Mathematics Department Stanford University
Math 51H – Vector spaces and linear maps

We start with the definition of a vector space; you can find this in Section A.8 of the text (over $\mathbb{R}$, but it works over any field).

Definition 1 *A vector space $(V, +, \cdot)$ over a field F is a set V with two maps $+ : V \times V \to V$ and $\cdot : F \times V \to V$ such that*

1. *$(V, +)$ is a commutative group, with unit 0 and with the inverse of x denoted by $-x$,*
2. *for all $x \in V$, $\lambda, \mu \in F$,*
$$(\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x), \ 1 \cdot x = x$$
(here 1 is the multiplicative unit of the field; also note that in the expression $\lambda\mu$ on the left hand side, we are using the field mutliplication; everywhere else the operation $\cdot : F \times V \to V$),
3. *for all $x, y \in V$, $\lambda, \mu \in F$,*
$$(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$$
and
$$\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y,$$
i.e. two distributive laws hold.

Notice that the two distributive laws are different: in the first, on the left hand side, $+$ is in F, in the second in V.

The 'same' argument as for fields shows that $0 \cdot x = 0$, and $\lambda \cdot 0 = 0$, where in the first case on the left 0 is the additive unit of the field, and all other 0's are the zero vector, i.e. the additive unit of the vector space. Another example of what one can prove in a vector space:

Lemma 1 *Suppose $(V, +, \cdot)$ is a vector space over F. Then $(-1) \cdot x = -x$.*

Proof: We have, using the distributive law, and that $0 \cdot x = 0$, observed above,
$$0 = 0 \cdot x = (1 + (-1)) \cdot x = 1 \cdot x + (-1) \cdot x = x + (-1) \cdot x,$$
which says exactly that $(-1) \cdot x$ is the additive inverse of x, which we denote by $-x$. □

Now, if F is a field, $F^n = F \times \ldots \times F$, with n factors, can be thought of as the set of maps $x : \{1, \ldots, n\} \to F$ where one writes $x_j = x(j) \in F$; of course one may just write $x = (x_1, x_2, \ldots, x_n)$ as usual, i.e. think of elements of F^n as ordered n-tuples of elements of F. Then F^n is a vector space with the definition of componentwise addition and componentwise multiplication by scalars $\lambda \in F$, as is easy to check.

A more complicated (but similar, in that it is also a function space!) vector space, over $\mathbb{R}$, is the set $C([0, 1])$ of continuous real-valued functions on the interval $[0, 1]$. Here addition and multiplication by elements of $\mathbb{R}$ is defined by:
$$(f + g)(x) = f(x) + g(x), \ (\lambda \cdot f)(x) = \lambda f(x), \ \lambda \in \mathbb{R}, \ f, g \in C([0, 1]), \ x \in [0, 1],$$
i.e. $f + g$ is the continuous function on $[0, 1]$ whose value at any $x \in [0, 1]$ is the sum of the values of f and g at that point (this is a sum in $\mathbb{R}$!), and λf is the continuous function on $[0, 1]$ whose value at any $x \in [0, 1]$ is the product of λ and the value of f at x (this is a product in $\mathbb{R}$). (Notice that this is the 'same' definition as $+, \cdot$ in F^n, if one writes elements of the latter as maps from $\{1, \ldots, n\}$: e.g. one adds components, i.e. values at $j \in \{1, \ldots, n\}$, so e.g. $(x + y)(j) = x(j) + y(j)$ in the function notation corresponds to $(x + y)_j = x_j + y_j$ in the component notation.)

For a vector space V (one often skips the operations and the field when understood), the notion of subspace, linear (in)dependence, spanning, etc., are just as for subspaces of $\mathbb{R}^n$. In particular, as pointed out in lecture, the linear dependence lemma holds. However, not all vector spaces have bases in the sense we discussed.

Definition 2 *A vector space V over F is called finite dimensional if there is a finite subset $\{v_1, \dots, v_n\}$ of elements of V such that* $\mathrm{Span}\{v_1, \dots, v_n\} = V$.

Notice that F^n is finite dimensional (the standard basis spans), while $C([0,1])$ is not.

Just as for subspaces of $\mathbb{R}^n$:

Definition 3 *A basis of a general vector space V is a finite subset of V which is linearly independent and which spans V.*

As pointed out in lecture, the proof of the basis theorem goes through without change in arbitrary finite dimensional vector spaces. This includes the statement that all bases of a finite dimensional vector space have the same number of elements, so the notion of dimension is well-defined.

There is a more general notion of basis in which the requirement of being finite is dropped. (Linear independence and spanning still only involves *finite* linear combinations though — in algebra one cannot take infinite sums; one needs a notion of convergence for the latter!) Using sophisticated tools (such as the axiom of choice), one can prove that every vector space has such a basis. However, this is *not* a very useful notion in practice, and we will *not* consider this possible definition any longer.

We now turn to linear maps, which are again defined as for subspaces of $\mathbb{R}^n$.

Definition 4 *Suppose V, W are vector spaces over a field F. A linear map $T : V \to W$ is a map $V \to W$ such that*

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty,\ \ \lambda, \mu \in F,\ \ x, y \in V.$$

It is easy to check that $T0 = 0$ and $T(-x) = -Tx$ for any linear map using $0 = 0 \cdot 0$ (with the first zero being the zero in the field), and $-x = (-1) \cdot x$, pulling the 0, resp. -1 through T using linearity. Thus, a linear map is exactly one which preserves all aspects of the vector space structure (addition, multiplication, additive units and inverses).

Two basic notions associated to a linear map T are:

Definition 5 *The nullspace, or kernel, $N(T) = \mathrm{Ker}\,T$ of a linear map $T : V \to W$ is the set $\{x \in V : Tx = 0\}$.*

The range, or image, $\mathrm{Ran}\,T = \mathrm{Im}\,T$ of a linear map $T : V \to W$ is the set $\{Tx :\ x \in V\}$.

Both of these sets are subspaces, $N(T) \subset V$, $\mathrm{Ran}\,T \subset W$, as is easily checked.

Recall that any map $f : X \to Y$ between sets X, Y has a set-theoretic inverse if and only if it is one-to-one, i.e. injective, and onto, i.e. surjective (such a one-to-one and onto map is called bijective); one defines $f^{-1}(y)$ to be the unique element x of X such that $f(x) = y$. (Existence of x follows from being onto, uniqueness from being one-to-one.)

Lemma 2 *A linear map $T : V \to W$ is one-to-one if and only if $N(T) = \{0\}$.*

Proof: Since $T0 = 0$ for any linear map, if T is one-to-one, the only element of V it may map to 0 is 0, so $N(T) = \{0\}$.

Conversely, suppose that $N(T) = 0$. If $Tx = Tx'$ for some $x, x' \in V$ then $Tx - Tx' = 0$, i.e. $T(x - x') = 0$, so $x - x' = 0$ since $N(T) = \{0\}$, so $x = x'$ showing injectivity. $\square$

The following is used in the proof of the rank-nullity theorem in the book in the special setting considered there:

Lemma 3 *If $T : V \to W$ is a linear map (F the field), and $v_1, \ldots, v_n$ span V, then* $\operatorname{Ran} T = \operatorname{Span}\{Tv_1, \ldots .Tv_n\}$.

Notice that this lemma shows that if V is finite dimensional, then $\operatorname{Ran} T$ is finite dimensional, even if W is not.

Proof of lemma: We have

$$\begin{aligned}\operatorname{Ran} T = \{Tx : \ x \in V\} &= \{T(\sum c_j v_j) : \ c_1, \ldots, c_n \in F\} \\ &= \{\sum c_j T v_j : \ c_1, \ldots, c_n \in F\} = \operatorname{Span}\{Tv_1, \ldots .Tv_n\}.\end{aligned}$$

Here the second equality uses that $v_1, \ldots, v_n$ span V, and the third that T is linear. $\square$

Recall that if $T : \mathbb{R}^n \to \mathbb{R}^m$, written as a matrix, then the jth column of T is Te_j, e_j the jth standard basis vector. Thus, in this case, the column space $C(T)$ of T (which is by definition the span of the Te_j, $j = 1, \ldots, n$) is $\operatorname{Ran} T$ by the lemma.

The rank-nullity theorem then is valid in general, provided one defines the rank as the dimension of the range, nullity as the dimension of the nullspace, with the same proof. (The proof in the textbook on the third line of the long chain of equalities starting with $C(A)$, shows $C(A) = \operatorname{Ran} A$; just start here with the argument.) Thus:

Theorem 1 *(Rank-nullity) Suppose $T : V \to W$ is linear with V finite dimensional. Then*

$$\dim \operatorname{Ran} T + \dim N(T) = \dim V.$$

This can be re-written in a slightly different way. Recall that $N(T) = \{0\}$ is exactly the statement that T is injective. Thus, $\dim N(T)$ measures the failure of injectivity. On the other hand, if W is finite dimensional, $\dim W = \dim \operatorname{Ran} T$ if and only if $W = \operatorname{Ran} T$, i.e. if and only of T is surjective. Thus, $\dim W - \dim \operatorname{Ran} T$ measures the failure of surjectivity of T. By rank-nullity,

$$(\dim \operatorname{Ran} T - \dim W) + \dim N(T) = \dim V - \dim W,$$

i.e. the *difference*

$$\dim N(T) - (\dim W - \dim \operatorname{Ran} T) = \dim V - \dim W \tag{1}$$

of the measure of the failure of injectivity and surjectivity relates to the difference of the dimensions of the domain and the target space. The expression on the left hand side is often called the *index* of T; note that bijections have index 0, but of course not every map of index 0 is bijective. However, (1) shows that a map T between two vector spaces can only be a bijection if $\dim V = \dim W$, i.e. bijective (hence invertible) linear maps can only exist between two vector spaces of equal dimensions.

Notice that if $T : V \to W$ is linear and bijective, so it has a set theoretic inverse, this inverse is necessarily linear.

Lemma 4 *If $T : V \to W$ is linear and bijective, then the inverse map T^{-1} is linear.*

Proof: Indeed, let T^{-1} be this map. Then

$$T(T^{-1}(\lambda x + \mu y)) = \lambda x + \mu y$$

by the definition of T^{-1}, and

$$T(\lambda T^{-1}x + \mu T^{-1}y) = \lambda T T^{-1}x + \mu T T^{-1}y = \lambda x + \mu y$$

where the first equality is the linearity of T and the second the definition of T^{-1}. Thus,

$$T(T^{-1}(\lambda x + \mu y)) = T(\lambda T^{-1}x + \mu T^{-1}y),$$

so by injectivity of T,

$$T^{-1}(\lambda x + \mu y) = \lambda T^{-1}x + \mu T^{-1}y,$$

proving linearity. $\square$

As discussed above for $C([0,1])$, which works equally well for any other set of maps with values in a vector space, if one has two linear maps $S, T : V \to W$, one can add these pointwise

$$(S+T)(x) = Sx + Tx,\ x \in V,$$

and multiply them by elements of the field pointwise:

$$(cT)(x) = c \cdot Tx,\ c \in F,\ x \in V.$$

Then it is straightforward to check the following:

Lemma 5 *For V, W are vector spaces over a field F, $S, T : V \to W$ linear, $c \in F$, $S+T, cT$ are linear.*

Sketch of proof: For instance, we have

$$\begin{aligned}(S+T)(\lambda x + \mu y) &= S(\lambda x + \mu y) + T(\lambda x + \mu y) = \lambda Sx + \mu Sy + \lambda Tx + \mu Ty \\ &= \lambda(Sx + Tx) + \mu(Sy + Ty) = \lambda(S+T)x + \mu(S+T)y,\end{aligned}$$

where the first and last equalities are the definition of $S+T$, while the second one is the linearity of S and that of T, while the third one uses that W is a vector space. $\square$

The next lemma, which again is easy to check, is that the set of linear maps, with the addition and multiplication by scalars which we just defined, is also a vector space, i.e. the operations $+$, $\cdot$ satisfy all properties in the definition of a vector space.

Lemma 6 *If V, W are vector spaces over a field F, then the set of linear maps $\mathcal{L}(V, W)$ is a vector space itself.*

An important property of maps (in general) is that one can compose them if the domains/targets match: if $f : X \to Y$, $g : Y \to Z$, $(g \circ f)(x) = g(f(x))$, $x \in X$. For linear maps:

Lemma 7 *Suppose $T : V \to W$, $S : W \to Z$ are linear, then $S \circ T$ is linear as well.*

Proof:

$$(S \circ T)(\lambda x + \mu y) = S(T(\lambda x + \mu y)) = S(\lambda Tx + \mu Ty) = \lambda S(Tx) + \mu S(Ty) = \lambda(S \circ T)(x) + \mu(S \circ T)(y).$$

$\square$

In particular, linear maps $V \to V$ both have a vector space structure, and a multiplication. Recall from the Week 2 fun problem that a ring is a set R with two operations $+, \cdot$ with the same properties as those of a field except $\cdot$ is not required to be commutative, and non-zero elements are not required to have multiplicative inverses. With this, the following is easy to check:

Lemma 8 *If V is a vector space, $\mathcal{L}(V,V)$ is a ring.*

Finally recall that in lecture we discussed the matrix of a linear map between finite dimensional vector spaces, and used this to motivate matrix multiplication, wanting that the product of two matrices corresponding to linear maps be the matrix of the composite linear map. The matrix of a linear map is discussed in Section 3.6 of the textbook. Recall that if $e_1, \ldots, e_n$ form a basis of V, $f_1, \ldots, f_m$ a basis of W, $T : V \to W$ linear, we write

$$Te_j = \sum_{i=1}^{m} a_{ij} f_i,$$

using that the f_i form a basis of W, and then $\{a_{ij}\}_{i=1,\ldots,m, j=1,\ldots,n}$ is the matrix of T. The computation in the lecture, see also Problem Set 3, Problem 1, shows that the matrix of $S \circ T$ has lj entry $\sum b_{li} a_{ij}$ if $\{b_{lk}\}_{l=1,\ldots,p,k=1,\ldots,m}$ is the matrix of S.

In general, two vector spaces V, W are considered the 'same as vector spaces', i.e. *isomorphic* if there is an invertible linear map $T : V \to W$ (so T^{-1} is also linear as discussed above). If W is a dimension n vector space, $V = F^n$, $e_1, \ldots, e_n$ the standard basis of V, $f_1, \ldots, f_n$ a basis of W (which exists, but one needs to pick one), then there is such a map T, namely the map $T(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j f_j$. This is linear (it is the unique linear map with $Te_j = f_j$), injective and surjective (the f_j are linearly independent and span). This *is* the sense in which any n-dimensional vector space is the 'same' as F^n, i.e. they are *isomorphic*; notice however that this isomorphism requires a choice, that of a basis of W. Notice now that if W, Z are n-dimensional vector spaces, they are also isomorphic: if $T : F^n \to W$, $S : F^n \to Z$ are isomorphism (such as those above, depending on choices of bases) then $S \circ T^{-1} : W \to Z$ is the desired isomorphism.