23. Kernel, Rank, Range

We now study linear transformations in more detail. First, we establish some important vocabulary.

The *range* of a linear transformation $f : V \to W$ is the set of vectors the linear transformation maps to. This set is also often called the *image* of f, written

$$\operatorname{ran}(f) = \operatorname{Im}(f) = L(V) = \{L(v) | v \in V\} \subset W.$$

The *domain* of a linear transformation is often called the *pre-image* of f. We can also talk about the pre-image of any subset of vectors $U \in W$:

$$L^{-1}(U) = \{v \in V | L(v) \in U\} \subset V.$$

A linear transformation f is *one-to-one* if for any $x \neq y \in V$, $f(x) \neq f(y)$. In other words, different vector in V always map to different vectors in W. One-to-one transformations are also known as *injective* transformations. Notice that injectivity is a condition on the pre-image of f.

A linear transformation f is *onto* if for every $w \in W$, there exists an $x \in V$ such that $f(x) = w$. In other words, every vector in W is the image of some vector in V. An onto transformation is also known as an *surjective* transformation. Notice that surjectivity is a condition on the image of f. [1]

Suppose $L : V \to W$ is *not* injective. Then we can find $v_1 \neq v_2$ such that $Lv_1 = Lv_2$. Then $v_1 - v_2 \neq 0$, but

$$L(v_1 - v_2) = 0.$$

Definition Let $L : V \to W$ be a linear transformation. The set of all vectors v such that $Lv = 0_W$ is called the *kernel of* L:

$$\ker L = \{v \in V | Lv = 0\}.$$

[1] The notions of one-to-one and onto can be generalized to arbitrary functions on sets. For example if g is a function from a set S to a set T, then g is one-to-one if different objects in S always map to different objects in T. For a linear transformation f, these sets S and T are then just vector spaces, and we require that f is a linear map; *i.e.* f respects the linear structure of the vector spaces.

The linear structure of sets of vectors lets us say much more about one-to-one and onto functions than one can say about functions on general sets. For example, we always know that a linear function sends 0_V to 0_W. Then we can show that a linear transformation is one-to-one if and only if 0_V is the only vector that is sent to 0_W: by looking at just one (very special) vector, we can figure out whether f is one-to-one. For arbitrary functions between arbitrary sets, things aren't nearly so convenient!

Theorem. *A linear transformation L is injective if and only if* $\ker L = \{0_V\}$.

The proof of this theorem is an exercise.

Notice that if L has matrix M in some basis, then finding the kernel of L is equivalent to solving the homogeneous system

$$MX = 0.$$

Example Let $L(x, y) = (x + y, x + 2y, y)$. Is L one-to-one?

To see, we can solve the linear system:

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

Then all solutions of $MX = 0$ are of the form $x = y = 0$. In other words, $\ker L = 0$, and so L is injective.

Theorem. *Let $L : V \to W$. Then* $\ker L$ *is a subspace of V.*

Proof. Notice that if $L(v) = 0$ and $L(u) = 0$, then for any constants c, d, $L(cu+dv) = 0$. Then by the subspace theorem, the kernel of L is a subspace of V. □

This theorem has a nice interpretation in terms of the eigenspaces of L. Suppose L has a zero eigenvalue. Then the associated eigenspace consists of all vectors v such that $Lv = 0v = 0$; in other words, the 0-eigenspace of L is exactly the kernel of L.

Returning to the previous example, let $L(x, y) = (x + y, x + 2y, y)$. L is clearly not surjective, since L sends $\mathbb{R}^2$ to a plane in $\mathbb{R}^3$.

Notice that if $x = L(v)$ and $y = L(u)$, then for any constants c, d then $cx + dy = L(cx + du)$. Then the subspace theorem strikes again, and we have the following theorem.

Theorem. *Let $L : V \to W$. Then the image $L(V)$ is a subspace of W.*

To find a basis of the the image of L, we can start with a basis $S = \{v_1, \ldots, v_n\}$ for V, and conclude that

$$L(V) = \operatorname{span} L(S) = \operatorname{span}\{v_1, \ldots, v_n\}.$$

However, the set $\{v_1, \ldots, v_n\}$ may not be linearly independent, so we solve

$$c^1 L(v_1) + \ldots + c^n L(v_n) = 0.$$

By finding relations amongst $L(S)$, we can discard vectors until a basis is arrived at. The size of this basis is the dimension of the image of L, which is known as the *rank* of L.

Definition The *rank* of a linear transformation L is the dimension of its image, written $\operatorname{rank} L$.

The *nullity* of a linear transformation is the dimension of the kernel, written L.

Theorem (Dimension Formula). *Let $L : V \to W$ be a linear transformation, with V a finite-dimensional vector space*[2]*. Then:*

$$\begin{aligned} \dim V &= \dim \ker V + \dim L(V) \\ &= L + \operatorname{rank} L. \end{aligned}$$

Proof. Pick a basis for V:

$$\{v_1, \ldots, v_p, u_1, \ldots, u_q\},$$

where $v_1, \ldots, v_p$ is also a basis for $\ker L$. This can always be done, for example, by finding a basis for the 0-eigenspace of L. Then $p = \dim \ker L$ and $p + q = \dim V$. Then we need to show that $q = \operatorname{rank} L$. To accomplish this, we show that $\{L(u_1), \ldots, L(u_q)\}$ is a basis for $L(V)$.

To see that $\{L(u_1), \ldots, L(u_q)\}$ spans $L(V)$, consider any vector w in $L(V)$ and find constants c^i, d^j such that:

$$\begin{aligned} w &= L(c^1 v_1 + \ldots + c^p v_p + d^1 u_1 + \ldots + d^q u_q) \\ &= c^1 L(v_1) + \ldots + c^p L(v_p) + d^1 L(u_1) + \ldots + d^q L(u_q) \\ &= d^1 L(u_1) + \ldots + d^q L(u_q) \text{ since } L(v_i) = 0, \end{aligned}$$

$\Rightarrow L(V) = \operatorname{span}\{L(u_1), \ldots, L(u_q)\}$.

Now we show that $\{L(u_1), \ldots, L(u_q)\}$ is linearly independent. We argue by contradiction: Suppose there exist constants c^j (not all zero) such that

$$\begin{aligned} 0 &= d^1 L(u_1) + \ldots + d^q L(u_q) \\ &= L(d^1 u_1 + \ldots + d^q u_q). \end{aligned}$$

[2]The formula still makes sense for infinite dimensional vector spaces, such as the space of all polynomials, but the notion of a basis for an infinite dimensional space is more sticky than in the finite-dimensional case. Furthermore, the dimension formula for infinite dimensional vector spaces isn't useful for computing the rank of a linear transformation, since an equation like $\infty = 3 + x$ cannot be solved for x. As such, the proof presented assumes a finite basis for V.

But since the u^j are linearly independent, then $d^1 u_1 + \ldots + d^q u_q \neq 0$, and so $d^1 u_1 + \ldots + d^q u_q$ is in the kernel of L. But then $d^1 u_1 + \ldots + d^q u_q$ must be in the span of $\{v_1, \ldots, v_p\}$, since this was a basis for the kernel. This contradicts the assumption that $\{v_1, \ldots, v_p, u_1, \ldots, u_q\}$ was a basis for V, so we are done. $\square$

References

- Hefferon, Chapter Three, Section II.2: Rangespace and Nullspace (Recall that 'homomorphism' is is used instead of 'linear transformation' in Hefferon.)

Wikipedia:

- Rank
- Dimension Theorem
- Kernel of a Linear Operator

Review Questions

1. Let $L : V \to W$ be a linear transformation. Prove that $\ker L = \{0_V\}$ if and only if L is one-to-one.

2. Let $\{v_1, \ldots, v_n\}$ be a basis for V. Explain why

$$L(V) = \mathrm{span}\{L(v_1), \ldots, L(v_n)\}.$$

3. Suppose $L : \mathbb{R}^4 \to \mathbb{R}^4$ whose matrix M in the standard basis is row equivalent to the folowing matrix:

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Explain why the first three columns of of the original matrix M form a basis for $L(V)$.
Find and describe and algorithm (*i.e.* a general procedure) for finding a basis for $L(V)$ when $L : \mathbb{R}^n \to \mathbb{R}^m$.
Finally, provide an example of the use of your algorithm.

4. Claim: If $\{v_1, \ldots, v_n\}$ is a basis for $\ker L$, where $L : V \to W$, then it is always possible to extend this set to a basis for V.

   Choose a simple yet non-trivial linear transformation with a non-trivial kernel and verify the above claim for the transformation you choose.

5. Let $P_n(x)$ be the space of polynomials in x of degree less than or equal to n, and consider the derivative operator $\frac{d}{dx}$. Find the dimension of the kernel and image of $\frac{d}{dx}$.

   Now, consider $P_2(x, y)$, the space of degree two polynomials in x and y. (Recall that xy is degree two, and x^2y is degree three, for example.) Let $L = \frac{d}{dx} + \frac{d}{dy}$. (For example, $L(xy) = \frac{d}{dx}(xy) + \frac{d}{dy}(xy) = y + x$.) Find a basis for the kernel of L. Verify the dimension formula in this case.