NOTES ON QUOTIENT SPACES

SANTIAGO CAÑEZ

Let V be a vector space over a field $\mathbb{F}$, and let W be a subspace of V. There is a sense in which we can "divide" V by W to get a new vector space. Of course, the word "divide" is in quotation marks because we can't really divide vector spaces in the usual sense of division, but there is still an analog of division we can construct. This leads the notion of what's called a *quotient vector space*. This is an incredibly useful notion, which we will use from time to time to simplify other tasks. In particular, at the end of these notes we use quotient spaces to give a simpler proof (than the one given in the book) of the fact that operators on finite dimensional complex vector spaces are "upper-triangularizable".

For each $v \in V$, denote by $v + W$ the following subset of V:

$$v + W = \{v + w \mid w \in W\}.$$

So, $v + W$ is the set of all vectors in V we can get by adding v to elements of W. Note that v itself is in $v + W$ since $v = v + 0$ and $0 \in W$. We call a subset of the form $v + W$ a *coset* of W in V. The first question to ask is: when are two such subsets equal? The answer is given by:

Proposition 1. *Let $v, v' \in V$. Then $v + W = v' + W$ if and only if $v - v' \in W$.*

Proof. Suppose that $v + W = v' + W$. Since $v \in v + W$, v must also be in $v' + W$. Thus there exists $w \in W$ so that

$$v = v' + w.$$

Hence $v - v' = w \in W$ as claimed.

Conversely, suppose that $v - v' \in W$. Without loss of generality, it is enough to show that $v \in v' + W$. Denote by w the element $v - v'$ of W. Then

$$v - v' = w, \text{ so } v = v' + w.$$

Hence $v \in v' + W$ as required. □

Definition 1. *The set V/W, pronounced "V mod W", is the set defined by*

$$V/W = \{v + W \mid v \in V\}.$$

That is, V/W is the collection of cosets of W in V.

According to Proposition 1, two elements v and v' of V determine the same element of V/W if and only if $v - v' \in W$. Because of this, it is also possible, and maybe better, to think of V/W as consisting of all elements of V under the extra condition that two elements of V are declared to be the same if their difference is in W. This is the point of view taken when defining V/W be the set of *equivalence classes* of the *equivalence relation* $\sim$ on V defined by

$$v \sim v' \text{ if } v - v' \in W.$$

For further information about equivalence relations, please consult the "Notes on Relations" paper listed on the website for my previous Math 74 course. We will not use this terminology here, and will just use the coset point of view when dealing with quotient spaces.

We want to turn V/W into a vector space, so we want to define an addition on V/W and a scalar multiplication. Since we already have such things defined on V, we have natural candidates for the required operations on V/W:

Definition 2. *The sum of two elements $v+W$ and $v'+W$ of V/W is defined by*

$$(v+W)+(v'+W):=(v+v')+W.$$

For $a\in\mathbb{F}$, the scalar multiplication of a on $v+W$ is defined by

$$a\cdot(v+W):=av+W.$$

Now we run into a problem: namely, these operations should not depend on which element of $v+W$ we choose to represent $v+W$. Meaning, we know that if $v-v'\in W$, then $v+W=v'+W$. For the above definition of addition to make sense, it should be true that for any other $v''+W$, we should have

$$(v+W)+(v''+W)=(v'+W)+(v''+W).$$

After all, $v+W$ and $v'+W$ are supposed to be equal, so computing sums with either one should give the same result. Similarly, performing scalar multiplication on either one should give the same result. Both of these requirements are true:

Proposition 2. *Suppose that $v+W=v'+W$. Then $(v+W)+(v''+W)=(v'+W)+(v''+W)$ for any $v''+W\in V/W$ and $a\cdot(v+W)=a\cdot(v'+W)$ for any $a\in\mathbb{F}$. Because of this, we say that addition and scalar multiplication on V/W are "well-defined".*

Proof. We first show that addition is well-defined. Without loss of generality, it is enough to show that $(v+W)+(v''+W)\subseteq(v'+W)+(v''+W)$. To this end, let

$$u\in(v+W)+(v''+W)=(v+v'')+W.$$

Then there exists $w\in W$ such that $u=(v+v'')+w$. We want to show that

$$u\in(v'+W)+(v''+W)=(v'+v'')+W.$$

Since $v+W=v'+W$, we know that $v-v'\in W$. Denote this element by w', so then $v=v'+w$. Thus

$$u=(v+v'')+w=((v'+w')+v'')+w=(v'+v'')+(w'+w),$$

showing that $u\in(v'+v'')+W$ since $w'+w\in W$. We conclude that the sets

$$(v+v'')+W \text{ and } (v'+v'')+W$$

are equal, so addition on V/W is well-defined.

Now, let $a\in\mathbb{F}$. To show that scalar multplication is well-defined, it is enough to show that $a\cdot(v+W)\subseteq a\cdot(v'+W)$. So, let

$$u\in a\cdot(v+W)=av+W.$$

Then $u=av+w$ for some $w\in W$. Again, denote by w' the element $v-v'\in W$. Then

$$u=av+w=a(v'+w')+w=av'+(aw'+w)\in av'+W,$$

since $aw'+w\in W$ as W is closed under addition and scalar multiplication. Thus the sets

$$a\cdot(v+W) \text{ and } a\cdot(v'+W)$$

are equal, so scalar multiplication on V/W is well-defined. ∎

Note that the above depends crucially on the assumption that W is a subspace of V; if W were just an arbitrary subset of V, addition and scalar multiplication on V/W would not be well-defined.

With these operations then, V/W becomes a vector space over $\mathbb{F}$:

Theorem 1. *With addition and scalar multiplication defined as above, V/W is an $\mathbb{F}$-vector space called the quotient space of V by W, or simply the quotient of V by W.*

Proof. All associative, commutative, and distributive laws follow directly from those of V. We only have to check the existence of an additive identity and of additive inverses.

We claim that $W = 0 + W$ is an additive identity. Indeed, let $v + W \in V/W$. Then

$$(v+W)+W = (v+0)+W = v+W \text{ and } W+(v+W) = (0+v)+W = v+W,$$

which is what we needed to show. Also, given $v + W \in V/W$, we have

$$(v+W)+(-v+W) = (v-v)+W = 0+W = W$$

and

$$(-v+W)+(v+W) = (-v+W)+W = 0+W = W,$$

so $-v+W$ is an additive inverse of $v+W$. We conclude that V/W is a vector space under the given operations. $\square$

Just to emphasize, note that the zero vector of V/W is W itself, which is the coset corresponding to the element 0 of W.

Now we move to studying these quotient spaces. The first basic question is: when is V/W finite-dimensional, and what is its dimension? The basic result is the following:

Proposition 3. *Suppose that V is finite-dimensional. Then V/W is finite-dimensional and*

$$\dim(V/W) = \dim V - \dim W.$$

So, when everything is finite-dimemsional, we have an easy formula for the dimension of the quotient space. Of course, even if V is infinite-dimesionsal, it is still possible to get a finite-dimensional quotient. We will see an example of this after the proof of the above proposition.

Proof of Proposition. We will construct an explicit basis for V/W. Let $(w_1, \ldots, w_n)$ be a basis of W, and extend it to a basis

$$(w_1, \ldots, w_n, v_1, \ldots, v_k)$$

of V. Note that with this notation, $\dim W = n$ and $\dim V = n + k$. We claim that

$$(v_1 + W, \ldots, v_k + W)$$

forms a basis of V/W. If so, we will then have

$$\dim(V/W) = k = (n+k) - n = \dim V - \dim W$$

as claimed.

First we check linear independence. Suppose that

$$a_1(v_1+W) + \cdots a_k(v_k+W) = W$$

for some scalars $a_1, \ldots, a_k$. Recall that W is the zero vector of V/W, which is why we have W on the right side of the above equation. We want to show that all a_i are zero. We can rewrite the above equation as

$$(a_1v_1 + \cdots a_kv_k) + W = W.$$

By Proposition 1 then, it must be that

$$a_1v_1 + \cdots a_kv_k \in W,$$

so we can write this vector in terms of the chosen basis of W; i.e.

$$a_1v_1 + \cdots a_kv_k = b_1w_1 + \cdots + b_nw_n$$

for some $b_1, \ldots, b_n \in \mathbb{F}$. But then

$$a_1v_1 + \cdots a_kv_k - b_1w_1 - \cdots - b_nw_n = 0,$$

so all coefficients are zero since $(w_1, \ldots, w_n, v_1, \ldots, v_k)$ is linearly independent. In particular, all the a_i are zero, showing that

$$(v_1 + W, \ldots, v_k + W)$$

is linearly independent.

Now, let $v + W \in V/W$. Since the w_i's and v_j's span V, we know that we can write

$$v = a_1w_1 + \cdots + a_nw_n + b_1v_1 + \cdots + b_kv_k$$

for some scalars $a_i, b_j \in \mathbb{F}$. We then have

$$\begin{aligned} v + W &= (a_1w_1 + \cdots + a_nw_n + b_1v_1 + \cdots + b_kv_k) + W \\ &= [(b_1v_1 + \cdots + b_kv_k) + (a_1w_1 + \cdots + a_nw_n)] + W \\ &= (b_1v_1 + \cdots + b_kv_k) + W \\ &= b_1(v_1 + W) + \cdots + b_k(v_k + W), \end{aligned}$$

where the third equality follows from the fact that $a_1w_1 + \cdots + a_nw_n \in W$. Hence

$$(v_1 + W, \ldots, v_k + W)$$

spans V/W, so we conclude that this forms a basis of V/W. □

Example 1. Let $W = \{0\}$. Then two elements v and v' of V determine the same element of V/W if and only if $v - v' \in \{0\}$; that is, if and only if $v = v'$. Hence $V/\{0\}$ is really just V itself.

Similarly, if W is all of V, then any two things in V determine the same element of V/W, so we can say V/V is the vector space containing one element; i.e. V/V is $\{0\}$.

Example 2. Let $V = \mathbb{R}^2$ and let W be the y-axis. We want to give a simple description of V/W. Recall that two elements (x, y) and (x', y') of $\mathbb{R}^2$ give the same element of V/W, meaning $(x, y) + W = (x', y') + W$, if and only if

$$(x, y) - (x', y') = (x - x', y - y') \in W.$$

This means that $x - x' = 0$ since W is the y-axis. Thus, two elements (x, y) and (x', y') of V determine the same element of V/W if and only if $x = x'$.

So, a vector of V/W is completely determined by specifying the x-coordinate since the value of the y-coordinate does not matter. In particular, any element of V/W is represented by exactly one element of the form $(x, 0)$, so we can "identify" V/W with the set of vectors of the form $(x, 0)$ — i.e. with the x-axis. Of course, the precise meaning of "identify" is that the map from V/W to the x-axis which sends

$$(x, y) + W \mapsto (x, 0)$$

is an isomorphism. We will not spell out all the details of this here, but the point is that it really does make sense to think of the quotient V/W in this case as being the same as the x-axis.

Example 3. Let $V = \mathbb{F}^\infty$ and let W be the following subspace of V:

$$W := \{(0, x_2, x_3, \ldots) \mid x_i \in \mathbb{F}\}.$$

As above, you can check that two elements of V determine the same element of V/W if and only if they have the same first coordinate. Hence an element of V/W is just determined by the value of that first coordinate x_1. This gives an identification of V/W with $\mathbb{F}$ itself (or, more precisely, with the "x_1-axis" of $\mathbb{F}^\infty$). Note that in this case, even though V and W are infinite-dimensional, the quotient V/W is finite-dimsneional — indeed it is in fact one-dimensional.

Here is a fundamental fact which we can now prove:

Theorem 2. *Let $T : V \to W$ be linear. Define a map $S : V/\operatorname{null} T \to \operatorname{range} T$, by*

$$S(v + W) = Tv.$$

Then S is well-defined and is an isomorphism.

Proof. As in the case of the definition of addition and scalar multiplication, we may have different elements v and v' of V giving the same element of V/W. For S to be well-defined, we must know that the definition of S does not depend on such v and v'.

So, suppose that $v + \operatorname{null} T = v' + \operatorname{null} T$. Then $v - v' \in \operatorname{null} T$, so $T(v - v') = 0$. Hence $Tv = Tv'$, so $S(v + \operatorname{null} T) = S(v' + \operatorname{null} T)$, showing that S is well-defined. Also, linearity of S follows from that of T.

Now, note that surjectivity of S follows from the fact that T is surjective onto its range. We claim that S is injective. Indeed, let $v + \operatorname{null} T \in \operatorname{null} S$. This means that

$$0 = S(v + \operatorname{null} T) = Tv.$$

Hence $v \in \operatorname{null} T$, and thus $v + \operatorname{null} T = \operatorname{null} T$ is the zero vector of V/W, showing that $\operatorname{null} S = \{0\}$. Therefore S is injective, and we conclude that S is an isomorphism as claimed. □

To get a sense of the previous theorem, note that given any linear map $T : V \to W$, we can construct a surjective linear map $T : V \to \operatorname{range} T$ simply by restricting the target space. Now, this new map is not necessarily invertible since it may have a nonzero null space. However, by taking the quotient of V by $\operatorname{null} T$, we are forcing the part that makes T possibly not injective to be zero. Thus the resulting map on $V/\operatorname{null} T$ should be injective. So, from any linear map $T : V \to W$ we can construct an isomorphism between some possibly different spaces that still reflects many of the properties of T. In particular, since we know what the dimension of a quotient space is, one can use this theorem (or more precisely the consequence that $\dim(V/\operatorname{null} T) = \dim \operatorname{range} T$) to give a different proof that when V and W are finite dimensional, $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$.

Now, given an operator $T \in \mathcal{L}(V)$, we can try to define an operator on V/W, which by abuse of notation we will still denote by T, by setting

$$T(v + W) := Tv + W.$$

However, as should now be seen to be a common thing when dealing with quotient spaces, we must check that this is well-defined . To see when this is the case, first recall that $v + W = v' + W$ if and only if $v - v' \in W$. Call this vector w. Then, we have $v = v' + w$ so

$$T(v + W) = Tv + W = (Tv' + Tw) + W,$$

which should equal $Tv' + W$. We see that this is the case only when $Tw \in W$. Thus, for T to descend to a well-defined operator on V/W, we must require that

$$Tw \in W \text{ for any } w \in W.$$

In other words, we have shown:

Proposition 4. *Let $T \in \mathcal{L}(V)$. Define $T : V/W \to V/W$ by*

$$T(v+W) = Tv + W \text{ for any } v + W \in V/W.$$

Then $T : V/W \to V/W$ is a well-defined operator if and only if W is T-invariant.

One can also show that any operator on V/W arises in this way, so the operators on V/W come from operators on V under which W is invariant.

Finally, we use this new fancy language of quotient spaces to give a relatively simple proof that any operator on a finite-dimensional complex vector space has an upper-triangular matrix in some basis.

Theorem 3. *Suppose that V is a finite dimensional complex vector space and let $T \in \mathcal{L}(V)$. Then there exists a basis of V with respect to which $\mathcal{M}(T)$ is upper-triangular.*

Proof. Say $\dim V = n$. We know that T has at least one nonzero eigenvector, call it v_1. This will be the first element of our basis. The idea of the proof is to construct the rest of the basis by taking eigenvectors of the operator induced by T on certain quotients of V. We proceed as follows.

Since v_1 is an eigenvector of T, $\operatorname{span}(v_1)$ is T-invariant. Thus T descends to a well-defined operator

$$T : V/\operatorname{span}(v_1) \to V/\operatorname{span}(v_1).$$

This quotient space is again a finite dimensional complex vector space, so this induced map also has a nonzero eigenvector, call it $v_2 + \operatorname{span}(v_1)$. Then if λ is the eigenvalue of this eigenvector, we have

$$T(v_2 + \operatorname{span}(v_1)) = \lambda(v_2 + \operatorname{span}(v_1)).$$

Let us see what this equation really means. Rewriting both sides, we have

$$Tv_2 + \operatorname{span}(v_1) = \lambda v_2 + \operatorname{span}(v_1),$$

so $Tv_2 - \lambda v_2 \in \operatorname{span}(v_1)$. This implies that $Tv_2 \in \operatorname{span}(v_1, v_2)$, so $\operatorname{span}(v_1, v_2)$ is T-invariant. Note also that since this $v_2 + \operatorname{span}(v_1)$ is nonzero in $V/\operatorname{span}(v_1)$, $v_2 \notin \operatorname{span}(v_1)$, so (v_1, v_2) is linearly independent. This $v_2 \in V$ gives us the second basis element we are looking for.

Now T gives a well-defined operator

$$T : V/\operatorname{span}(v_1, v_2) \to V/\operatorname{span}(v_1, v_2).$$

Pick an eigenvector $v_3 + \operatorname{span}(v_1, v_2)$ of this, and argue as above to show that $\operatorname{span}(v_1, v_2, v_3)$ is T-invariant and that (v_1, v_2, v_3) is linearly independent. Continuing this process, we construct a basis of V:

$$(v_1, v_2, \ldots, v_n),$$

with the property that for any k, $\operatorname{span}(v_1, \ldots, v_k)$ is T-invariant. By one of the characterizations of $\mathcal{M}(T)$ being upper-triangular, we conclude that $\mathcal{M}(T)$ is upper-triangular with respect to this basis. □

Compare this proof with the one given in the book. There it is not so clear where the proof comes from, let alone where the basis constructed itself really comes from. Here we see that this result is simply a consequence of the repeated application of the fact that operators on finite dimensional complex vector spaces have eigenvectors. The basis we constructed comes exactly from eigenvectors of successive quotients of V.