Math 201 lecture for Friday, Week 5

Rank-nullity theorem; Isomorphisms.

Let $f\colon V \to W$ be a linear mapping between vectors spaces V and W over a field F. Recall the definitions from last time:

Definition. The *kernel* or *null space* of f is

$$\mathcal{N}(f) := \ker(f) := f^{-1}(\{0_W\}) := \{v \in V : f(v) = 0\}\,.$$

The *nullity*[1] of f is the dimension of the kernel.

The *image* or *range* of f is

$$\mathcal{R}(f) = \operatorname{im}(f) = f(V) = \{f(v) \in W : v \in V\}\,.$$

The *rank* of f is the dimension of the image.

Theorem. (Rank-nullity theorem) Let $f\colon V \to W$ be a linear mapping, and suppose that V is finite-dimensional. Then

$$\operatorname{rank}(f) + \operatorname{nullity}(f) = \dim V.$$

In other words, the $\dim(\operatorname{im}(f)) + \dim(\ker(f)) = \dim V$.

Proof. Let $K = \{v_1, \dots, v_k\}$ be a basis for $\ker(f)$ (and therefore, $\operatorname{nullity}(f) = k$). Complete K to a basis for V:

$$B = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}\,.$$

To prove the theorem, it suffices to show that $\{f(v_{k+1}), \dots, f(v_n)\}$ is a basis for $\operatorname{image}(f)$. We first show linear independence. Suppose that

$$a_{k+1} f(v_{k+1}) + \cdots + a_n f(v_n) = 0_W.$$

Since f is linear, it follows that

$$f(a_{k+1}v_{k+1} + \cdots + a_n v_n) = a_{k+1} f(v_{k+1}) + \cdots + a_n f(v_n) = 0_W.$$

Therefore, $a_{k+1}v_{k+1} + \cdots + a_n v_n \in \ker(f)$. Since $K = \{v_1, \dots, v_k\}$ is a basis for $\ker(f)$, there are scalars $a_1, \dots, a_k$ such that

$$a_{k+1}v_{k+1} + \cdots + a_n v_n = a_1 v_1 + \cdots + a_k v_k,$$

i.e.,

$$a_1 v_1 + \cdots + a_k v_k - a_{k+1}v_{k+1} - \cdots - a_n v_n = 0_V.$$

[1] Don't confuse this concept with the *mullity* of f, defined as follows: $\operatorname{mullity}(f) = p(f) + b(f)$ where $p(f)$ is the amount of party of f in the back and $b(f)$ is the amount of business of f in the front.

This is a linear relation among the vectors of B, the basis we constructed for V. Since B is a linearly independent set, all of the a_i must be 0. In particular, $a_{k+1} = \cdots = a_n = 0$, as we were trying to show.

Next, we show that $\{f(v_{k+1}), \ldots, f(v_n)\}$ spans $\operatorname{im}(f)$. We know that since $B = \{v_1, \ldots, v_n\}$ is a basis for V that

$$\{f(v_1), \ldots, f(v_n)\}$$

spans the image of f. However, $v_1, \ldots, v_k$ are in $\ker(f)$, so

$$\begin{aligned}\operatorname{im}(f) &= \operatorname{Span}\{f(v_1), \ldots, f(v_k), f(v_{k+1}), \ldots, f(v_n)\} \\ &= \operatorname{Span}\{0_W, \ldots, 0_W, f(v_{k+1}), \ldots, f(v_n)\} \\ &= \operatorname{Span}\{f(v_{k+1}), \ldots, f(v_n)\}.\end{aligned}$$

□

Proposition 1. The linear mapping $f\colon V \to W$ is injective (i.e., one-to-one) if and only if $\ker(f) = \{0_V\}$.

Proof. ($\Rightarrow$) First suppose that f is injective, and let $v \in \ker(f)$. Therefore, $f(v) = 0_W$. We also know that since f is linear, $f(0_V) = 0_W$. So $f(v) = 0_W = f(0_V)$. Since f is injective and $f(v) = f(0_V)$, it follows that $v = 0_V$. We have shown that $\ker(f) = \{0_V\}$.

($\Leftarrow$) For the converse, now suppose that $\ker(f) = \{0_V\}$, and let $u, v \in V$ with $f(u) = f(v)$. It follows that $f(u - v) = f(u) - f(v) = 0_W$. Hence, $u - v \in \ker(f)$. However, we are assuming $\ker(f) = \{0_V\}$. So $u - v = 0_V$, which means $u = v$. Therefore, f is injective. □

Proposition 2. Let $S \subseteq V$.

(a) If S is linearly dependent, then $f(S) := \{f(s) : s \in S\} \subseteq W$ is linearly dependent. (The image of a dependent set is dependent.)

(b) If f is injective and S is linearly independent, then $f(S) \subseteq W$ is linearly independent. (The image of an independent set is independent *provided* f is injective.)

Proof. Suppose that $\sum_{i=1}^k a_i s_i = 0_V$ for some $a_i \in F$ and $s_i \in S$. Since f is linear, we have

$$0_W = f(0_V) = f(\textstyle\sum_{i=1}^k a_i s_i) = \sum_{i=1}^k a_i f(s_i).$$

Thus, f preserves linear dependencies, as claimed in part (a).

Suppose now that f is injective and S is linearly independent. If $\sum_{i=1}^k a_i f(s_i) = 0_W$ for some $a_i \in F$ and $s_i \in S$, then since f is linear,

$$0_W = \textstyle\sum_{i=1}^k a_i f(s_i) = f(\sum_{i=1}^k a_i s_i).$$

Therefore, $\sum_{i=1}^k a_i s_i$ is in the kernel of f. Since, f is injective, $\ker(f) = \{0_V\}$ by Proposition 1. It follows that $\sum_{i=1}^k a_i s_i = 0_V$. Then, since S is linearly independent, it follows that $a_i = 0$ for all i. This shows that $f(S)$ is linearly independent. □

Definition. The linear function $f\colon V \to W$ is an *isomorphism* if there exists a linear function $g\colon W \to V$ such that $g \circ f = \mathrm{id}_V$ and $f \circ g = \mathrm{id}_W$. The function g is called the *inverse of* f.

Remark. Suppose that $f\colon V \to W$ is an isomorphism. Then, just as proved in Math 112 for mappings of sets, it follows that f is bijective, i.e., both injective and surjective. For mappings of sets, being bijective is equivalent to having an inverse. The same is true for mappings of vector spaces: A linear function $f\colon V \to W$ is an isomorphism if and only if it is bijective. It turns out that if a linear function is bijective, then its inverse mapping (as a mapping of sets) is automatically linear. (Check this for yourself.)

Example. The space of 2×2 matrices over F is isomorphic to F^4. One isomorphism is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, c, d).$$

Exercise. Write $V \sim W$ if there is an isomophism $V \to W$. Check that $\sim$ is an equivalence relation.

Proposition 3. A linear mapping $f\colon V \to W$ is an isomorphism if and only if $\ker(f) = \{0_V\}$ and $\mathrm{im}(f) = W$, (i.e., if and only if its kernel is trivial and it is surjective).

Proof. We have just seen that $\ker f = \{0_V\}$ if and only if f is injective, and by definition of surjectivity, f is surjective if and only if $\mathrm{im}(f) = W$. Thus, the condition that $\ker(f)$ is trivial and $\mathrm{im}(f) = W$ is equivalent to the bijectivity of f. $\square$

Theorem 4. Let V be a vector space over F. Then V is isomorphic to F^n if and only if $\dim V = n$.

Proof. ($\Rightarrow$) Suppose that $f\colon V \to F^n$ is an isomorphism with inverse $g\colon F^n \to V$, and let $e_1, \dots, e_n$ be the standard basis for F^n. Define $v_i = g(e_i) \in V$ for $i = 1, \dots, n$. We claim that $B := \{v_1, \dots, v_n\}$ is a basis for V (and hence, $\dim V = n$). First note that B is linearly independent by Proposition 2 (a) since $\{e_1, \dots, e_n\}$ is linearly independent. Next, to see that B spans, let $v \in V$, and write

$$f(v) = \sum_{i=1}^{n} a_i e_i$$

for some $a_i \in F$. It follows that

$$v = g(f(v)) = g\left(\textstyle\sum_{i=1}^n a_i e_i\right) = \textstyle\sum_{i=1}^n a_i g(e_i) = \sum_{i=1}^n a_i v_i \in \mathrm{Span}(B).$$

($\Leftarrow$) Now suppose $\dim V = n$. Choose a basis $\{b_1, \dots, b_n\}$ for V, and let $\{e_1, \dots, e_n\}$ be the standard basis for F^n. Define $f\colon V \to F^n$ by $f(b_i) = e_i$ for $i = 1, \dots, n$ and extending linearly. Recall what this means: given $v \in V$, there are unique $\alpha_i \in F$ such that $v = \sum_{i=1}^n \alpha_i b_i$. Then by definition of "extend linearly",

$$f(v) = \sum_{i=1}^n \alpha_i f(b_i) = \sum_{i=1}^n \alpha_i e_i = (\alpha_1, \dots, \alpha_n) \in F^n.$$

Earlier, we called $(\alpha_1, \dots, \alpha_n)$ the *coordinates* of v with respect to the ordered basis $\langle b_1, \dots, b_n \rangle$.

Suppose $v \in \ker(f)$, and write $v = \sum_{i=1}^n \alpha_i b_i$. Then $0_W = f(v) = \sum_{i=1}^n \alpha_i e_i$ implies $\alpha_i = 0$ for all i since the e_i are linearly independent. So $v = 0_V$. This shows that the kernel of f is trivial, and hence, f is injective. For surjectivity, note that the image contains all linear combinations of the standard basis vectors, $e_1, \dots, e_n$ for F^n. □

Remarks: Theorem 4 says that for each $n = 0, 1, 2, \dots$, there is essentially only one vector space over F of dimension n. More precisely, under the equivalence relation $V \sim W$ defined earlier, there is one equivalence class for each natural number n. Theorem 4 and its proof say that the difference between a vector space V of dimension n and F^n is the choice of a basis. Once a basis B is chosen, we get an isomorphism $V \to F^n$ by sending each vector to its coordinates with respect to B:

$$\begin{aligned} V &\to F^n \\ v &\mapsto [v]_B. \end{aligned}$$

The practical importance of this result is that if we have a problem involving vectors in V, we can use the isomorphism to translate problem into one about n-tuples in F^n. We apply our algorithms, e.g., Gaussian elimination, to solve the problem in F^n and then use the inverse of the isomorphism to translate the solution back to V.

Corollary 5. *Let V and W be finite-dimensional vectors spaces. Then V and W are isomorphic if and only if they have the same dimension.*

Proof. First, suppose that $f\colon V \to W$ is an isomorphism, and let $b_1, \dots, b_n$ be a basis for V. By Proposition 2, $f(b_1), \dots, f(b_n)$ are linearly independent, and since f is surjective, they span W. So $\{f(b_1), \dots, f(b_n)\}$ is a basis for W. Thus, the number of elements in a basis for V is the same as the number of elements in a basis for W, which says that $\dim V = \dim W$.

Conversely, suppose that $\dim V = \dim W = n$. By Theorem 4, we have isomorphisms $f_V\colon V \to F^n$ and $f_W\colon W \to F^n$. Let $f_W^{-1}\colon F^n \to W$ be the inverse of f_W. It follows that the composition,

$$V \xrightarrow{f_V} F^n \xrightarrow{f_W^{-1}} W$$

is an isomorphism. (From Math 112, you know that a composition of bijections of sets is a bijection of sets, and you should do the easy check that a composition of linear functions is linear.) □

Proposition 6. Let $f\colon V \to W$ be a linear function, and let $\dim V = \dim W < \infty$. (An important special case is $f\colon V \to V$ when $\dim V < \infty$.) Then the following are equivalent:

(a) f is injective (1-1),

(b) f is surjective (onto),

(c) f is an isomorphism.

Proof. The proof is left as an exercise. The central idea is to use the rank-nullity theorem to relate injectivity and surjectivity. $\square$

Note: Proposition 6 is not true if the dimensions of V and W are not finite. For instance, consider the infinite-dimensional vector space $\mathcal{P}(F) = F[x]$ and the mapping

$$\begin{aligned} F[x] &\to F[x] \\ f &\mapsto xf, \end{aligned}$$

given by multiplication by x. For instance, under this mapping, $1 + x + x^2 \mapsto x + x^2 + x^3$. This mapping is linear and injective, but not surjective. For instance, 1 is not in the image (nor is any other constant besides 0).