(III.A) Linear transformations without matrices

Linear algebra is, in a nutshell, the study of the objects in the following

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

\[ T : V \rightarrow W \]

obeying linearity: \( T(\alpha \vec{v}_1 + \beta \vec{v}_2) = \alpha T \vec{v}_1 + \beta T \vec{v}_2 \). If \( V = W \) then \( T \) is called an *endomorphism* of \( V \).

**Example 2.** 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]
- \( T(a, b) = (a \cos \theta - b \sin \theta, a \sin \theta + b \cos \theta) \) [rotation]
- \( T(a, b) = (0, b) \) [projection]

**Example 3.** Let \( V \) be the vector space of all continuous functions \( f : \mathbb{R} \rightarrow \mathbb{R} \). (Note that this space is not finitely generated, 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 \( I : V \rightarrow \mathbb{R} \) by sending \( f(x) \mapsto \int_a^b f(x)dx \).
THEOREM 4. \( T \) is determined uniquely by where it sends an (ordered) basis. More precisely, let \( V \) and \( W \) be vector spaces over \( \mathbb{R} \), \( V \) finite dimensional (\( \dim n \)), \( \{ \vec{v}_1, \ldots, \vec{v}_n \} \) a basis for \( V \), \( \{ \vec{w}_1, \ldots, \vec{w}_n \} \) 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:
\[
T \vec{v} = T(\sum \alpha_i \vec{v}_i) = \sum \alpha_i T \vec{v}_i = \sum \alpha_i \vec{w}_i.
\]
That is, we “linearly extend” the assignment \( \vec{v}_i \mapsto \vec{w}_i \) to all of \( V \). \( \square \)

DEFINITION. The image of \( T \) consists of all the vectors in \( W \) that are “hit” by the transformation:
\[
\text{im}(T) = \{ \vec{w} \in W \mid \vec{w} = T \vec{v} \text{ for some } \vec{v} \in V \}.
\]
The kernel of \( T \) (or “null space”) is the set of vectors in \( V \) “killed”\(^1\) by \( T \):
\[
\ker(T) = \{ \vec{v} \in V \mid T \vec{v} = 0 \}.
\]
Both are vector (sub)spaces: \( \text{im}(T) \subseteq W, \ker(T) \subseteq V \). We shall call the dimension of \( \text{im}(T) \) the rank of \( T \), and the dimension of \( \ker(T) \) the nullity of \( T \).

EXAMPLE 5. Consider \( T : P_n(\mathbb{R}) \to P_n(\mathbb{R}) \) (polynomials of degree \( n \)) defined by taking the derivative:
\[
f = \sum a_j x^j \mapsto \sum j a_j x^{j-1} = f'.
\]
The kernel and image have dimensions 1 and \( n \) respectively. (We could also take the range space \( W \) to be \( P_{n-1}(\mathbb{R}) \), in which case these dimensions don’t change.)

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

THEOREM 6 (Rank+Nullity). For any \( T : V \to W \), with \( V \) finite-dimensional,
\[
\dim(\text{im}(T)) + \dim(\ker(T)) = \dim(V).
\]
\(^1\)apparently, algebraists are rather bloodthirsty individuals
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 \( \vec{w} \in \text{im}(T) \subseteq W \); then for some \( \vec{v} \in V \),

\[
\vec{w} = T\vec{v} = T \left( \sum_{i=1}^{n} a_i \vec{v}_i \right) = \sum_{i=1}^{n} a_i T\vec{v}_i = \sum_{i=k+1}^{n} a_i (T\vec{v}_i) \uparrow \\
T\vec{v}_i = 0, \ i \leq k
\]

\[\Rightarrow \ {\{ T\vec{v}_i \}_{i=k+1}^{n}} \text{ span } \text{im}(T) .\]

To show they are linearly independent, first notice that if

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

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,

\[
0 = \sum_{i=1}^{n} \left\{ \begin{array}{ll}
 b_i, & i \leq k \\
 -a_i, & i > k 
\end{array} \right\} \vec{v}_i .
\]

Since \( \{ \vec{v}_1, \ldots, \vec{v}_n \} \) are all linearly independent, all the coefficients \((a_i \text{ and } b_i)\) 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 \( \text{im}(T) \).

Since there are \( n - k \) of them, \( \dim(\text{im}(T)) = n - k \). Adding this to \( \dim(\ker(T)) = k \) gives the theorem. \( \square \)

COROLLARY 7. \( \dim(T(V)) \leq \dim(V) \).

EXAMPLE 8. Basic evil exam question: “Given \( T : V \rightarrow V \), we have \( \ker(T) \cap \text{im}(T) = \{0\} \). True or False?”

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 the exercises
below.)

**Projections and reflections.** Now we consider two types of en-
domorphisms which are based on having a “decomposition” of a
vector space $V$ into two subspaces:

**Definition 9.** $V$ 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
$\vec{v} = w_1 + w_2$ with $w_i \in W_i$ ($i = 1, 2$). Moreover, the $w_i$ are unique: if
$$w_1 + w_2 = \vec{v} = \vec{v}_1 + \vec{v}_2$$
are two such decompositions, we have $w_1 - w'_1 = w_2 - w'_2 \in W_1 \cap
W_2 = \{0\} \implies w_i = w'_i$ ($i = 1, 2$). Consequently, we may write
elements of $V$ as ordered pairs $(w_1, w_2)$.

By the projection of $V$ onto $W_1$ along $W_2$, we will mean the endo-
morphism $\pi : V \to V$ sending $(w_1, w_2) \mapsto (w_1, 0)$. (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
$$\vec{v} \mapsto \vec{v} - 2\pi(\vec{v})$$
these are closely related to to the projections.

**Proposition 10.** The projections are exactly the endomorphisms $T$ of
$V$ with $T^2 = T$, and the reflections are those with $T^2 = \text{Id}_V$.

**proof.** Suppose $T^2 = T$. Put $W_1 := \ker(T)$ and $W_1 := \text{im}(T)$.
Given any $\vec{v} \in V$, write $w_1 := T(\vec{v})$ and $\vec{w}_2 := \vec{v} - T(\vec{v})$. Clearly
$w_1 \in W_1$, while $T(\vec{w}_2) = T(\vec{v}) - T(T(\vec{v})) = T(\vec{v}) - T^2(\vec{v}) = T(\vec{v}) -
T(\vec{v}) = 0 \implies \vec{w}_2 \in W_2$. So $W_1 + W_2 = V$. Now if $\vec{v} \in W_1 \cap W_2$,
then $\vec{v} = T(\vec{v}')$ for some $\vec{v}'$ (as $\vec{v} \in \text{im}(T)$) and $T(\vec{v}) = 0$ (as $\vec{v} \in
\ker(T)$). So $\vec{v} = T(\vec{v}') = T^2(\vec{v}') = T(T(\vec{v}')) = T(\vec{v}) = 0$, which

\footnote{one should assume here that $F$ has characteristic different from 2}
shows $W_1 \cap W_2 = V$. Hence $V = W_1 \oplus W_2$ and $T$ is precisely the $\pi$ constructed above.

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

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.

**Isomorphisms (invertible transformations).** Now notice that

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

(or “injective”), because then

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

$$\implies \bar{v}_1 - \bar{v}_2 = 0$$

or $\bar{v}_1 = \bar{v}_2$. On the other hand, if $im(T) = W$ we say that $T$ is “onto $W$” (or “surjective”). A one-to-one and onto transformation is called an isomorphism. Given $V$, $W$ if there exists an isomorphism $T$ between them, we call $V$ and $W$ 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 1), is an isomorphism.

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

**Proposition 11.** 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 there’s a map $T : V \rightarrow W$ with $\ker(T) = 0$, $\text{im}(T) = W$. Applying rank+nullity to $T$, $\dim(V) = \dim(\text{im}(T)) + 0 = \dim(W)$. \hfill \square

REMARK 12. As you will see in the exercises, the converse of Proposition 11 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 $\mathbb{P}_2(\mathbb{R})$.

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

\[ U \stackrel{R}{\longrightarrow} V \stackrel{T}{\longrightarrow} W \]

is written $T \circ R$ (or $TR$). This is linear because

\[ T(R(\alpha \vec{u} + \beta \vec{v})) = T(\alpha R(\vec{u}) + \beta R(\vec{v})) = \alpha T(R(\vec{u})) + \beta T(R(\vec{v})). \]

DEFINITION 13. Given $T : V \rightarrow W$, if there exists $S : W \rightarrow V$ such that

\[ T \circ S = \text{Id}_W, \quad S \circ T = \text{Id}_V, \]

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

Now

(i) $T \circ S = \text{Id}_W \implies T$ is onto $W$ (clear)
(ii) $S \circ T = \text{Id}_V \implies \ker(T) = 0$.

[To see this, observe that if $T(\vec{v}) = 0$, then $\vec{v} = \text{Id}_V(\vec{v}) = (S \circ T)(\vec{v}) = S(T(\vec{v})) = \vec{0}$. Alternatively, if the spaces are finite-dimensional, you can use Corollary 7: 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

PROPOSITION 14. The range and domain of an invertible linear transformation have the same dimension. That is, $T : V \rightarrow W$ invertible $\implies \dim(V) = \dim(W)$. 
In any case, the converse \((T\text{ isomorphism } \implies T\text{ invertible})\) is also true: given \(T : V \cong W\), you can show \(T\) takes a basis \(\{\vec{v}_i\}\) to a basis \(\{\vec{w}_i\}\), and use Theorem 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.

**Example 15 (Lagrange Interpolation).** Let \(c_0, \ldots, c_n \in F\) be distinct elements, and consider the map
\[
T : P_n(F) \to F^{n+1}
\]
given by \(T(g(x)) := \{g(c_0), \ldots, g(c_n)\}\). (It evaluates a given function at each of the \(c_i\).) We will show that this is an isomorphism by constructing an inverse
\[
L : F^{n+1} \to P_n(F).
\]
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 \(f_i(c_j) = \delta_{ij}\). Now define
\[
L(\{b_0, \ldots, b_n\}) := \sum_{i=0}^{n} b_i f_i(x).
\]
Then
\[
(T \circ L)(\{b_0, \ldots, b_n\}) = T \left( \sum_{i=0}^{n} b_i f_i(x) \right) = \sum_{i=0}^{n} b_i T(f_i(x))
\]
\[
= \sum_{i=0}^{n} b_i f_i(c_0), \ldots, f_i(c_n) = \{ \sum_{i=0}^{n} b_i f_i(c_0), \ldots, \sum_{i=0}^{n} b_i f_i(c_n) \}
\]
\[\delta_{ij}\] is the so-called “Kronecker delta”: 0 if \(i \neq j\) and 1 if \(i = j\).
\[ (\sum_{i=0}^{n} b_i \delta_{i0}, \ldots, \sum_{i=0}^{n} b_i \delta_{in}) = (b_0, \ldots, b_n) \]

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

Of course, it was obvious that \( P_n(F) \) 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 between 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 \( \{ (c_i, b_i) \}_{i=0}^{n} \).

Further remarks. We want to emphasize the following important point: for a transformation \( T : V \to W \) to be invertible means invertible on all of the given range space \( W \), which means all of \( W \) is going to have to be in the image. 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 \( V \to \text{im}(T) \), but 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, depending on its domain, \( f \) can be differentiable (=differentiable at every point in its domain), or not. As a function on \((0, \infty), f(x) = |x|\) is differentiable. As a function on \((-1, 1), \) it is not.

Finally, the attentive reader will have noticed that our Definition 1 was rather imprecise. Our vector spaces are over a field \( F \), and \( \alpha, \beta \) 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 \( T(\bar{u} + \bar{v}) = T\bar{u} + T\bar{v} \) has nothing to do with the choice of field, so what’s at stake is whether we have \( T(\alpha \bar{u}) = \alpha T(\bar{u}) \) for all \( \alpha \in F \).

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

**Exercises**

(1) Prove the converse of Proposition 11: 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 all \( n \) -dimensional vector spaces over \( \mathbb{R} \) are \( \cong \mathbb{R}^n \) !

(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 identical. (a) Show that \( n \) has to be even. (b) Give an example of such a transformation.

(3) 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 a rule for \( T^{-1} \) like the one which defines \( T \).

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

(5) 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) \( T(x_1, x_2) = (x_1 + 1, x_2) \).

(6) Compute the rank and nullity of the linear transformation \( \text{tr} : M_{n \times n}(F) \to F \) defined by \( \text{tr}(A) = \sum A_{ii} \). Is it one-to-one? onto?

(7) 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).