Math 4310 Handout - Quotient Vector Spaces

Dan Collins

The textbook defines a *subspace* of a vector space in Chapter 4, but it avoids ever discussing the notion of a *quotient space*. This is understandable since quotient spaces can be a bit harder to wrap one's head around, but they're useful for a deeper understanding of certain concepts, and they're essential in other areas of abstract algebra and in many uses of linear algebra elsewhere in mathematics.

The definition of a quotient space is a lot like the definition of $\mathbb{Z}/n\mathbb{Z}$ - it proceeds by first defining an equivalence relation, and then working with the resulting quotient set.

Definition 1. Let V be a vector space and W a subspace. We can define an equivalence relation on V (with we'll denote $\sim$ for now) by $v \sim v'$ if $v' - v \in W$.

It's an easy exercise to see that the axioms of W being a subspace implies this is actually an equivalence relation. (You should think through why this is true, and especially where the hypothesis "W is nonempty" is important).

Definition 2. Let V be a vector space and W a subspace. Then the equivalence class of a vector v under the relation $\sim$ above is the set

$$v + W = \{v + w : w \in W\};$$

this is called an *affine subspace* a *coset* for v and W.

Again, it's pretty straightforward to see that the set we've called $v + W$ is actually the equivalence class of v: certainly we have $v \sim v + w$ for any $w \in W$, and conversely if $v \sim v'$ then $v - v' \in W$ means $v' = v + (v' - v) \in v + W$.

A geometric example of this to keep in mind is if $V = \mathbb{R}^2$ is the plane and W is some line through the origin. Then the cosets of W are all of the translates of that line; so we have a partition of the plane by separating it into a bunch of parallel lines. Similarly if $V = \mathbb{R}^3$ and W is a plane through the origin, its cosets are all of the parallel plane, and we can imagine this as filling up $\mathbb{R}^3$ with a stack of planes.

Definition 3. Let V be a vector space and W a subspace. We let V/W denote the quotient set $V/\sim$, i.e. the set of all cosets $v + W$; we call it the *quotient vector space*.

So in our two examples above, V/W is a set of parallel lines or a set of parallel planes, respectively. As the notation suggests, this is actually a vector space.

Proposition 4. *The quotient set V/W is a vector space itself, with addition and scalar multiplication defined by*

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

Proof. The main thing we need to do here is check that these operations are well-defined. So suppose $v + W = v' + W$ and $u + W = u' + W$; for well-definedness of addition we need to check that $(v + u) + W = (v' + u') + W$. But $v + W = v' + W$ means $v - v' \in W$, and similarly $u + W = u' + W$ implies $u - u' \in W$. Since W is a subspace and thus closed under addition, we can conclude

$$(v - v') + (u - u') = (v + u) - (v' + u') \in W,$$

and thus $(v + u) + W = (v' + u') + W$. Similarly, for any $a \in F$, $v - v' \in W$ means $a(v - v') = av - av' \in W$ and thus $av + W = av' + W$, so scalar multiplication is well-defined.

So this justifies that V/W has well-defined addition and scalar multiplication operations. To check it's a vector space we then need to verify the eight axioms. But these are all essentially immediate from the appropriate axioms in V; for instance we can prove "distributivity of scalar multiplication over field addition" by

$$(a + b) \cdot (v + W) = ((a + b)v) + W = (av + bv) + W = (av + W) + (bv + W) = a(v + W) + b(v + W),$$

where the second equality uses the same axiom in V, and the rest just come from the definitions of the operations on V/W. □

Since this is a quotient set, it can be a bit hard to think about clearly - it's a set of sets, which we're somehow defining operations on. But in this case the sets are pretty simple: just like in the examples above V/W can be pictured as a set or parallel lines or parallel planes, in the general case V/W we can think of as being the set of W and everything parallel to it that we can fit inside of V.

The notation is also fairly suggestive, and fortunately the things it seems to suggest actually do make sense. Writing $v+W$ makes it sound like we're adding a vector v to a plane W, and thinking about it that's what we're doing: we're taking v plus each vector in the plane W! And similarly, $(v+W)+(v'+W)$ actually equals the set of all sums $(v+w)+(v'+w')$ for $v+w \in v+W$ and $v'+W' \in v'+W'$. So we really are just doing addition, just we've expanded "addition" to let us add together sets (by adding together all of their elements). Likewise, if $a \neq 0$ is a scalar then $a \cdot (v+W)$ is the set of all multiples $a(v+w)$. (We do need to be a bit careful, though: $0 \cdot (v+W)$ is defined to be the coset $0+W=W$, but the set of all products $0 \cdot (v+w)$ is just $\{0\}$).

So now we have this abstract definition of a quotient vector space, and you may be wondering *why* we're making this definition, and what are some useful examples of it. One reason will be in our study of homomorphisms (in a couple weeks); but even with that in mind it can be hard to give interesting examples *within* the subject of linear algebra itself. However, there are many important examples that come up when we use linear algebra in other subjects, and especially when we study vector spaces with more structure on them. I'll give two examples where quotient spaces do naturally come up - these may be a bit challenging given where we are in the course, but they can provide a view of where we might go with these abstract definitions.

Extended example 1: L^2 **space.** One example that comes up naturally is from real analysis. Let $I=[0,1]$ be the unit interval, and then let $\mathcal{L}^2(I) \supseteq \mathcal{F}(I,\mathbb{R})$ be the set of functions $f:[0,1] \to \mathbb{R}$ that are integrable, and such that $\int_0^1 |f(x)|^2 dx < \infty$. (Strictly speaking the definition usually uses the requirement that f is *Lebesgue* integrable, which includes more functions than the definition of *Riemann* integrability you've likely seen in calculus or analysis classes. But this isn't really important for what I'm going to say here).

So, $\mathcal{L}^2(I)$ is a set of functions, and it's not too hard to check that it's a real vector space of the space $\mathcal{F}(I,\mathbb{R})$ of all functions $I \to \mathbb{R}$. What we'd like to do is put a "norm" on it, much like the Euclidean norm $\|\cdot\|$ on $\mathbb{R}^n$. Recall that the Euclidean norm is given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Thus $\mathbf{x} \mapsto \|\mathbf{x}\|$ is a function taking a vector as an input and returning a nonnegative real number as an output, which satisfies the following properties:

- (Positive-definiteness): If $\mathbf{x} \neq 0$ then $\|\mathbf{x}\| \neq 0$.
- (Homogeneity): If $\mathbf{x} \in \mathbb{R}^n$ and $a \in \mathbb{R}$, we have $\|a\mathbf{x}\| = |a| \cdot \|\mathbf{x}\|$.
- (Triangle inequality): If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

In fact these three properties are taken as the abstract definition of a *norm* on a vector space.

So we want to put a norm on $\mathcal{L}^2(I)$. Since $\mathcal{L}^2(I)$ is defined in terms of integrability of functions, and in particular by requiring that the integral $\int |f|^2 dx$ is finite, it's natural to try to define a norm on it by using this integral; so for $f \in \mathcal{L}^2(I)$ we define

$$\|f\| = \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}.$$

The reason we take a square root is so that the homogeneity property is satisfied. It also turns out to not be too difficult to prove the triangle inequality $\|f+g\| \leq \|f\| + \|g\|$ (though it would be a bit out of the way to do it here). But what about positive-definiteness? That axiom actually fails - for instance a function like

$$h(x) = \left\{ \begin{array}{ll} 1 & x = 1/2 \\ 0 & x \neq 1/2 \end{array} \right. ,$$

which is zero except at a single point, is certainly integrable and not the zero function, but the integral of $|f|^2$ is zero.

There's a few ways we could try to fix this. One thing we could do is replace $\mathcal{L}^2(I)$ with the space of *continuous* functions $\mathcal{C}(I, \mathbb{R})$ - on this space our function is actually positive-definiteness. But this turns out not to be such a good idea, because when we work with these spaces in analysis we want to be able to take limits, and it's possible to have a limit of continuous functions that isn't continuous. Another possibility might be to just try to throw out "obviously" bad functions like the $h(x)$ mentioned above. But there isn't really a reasonable way to do this - for instance neither functions

$$h_1(x) = \left\{ \begin{array}{ll} 1 & x \geq 1/2 \\ 0 & x < 1/2 \end{array} \right. \qquad\qquad h_2(x) = \left\{ \begin{array}{ll} 1 & x > 1/2 \\ 0 & x \leq 1/2 \end{array} \right. ,$$

seems particularly "better" or worse than the other, but their difference is $h(x)$ so we'd need to exclude one of them.

We could also consider just working with the full set $\mathcal{L}^2(I)$, and dropping the positive-definiteness condition. But this also isn't an ideal thing to do. For instance, one thing we'd ultimately like to get out of this theory is being able to write Fourier series expansions

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \sin(2\pi n x) + \sum_{n=1}^{\infty} b_n \cos(2\pi n x),$$

where the coefficients a_i and b_i come from certain integrals involving the function f. But when we're doing anything with integrals, the function h I wrote down above behaves exactly the same as the zero function, so we'd find ourselves writing down the same expansion for both of them. And if we're interpreting things as actual equalities of functions, we certainly don't want to say that two different functions are equal to the same thing!

So what's the way to fix this? It turns out that $\mathcal{L}^2(I)$ itself isn't the thing we actually want to work with - we instead want the quotient space $L^2(I) = \mathcal{L}^2(I)/L_0$, where L_0 is the collection of all functions that have norm zero:

$$L_0 = \{f : \textstyle\int_0^1 |f(x)|^2 = 0\}.$$

In the Lebesgue integration theory this is equal to the set of functions which are "almost everywhere equal to zero", and thus don't affect integrals at all. It's not too hard to check that this is a subspace of $\mathcal{L}^2(I)$, so we can actually form the quotient we want. Then the norm function $\|\cdot\|$ is well-defined on the quotient space $L^2(I)$ because if two functions f, g are in the same coset their difference doesn't affect integrals!

So now we can talk about $L^2(I)$ as a normed vector space; in fact it's what's called a *Hilbert space*, and the theory of Hilbert spaces is well studied. What the theory tells us is then that for any function f, we do have a Fourier series decomposition

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \sin(2\pi n x) + \sum_{n=1}^{\infty} b_n \cos(2\pi n x),$$

where the coefficients can be written explicitly as integrals involving f. But this is an equality that holds *in the space* $L^2(I)$ - so on the level of functions the equality is only up to adding a function in L_0. But this is all we could hope to ask for, because the coefficients being defined as integrals means that changing by something in L_0 can't affect them! (And it's also important to have this restriction because there are Fourier series expansions where the infinite series diverge at a few specific values of x).

The upshot is that in this context, talking about equality in our quotient space $L^2(I)$ is the same as talking about equality "almost everywhere" of actual functions in $\mathcal{L}^2(I)$ - and when working with integrals and measure theory, this is what really makes sense to talk about.

Extended example 2: Adjoining roots to fields. In the Fields handout from last week, I defined a field

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\};$$

we can say this field is obtained from the rationals $\mathbb{Q}$ by adjoining the element $\sqrt{2}$, i.e. a root to the polynomial $p(x) = x^2 - 2$. But the reason we could do this is we *already* have a field $\mathbb{R} \supseteq \mathbb{Q}$ in which $\sqrt{2}$ exists that we can work inside of. What would have happened if we didn't already know we had a field $\mathbb{R}$ in which a square root of 2 existed? For instance, what if we want to add a square root of 2 to the finite field $\mathbb{F}_3$ (where you can check that none of the three elements satisfy $a^2 = 2$)? We can formulate this problem as follows

Question 5. Given a field F and a polynomial $p(x) \in F[x]$ without a root in F, can we construct a field $K \supseteq F$ containing a root of $p(x)$?

It turns out the most direct way to do this is to work with a quotient space of $F[x]$. If we take a quotient $F[x]/I$ we know it's a vector space over F, and if we set up the subspace I appropriately we can make it so that $F[x]/I$ inherits the multiplication operation from $F[x]$. Namely, we have the following theorem (which I'll omit the proof since it's outside of the scope of what we're doing now):

Theorem 6. *Let F be a field and $p(x) \in F[x]$ be an irreducible polynomial (one that doesn't factor as a product of two smaller-degree polynomials). Let $I \subseteq F[x]$ be the set of all polynomials divisible by $p(x)$. Then:*

1. *The set I is a subspace of $F[x]$, so we can form a quotient space $F[x]/I$.*
2. *There is a well-defined multiplication operation $F[x]/I \times F[x]/I \to F[x]/I$ given by $(f + I) \cdot (g + I) = fg + I$.*
3. *With this multiplication operation (plus addition coming from its vector space structure, $F[x]/I$ is a field).*
4. *$F[x]/I$ contains a copy of F as a subfield (with an element $a \in F$ arising as the coset $a + I$).*
5. *The element $x + I$ in $F[x]/I$ is a root of the polynomial $p(x)$.*

This is a bit of an abstract theorem, so we'll work with the concrete example above: taking the field $F = \mathbb{F}_3$ (which is $\mathbb{Z}/3\mathbb{Z}$, so has three elements we'll call $0, 1, 2$ for simplicity) and the polynomial $p(x) = x^2 - 2$. First of all, in this case I consists of multiples of a degree-2 polynomial, so it *doesn't* contain any constant polynomials (besides 0) or any degree-1 polynomials. Thus we can see that the following nine elements of $F[x]/I$ are all distinct:

$$0 + I, 1 + I, 2 + I, x + I, (x + 1) + I, (x + 2) + I, 2x + I, (2x + 1) + I, (2x + 2) + I.$$

Here we can see $\mathbb{F}_3$ "sits inside" of the quotient $F[x]/I$ as the cosets of the constant polynomials.

Our theorem tells us that the element $x + I$ is a root of the polynomial p. We can see this directly because we have

$$p(x + I) = (x + I)^2 - 2 = (x^2 - 2) + I = 0 + I$$

is the zero element of this quotient ring (remember that addition and multiplication in $F[x]/I$ are both defined by just doing the addition or multiplication in $F[x]$ and then taking the coset).

Moreover, we claim that the nine elements are *all* of the elements of $F[x]/I$. A general element of $F[x]/I$ is a coset of a polynomial,

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + I = (a_0 + I) + (a_1 + I)(x + I) + (a_2 + I)(x + I)^2 + \cdots + (a_n + I)(x + I)^n.$$

But since $(x+I)^2 = 2+I$ (because $x+I$ satisfies the polynomial $p(x)$), so we can reduce any power $(x+I)^n$ to something already on our list!

So summarizing, we've constructed a field $F[x]/I$ with nine elements, in which $\mathbb{F}_3$ sits inside in an obvious way (strictly speaking $\mathbb{F}_3$ isn't a subset, but there's a natural *embedding* as the constant polynomials), and in which we also have a root of the polynomial $x^2 - 2$. We can call this field $\mathbb{F}_9$; if we let $\alpha = x + I$ we can consider α to be an element that's "the square root of 2 over $\mathbb{F}_3$" and our field consists of the 9 elements

$$0, 1, 2, \alpha, \alpha + 1, \alpha + 2, 2\alpha, 2\alpha + 1, 2\alpha + 2$$

(and at this point we'll generally abuse notation and think of the subset $\{0, 1, 2\}$ as *being* the field $\mathbb{F}_3$).

In contrast to the first example, in this case we don't ultimately care what our field was constructed as a quotient of! We just want to know there *exists* a field $\mathbb{F}_9$ containing $\mathbb{F}_3$ and containing an element α satisfying $\alpha^2 = 2$ (and which necessarily consist exactly of the 9 elements listed). At this point we just want to think of $\mathbb{F}_9$ as an abstract 9-element set with a certain addition and multiplication table; thinking about what's "under the hood" (i.e. that the elements are cosets in $F[x]$) just ends up complicating things.