions 1 and n respectively. (We could also take the range space W to be $\mathcal{P}_{n-1}(\mathbb{R})$, in which case these dimensions don't change.)

We will prove the next theorem once abstractly (here) and once using matrices (later).

III.A.7. THEOREM (Rank+Nullity). *For any $T : V \to W$, with V finite-dimensional,*

$$\dim(\mathrm{im}(T)) + \dim(\ker(T)) = \dim(V).$$

That is, the dimension of the image is just the dimension of the domain less the dimension of the subspace (of the domain) killed by T.

PROOF. Assume $\dim(\ker(T)) = k$, $\dim(V) = n$. (Of course, $n \geq k$.) Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be any basis for $\ker(T)$; by choosing vectors not in their span we can complete it to a basis $\{\vec{v}_1, \ldots, \vec{v}_k; \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ for V.

Take any $\vec{w} \in \mathrm{im}(T)$ $[\subseteq W]$; then for some $\vec{v} \in V$,

$$\vec{w} = T\vec{v} = T\left(\textstyle\sum_{i=1}^n \alpha_i \vec{v}_i\right) = \textstyle\sum_{i=1}^n \alpha_i T\vec{v}_i \underset{\substack{\uparrow \\ T\vec{v}_i = 0,\ i \leq k}}{=} \textstyle\sum_{i=k+1}^n \alpha_i (T\vec{v}_i)$$

$\implies$ $\{T\vec{v}_i\}_{i=k+1}^n$ *span* $\mathrm{im}(T)$.

To show they are linearly independent, first notice that if

$$0 = \textstyle\sum_{i=k+1}^n a_i T\vec{v}_i = T\left(\textstyle\sum_{i=k+1}^n a_i \vec{v}_i\right)$$

then $\sum_{i=k+1}^n a_i \vec{v}_i \in \ker(T)$ by definition. Since $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *span* the kernel, $\sum_{i=k+1}^n a_i \vec{v}_i = \sum_{i=1}^k b_i \vec{v}_i$ (for some b_i); that is,

$$0 = \sum_{i=1}^n \left\{ \begin{array}{c} b_i\,,\ \ i \leq k \\ -a_i\,,\ \ i > k \end{array} \right\} \vec{v}_{i.}$$

Since the *entire* collection $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly independent, all the coefficients (a_i and b_i) must be 0. So all the a_i are zero, which establishes independence of $\{T\vec{v}_{k+1}, \ldots, T\vec{v}_n\}$. So they are a basis for $\mathrm{im}(T)$.

Since there are $n - k$ vectors in this basis, $\dim(\mathrm{im}(T)) = n - k$. Adding this to $\dim(\ker(T)) = k$ gives the theorem. $\square$

III.A.8. COROLLARY. $\dim(T(V)) \leq \dim(V)$.

Projections and reflections. Now we consider two types of endomorphisms which are based on having a "decomposition" of a vector space V into two subspaces:

III.A.9. DEFINITION. V is the **direct sum** of two of its subspaces W_1, W_2 if $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$. We write $V = W_1 \oplus W_2$.

The point is that given any vector $\vec{v}$ in such a V, we can write $\vec{v} = \vec{w}_1 + \vec{w}_2$ with $\vec{w}_i \in W_i$ $(i = 1, 2)$. Moreover, the $\vec{w}_i$ are unique: if

$$\vec{w}_1 + \vec{w}_2 = \vec{v} = \vec{w}_1' + \vec{w}_2'$$

are two such decompositions, we have $\vec{w}_1 - \vec{w}_1' = \vec{w}_2' - \vec{w}_2 \in W_1 \cap W_2 = \{0\} \implies \vec{w}_i = \vec{w}_i'$ $(i = 1, 2)$. Consequently, we may write elements of V as ordered pairs $(\vec{w}_1, \vec{w}_2)$.

By the *projection of V onto W_1 along W_2*, we will mean the endomorphism $\pi : V \to V$ sending $(\vec{w}_1, \vec{w}_2) \mapsto (\vec{w}_1, \vec{0})$. (This is not the "perpendicular" projection; it really does depend on W_2, and what's more, we have no notion of orthogonality on V without an inner product.) Similarly we can produce a *reflection* $\rho : V \to V$ of V about W_1 by sending[2] $(\vec{w}_1, \vec{w}_2) \mapsto (\vec{w}_1, -\vec{w}_2)$. Since $\vec{v} + \rho(\vec{v}) = 2\pi(\vec{v})$, these are closely related to to the projections.

Notice that the projection restricts to the identity on W_1; that is, $\pi(w_1, 0) = (w_1, 0)$. Thus the composition $\pi \circ \pi =: \pi^2$ (of a projection with itself) equals π. The reflection also restricts to the identity on W_1, but has $\rho^2 = Id_V$. It turns out that the converse holds in each case, so that these properties characterize projections and reflections:

III.A.10. PROPOSITION. *The projections are exactly the endomorphisms T of V with $T^2 = T$, and the reflections are those with $T^2 = Id_V$.*

PROOF. Suppose $T^2 = T$. Put $W_2 := \ker(T)$ and $W_1 := \mathrm{im}(T)$. Given any $\vec{v} \in V$, write $\vec{w}_1 := T(\vec{v})$ and $\vec{w}_2 := \vec{v} - T(\vec{v})$, so that $\vec{v} = \vec{w}_1 + \vec{w}_2$. Clearly $\vec{w}_1 \in W_1$, while $T(\vec{w}_2) = T(\vec{v}) - T(T(\vec{v})) = T(\vec{v}) - T^2(\vec{v}) = T(\vec{v}) - T(\vec{v}) = \vec{0} \implies \vec{w}_2 \in W_2$. So $W_1 + W_2 = V$.

[2]One should assume here that our vector space is over a field F of characteristic different from 2 (since in that case $-1 = 1$ and the "reflection" is the identity).

Now if $\vec{v} \in W_1 \cap W_2$, then $\vec{v} = T(\vec{v}')$ for some $\vec{v}'$ (as $\vec{v} \in \mathrm{im}(T)$) and $T(\vec{v}) = \vec{0}$ (as $\vec{v} \in \ker(T)$). So $\vec{v} = T(\vec{v}') = T^2(\vec{v}') = T(T(\vec{v}')) = T(\vec{v}) = \vec{0}$, which shows $W_1 \cap W_2 = \{0\}$. Hence $V = W_1 \oplus W_2$ and T is precisely the projection π constructed above.

On the other hand, if $T^2 = \mathbb{I}$, then we set $S = \frac{1}{2}(T + \mathbb{I})$ and compute that $S^2 = S$, so that S is a projection and $T = 2S - \mathbb{I}$ sends $(\vec{w}_1, \vec{w}_2) \mapsto (2\vec{w}_1, 0) - (\vec{w}_1, \vec{w}_2) = (\vec{w}_1, -\vec{w}_2)$. □

Since we can "multiply" (compose) and add endomorphisms, they form a ring $End(V)$. The elements of a ring which square to themselves are called *idempotents*.

III.A.11. REMARK. Basic evil exam question: "Given an endomorphism $T : V \to V$, we have $\ker(T) \cap \mathrm{im}(T) = \{0\}$. True or False?"

Certainly, this is true for the projections we've just described. More generally, we might reason as follows if V is finite-dimensional: by "rank + nullity", we know that the sum of the dimensions of the kernel and the image is indeed the dimension of V. So maybe V is just their "vector space sum"?

But this turns out to be totally false! There are examples where the kernel and the image are exactly the same! (See Exercise (2) below.) What *is* true is that *if* $\ker(T) \cap \mathrm{im}(T) = \{0\}$, then $V = \ker(T) \oplus \mathrm{im}(T)$ (see Exercise (3)).

Isomorphisms (invertible transformations). Now notice that

$$\ker(T) = 0 \implies T \text{ is 1-to-1 (or "injective")},$$

because when the kernel is trivial

$$T\vec{v}_1 = T\vec{v}_2 \implies T(\vec{v}_1 - \vec{v}_2) = 0 \implies \vec{v}_1 - \vec{v}_2 \in ker(T)$$
$$\implies \vec{v}_1 - \vec{v}_2 = 0 \implies \vec{v}_1 = \vec{v}_2.$$

On the other hand, if $\mathrm{im}(T) = W$ we say that T is "*onto* W" (or "surjective").

III.A.12. DEFINITION. A 1-to-1 and onto transformation is called an **isomorphism**.

If there exists an isomorphism T between two vector spaces V and W, we call them *isomorphic* and write $V \cong W$. This indicates that they are essentially the same vector space — differing perhaps only in presentation. For example, the transformation

$$T : \mathbb{R}^3 \longrightarrow P_2 \, (= \text{polynomials of degree } \leq 2)$$

given by "linearly extending" (to all of $\mathbb{R}^3$) the assignments

$$\hat{e}_1 \mapsto 1\,, \quad \hat{e}_2 \mapsto x\,, \quad \hat{e}_3 \mapsto x^2$$

of the standard basis vectors (as in Theorem III.A.4), is an isomorphism.

I said an isomorphism indicates "equivalence" of vector spaces: in that case one would expect the following statement holds!

III.A.13. PROPOSITION. *Let V and W be finite dimensional vector spaces over some field F. Then $V \cong W \implies \dim_F(V) = \dim_F(W)$.*

PROOF. $V \cong W$ means that there is a map $T : V \to W$ with $\ker(T) = 0$, $\operatorname{im}(T) = W$. Applying rank + nullity to T, $\dim(V) = \dim(\operatorname{im}(T)) + 0 = \dim(W)$. □

III.A.14. REMARK. As you will see in Exercise (1), the converse of Proposition III.A.13 also holds. So for example, $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic as vector spaces over $\mathbb{R}$, as are $\mathbb{R}^3$ and $\mathcal{P}_2(\mathbb{R})$.

We may compose linear transformations (just as we would compose any functions): e.g., the composite of

$$R\colon U \to V \quad \text{and} \quad T\colon V \to W$$

is written $(T \circ R)\colon U \to W$, or simply TR. This is linear because

$$T(R(\alpha\vec{u} + \beta\vec{v})) = T\left(\alpha R(\vec{u}) + \beta R(\vec{v})\right) = \alpha T(R(\vec{u})) + \beta T(R(\vec{v})).$$

III.A.15. DEFINITION. Given $T : V \to W$, if there exists $S : W \to V$ such that

$$T \circ S = Id_W\,, \quad S \circ T = Id_V\,,$$

then T is *invertible*, with $T^{-1} = S$.

Now

(i) $T \circ S = Id_W \implies T$ is onto W (clear)
(ii) $S \circ T = Id_V \implies \ker(T) = 0$.
[To see (ii), observe that if $T(\vec{v}) = 0$, then

$$\vec{v} = Id_V(\vec{v}) = (S \circ T)(\vec{v}) = S(T(\vec{v})) = \vec{0}.$$

Alternatively, if the spaces are finite-dimensional, you can use Corollary III.A.8: we have

$$\dim(V) = \dim(S(T(\vec{v}))) \leq \dim(T(V)) \leq \dim(V)$$

hence $\dim(T(V)) = \dim(V)$. Therefore by "rank + nullity", $\dim(\ker(T)) = 0$.]

What (i) and (ii) show is that T invertible $\implies T$ is an isomorphism. In proving (ii) we have also seen the following useful

III.A.16. PROPOSITION. *The range and domain of an invertible linear transformation have the same dimension. That is, $T : V \to W$ invertible $\implies \dim(V) = \dim(W)$.*

In any case, the converse (T isomorphism $\implies T$ invertible) is also true: given $T : V \xrightarrow{\cong} W$, you can show T takes a basis $\{\vec{v}_i\}$ to a basis $\{\vec{w}_i\}$, and use Theorem III.A.4 to construct the (2-sided) inverse by sending $\vec{w}_i \mapsto \vec{v}_i$. (Try it!) For a less constructive approach (but one that works for spaces that aren't finite-dimensional), just define $S : W \to V$ as follows: given $\vec{w} \in W$, there exists a vector $\vec{v}_{\vec{w}} \in V$ with $T(\vec{v}_{\vec{w}}) = \vec{w}$ (since T is onto), and this vector is unique (since T is 1-1). Set $S(\vec{w}) := \vec{v}_{\vec{w}}$, and observe that $S \circ T$ and $T \circ S$ are the identity.

Lest the reader begin to form an opinion that isomorphisms are trivial identifications, or even embrace the heresy that algebra deals in mere equalities so says nothing of interest to an analyst or geometer, we submit the following application which uses an isomorphism to do something rather nontrivial in a single page:

III.A.17. EXAMPLE (Lagrange Interpolation). Let $c_0, \ldots, c_n \in F$ be distinct elements, and consider the map

$$T : \mathcal{P}_n(F) \to F^{n+1}$$

given by[3] $T(g(x)) := {}^t(g(c_0), \ldots, g(c_n))$. (It evaluates a given function at each of the c_i.) We will show that this is an isomorphism by constructing an inverse

$$L : F^{n+1} \to \mathcal{P}_n(F).$$

First set, for $i = 0, \ldots, n$,

$$f_i(x) := \prod_{k \neq i} \frac{x - c_k}{c_i - c_k}$$

where the product is over $k = 0, \ldots, n$ omitting i; notice that $f_i(c_j) = \delta_{ij}$ (Kronecker delta). Now define

$$L\left({}^t(b_0, \ldots, b_n)\right) := \sum_{i=0}^{n} b_i f_i(x). \tag{III.A.18}$$

Then

$$\begin{aligned}(T \circ L)\left({}^t(b_0, \ldots, b_n)\right) &= T\left(\textstyle\sum_{i=0}^n b_i f_i(x)\right) = \textstyle\sum_{i=0}^n b_i T(f_i(x)) \\ = \textstyle\sum_{i=0}^n b_i {}^t(f_i(c_0), \ldots, f_i(c_n)) &= {}^t\left(\textstyle\Sigma_{i=0}^n b_i f_i(c_0), \ldots, \Sigma_{i=0}^n b_i f_i(c_n)\right) \\ &= {}^t\left(\textstyle\Sigma_{i=0}^n b_i \delta_{i0}, \ldots, \Sigma_{i=0}^n b_i \delta_{in}\right) = {}^t(b_0, \ldots, b_n)\end{aligned}$$

shows that $T \circ L = Id_{F^{n+1}}$. In particular T is onto, and so by "rank + nullity" it must also be 1-to-1 (as our spaces have the same dimension), hence an isomorphism.

Of course, it was obvious that $\mathcal{P}_n(F)$ and F^{n+1} were isomorphic, because they have the same dimension. But that means there *exists* an isomorphism. To say that a particular transformation betweeen them is an isomorphism is a more precise statement, and in this case useful: it says that you can find a polynomial whose graph interpolates the points $\{(c_i, b_i)\}_{i=0}^n$, which is just what (III.A.18) provides.

[3]We write column vectors as transposes of row vectors to fit them on the page more neatly. $\mathcal{P}_n(F)$ means polynomials of degree $\leq n$ with coefficients in a field F.

Further remarks. We want to emphasize the following important point: invertibility of a transformation $T : V \to W$ means that it has an inverse *with domain* W. For this to be possible, the image of T must be the whole of W. If T is just 1-to-1 (but not onto all of W), it is *invertible on its image*, i.e. as a map from $V \to \mathrm{im}(T)$, not as a map from $V \to W$. But you say, "how can the same T be both invertible and not invertible?"

The point here is that a "linear transformation" (the object we are saying is invertible, or not) is really *all three objects*: T, V, and W. This isn't so "out there": recall from calculus that differentiability of a "given function" can depend on our choice of domain. As a function on $(0, \infty)$, $f(x) = |x|$ is differentiable. As a function on $(-1, 1)$, it is certainly not.

Finally, the attentive reader will have noticed that our Definition III.A.1 was rather imprecise. Our vector spaces are over a field F, and α, β come from that field. If there is any doubt what field we are thinking of V and W as being over, one says that T is an "F-linear transformation". Clearly $T(\vec{u} + \vec{v}) = T\vec{u} + T\vec{v}$ has nothing to do with the choice of field, so what's at stake is whether we have $T(\alpha\vec{u}) = \alpha T(\vec{u})$ for all $\alpha \in F$.

For instance, if $V = W = \mathbb{C}$, then $T(x + \mathbf{i}y) := x - \mathbf{i}y$ and $\tilde{T}(x + \mathbf{i}y) := x + 2y + \mathbf{i}y$ are $\mathbb{R}$-linear but not $\mathbb{C}$-linear: for $\alpha \in \mathbb{C}$, $T(\alpha(x + \mathbf{i}y)) = \bar{\alpha}T(x + \mathbf{i}y)$ etc. On the other hand, $S(x + \mathbf{i}y) := -y + \mathbf{i}x$, being (after all) "multiplication by $\mathbf{i}$", is $\mathbb{C}$-linear.

Exercises

(1) Prove the converse of Proposition III.A.13: given two vector spaces V, W over the same field and of the same dimension (over that field), show they are isomorphic. Explain why this means that all n-dimensional vector spaces over $\mathbb{R}$ are isomorphic to $\mathbb{R}^n$!!

(2) Let V be an n-dimensional vector space over the field F and let T be a linear transformation from V into V such that the range and kernel of T are the same subspace of V.

(a) Show that n has to be even.

(b) Give an example of such a transformation.

(3) Given an endomorphism $T\colon V \to V$ of a finite-dimensional vector space, with $\ker(T) \cap \operatorname{im}(T) = \{0\}$, show that $V = \ker(T) \oplus \operatorname{im}(T)$. [Hint: use the result of Exercise II.C.3.]

(4) Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3).$$

Is T invertible? If so, find an explicit description for T^{-1} like the one which defines T.

(5) Find two linear operators T and U on $\mathbb{R}^2$ such that $T \circ U = 0$ but $U \circ T \neq 0$.

(6) State why each of the following functions $T : \mathbb{R}^2 \to \mathbb{R}^2$ is *not* linear:
(a) $T(x_1, x_2) = (1, x_2)$;
(b) $T(x_1, x_2) = (\sin x_1, 0)$;
(c) $T(x_1, x_2) = (x_1 + 1, x_2)$.

(7) Compute the rank and nullity of the linear transformation

$$\operatorname{tr}\colon M_{n\times n}(F) \to F$$

given by sending a matrix to its *trace* $\operatorname{tr}(A) = \sum A_{ii}$. Is it one-to-one? onto?

(8) Let $T : V \to V$ be an endomorphism of an n-dimensional real vector space, with $T^2 = T$. Describe $\mathbb{I} - T$ and $(\mathbb{I} + T)^{-1}$, and say what their ranks are (in terms of dimensions of some subspaces).

(9) Let $T\colon V \to W$ and $S\colon U \to V$ be linear transformations, with U, V, W finite-dimensional. Show the following:
(a) $\dim(\ker(S)) \leq \dim(\ker(T \circ S))$
and $\dim(\operatorname{im}(T \circ S)) \leq \dim(\operatorname{im}(T))$.
(b) $\dim(\ker(T \circ S)) \leq \dim(\ker(T)) + \dim(\ker(S))$. [Hint: start with a basis of $\ker(S)$, then extend to a basis of $\ker(T \circ S)$.]
(c) If S is surjective, then we have equality in (b).