Math 4310 Handout - Isomorphism Theorems

Dan Collins

Now that we've talked about linear transformations, quotient spaces will (finally) start to show up more naturally. I'll start by going back and giving a careful proof of one of the theorems I mentioned in class:

Proposition 1. *Let V be a finite-dimensional vector space and W a subspace. Then V/W is finite-dimensional and* $\dim(V/W) = \dim V - \dim W$.

Proof. Let $w_1, \ldots, w_m$ be a basis for W; by the Basis Extension Theorem we can extend this to a basis $w_1, \ldots, w_m, v_{m+1}, \ldots, v_n$ for V (so $n = \dim V$ and $m = \dim W$). I claim that the cosets

$$v_{m+1} + W, \ldots, v_n + W \in V/W$$

form a basis for V/W; since this is a list of $n - m$ elements it will prove the theorem.

To prove that this is a basis, we need to show that it's linearly independent and that it spans. For spanning, suppose we have an arbitrary element $v+W \in V/W$, where we've written this with a representative v picked out. By using that $w_1, \ldots, w_m, v_{m+1}, \ldots, v_n$ is a basis for V we find we can (uniquely) write

$$v = a_1 w_1 + \cdots + a_m w_m + a_{m+1} v_{m+1} + \cdots + a_n v_n.$$

Passing from this equality in V to one in the quotient set V/W, we have

$$\begin{aligned} v + W &= \big(a_1 w_1 + \cdots + a_m w_m + a_{m+1} v_{m+1} + \cdots + a_n v_n\big) + W \\ &= a_1(w_1 + W) + \cdots + a_m(w_m + W) + a_{m+1}(v_{m+1} + W) + \cdots + a_n(v_n + W). \end{aligned}$$

But since each w_i is actually in W already, the coset $w_i + W$ is just the trivial coset W, i.e. the zero element of the vector space V/W. So we actually have

$$v + W = a_{m+1}(v_{m+1} + W) + \cdots + a_n(v_n + W).$$

Since $v + W$ is an arbitrary element of V/W this tells us the set $v_{m+1} + W, \ldots, v_n + W$ spans.

We next need to check linear independence. So assume we have a dependence relation

$$b_{m+1}(v_{m+1} + W) + \cdots + b_n(v_n + W) = (b_{m+1} v_{m+1} + \cdots + b_n v_n) + W = 0 + W$$

in V/W. By definition of equality of cosets this means

$$b_{m+1} v_{m+1} + \cdots + b_n v_n \in W,$$

and thus we can uniquely write this element as a linear combination $b_1 w_1 + \cdots + b_m w_m$ in using that $\{w_1, \ldots, w_m\}$ is a basis for W. But then the equality

$$b_1 w_1 + \cdots + b_m w_m = b_{m+1} v_{m+1} + \cdots + b_n v_n$$

rearranges to a dependence relation on the set $w_1, \ldots, w_m, v_{m+1}, \ldots, v_n$, which is a basis for V; so all of the coefficients b_i have to be trivial. □

The first isomorphism theorem. In class we've talked about the rank-nullity theorem: if $T : V \to W$ is a linear transformation and V is finite-dimensional, we have an equation that we can write as

$$\dim \operatorname{img}(T) = \dim V - \dim \ker(T).$$

Here, $\ker(T)$ is a subspace of V, so we can form the quotient space $V/\ker(T)$. If we look at the proposition I proved above, we also find

$$\dim V/\ker(T) = \dim V - \dim \ker(T).$$

So the rank-nullity theorem can be rephrased as saying "the image $\operatorname{img}(T)$ and the quotient space $V/\ker(T)$ always have the same dimension"! An abstract result known as the *first isomorphism theorem* says something even better, that $\operatorname{img}(T)$ and $V/\ker(T)$ are actually isomorphic in a very natural way.

Theorem 2 (First isomorphism theorem)**.** *Let V be a vector space and $T : V \to W$ a linear transformation. Then T induces an isomorphism $\tau : V/\ker(T) \to \operatorname{img}(T)$ defined by*

$$\tau(v + \ker(T)) = T(v).$$

Proof. First of all, we need to make sure this makes sense. Since we're defining τ on the *coset* $v + \ker(T)$ in terms of the *representative* v, we need to check well-definedness, i.e. that if $v + \ker(T) = v' + \ker(T)$ then the values $T(v)$ and $T(v')$ we're trying to assign as outputs are equal. But v, v' being in the same coset means $v' - v$ is in $\ker(T)$, and thus

$$T(v') = T((v' - v) + v) = T(v' - v) + T(v) = 0 + T(v) = T(v).$$

So the map is well-defined. And since all of the values $T(v)$ lie inside of $\operatorname{img}(T)$ by definition, we don't have any problems with the codomain either.

So τ is a well-defined function; to show it's an isomorphism we need to show that it's linear, that it's injective, and that it's surjective. All of these are pretty straightforward. Linearity follows from linearity of T:

$$\tau\big((v+\ker(T))+(v'+\ker(T))\big) = \tau(v+v'+\ker(T)) = T(v+v') = T(v)+T(v') = \tau(v+\ker(T))+\tau(v'+\ker(T)),$$

$$\tau\big(a(v + \ker(T))\big) = \tau(av + \ker(T)) = T(av) = a \cdot T(v) = a\tau(v + \ker(T)).$$

For injectivity, we need to check that if $\tau(v + \ker(T)) = 0$ then $v + \ker(T) = 0$; but this is basically trivial because if $\tau(v + \ker(T)) = T(v) = 0$ then $v \in \ker(T)$ by definition. For surjectivity, any element of $\operatorname{img}(T)$ can be written as $T(v)$ for some $v \in V$ and thus is equal to $\tau(v + \operatorname{img}(T))$. □

We can think of the first isomorphism theorem as a "refined version" of the rank-nullity theorem: it gives us an *explicit, specific* way of constructing an isomorphism $V/\ker(T) \cong \operatorname{img}(T)$, and knowing this isomorphism tells us $\dim V/\ker(T) = \dim \operatorname{img}(T)$ (which is a rephrasing of the rank-nullity theorem).

If we started with the rank-nullity theorem instead, the fact that $\dim V/\ker(T) = \dim \operatorname{img}(T)$ tells us that there is *some* way to construct an isomorphism $V/\ker(T) = \operatorname{img}(T)$, but doesn't tell us anything much about what such an isomorphism would look like. The first isomorphism theorem does tell us what the isomorphism is, and shows that it comes pretty directly from T itself.

The universal mapping property. If you go back to the proof of the first isomorphism theorem, really most of the work is in showing that if we start with $T : V \to W$, then we get a well-defined "induced map" $\tau : V/\ker(T) \to \operatorname{img}(T)$. That sort of argument works in a bit more generality, which gives us the following important result:

Theorem 3 (Universal mapping property for quotient spaces)**.** *Let F be a field, V, W vector spaces over F, $T : V \to W$ a linear transformation, and $U \subseteq V$ a subspace. If $U \subseteq \ker(T)$, then there is a unique well-defined linear transformation $\tau : V/U \to W$ given by $\tau(v + U) = T(v)$.*

If $\pi : V \to V/U$ is the canonical projection (i.e. the linear transformation given by $\pi(v) = v + W$), this can be rephrased as saying that there's a unique well-defined linear transformation τ satisfying $\tau \circ \pi = T$. We can think of this as saying T "factors through" the quotient space V/U: starting with a map $V \to W$, we can actually split it up as two maps $V \to V/U \to W$.

Proof. This is basically the same proof as above (minus the last few lines). To see τ is well-defined on a coset $v + U$ we need to check that if $v + U = v' + U$ then $T(v) = T(v')$; but this follows because $v + U = v' + U$ means $v - v' \in U$ and thus $v - v' \in \ker(T)$ because U is contained in the kernel. Then we have $T(v - v') = 0$ by definition, and rearranging and using linearity gives $T(v) = T(v')$. Linearity is then a formal consequence of linearity of T:

$$\tau\big((v + U) + (v' + U)\big) = \tau(v + v' + U) = T(v + v') = T(v) + T(v') = \tau(v + U) + \tau(v' + U),$$

$$\tau\big(a(v + U)\big) = \tau(av + U) = T(av) = a \cdot T(v) = a\tau(v + U).$$ □

This gives us a systematic way of constructing linear transformations on quotient spaces: to get a linear transformation $V/U \to W$, we just need to start with a linear transformation $V \to W$ which is trivial on U (i.e. has the kernel containing U). The terminology "universal mapping property" refers to any framework like this: starting with a function T satisfying some certain properties, we can conclude there exists a *unique* map τ defined in a certain way in terms of T.

Example 4. At this point, a fair question to ask is "why do I actually need to work with linear transformations defined on quotient spaces"? Like the question of "why do I actually need to work with quotient spaces", it's hard to give an answer entirely within linear algebra itself: most of the important uses of quotient spaces come up when you *apply* linear algebra to other subjects.

So I'll give an example building off of the Extended Example 1 (of the space $L^2(I)$, of "square-integrable functions $[0,1] \to \mathbb{R}$) in the "quotient vector spaces" handout. In that example, we defined $\mathcal{L}^2(I)$ to be the space of integrable functions $f : I \to \mathbb{R}$ such that $\int_0^1 |f|^2 dx < \infty$, and then as the quotient of this by the subspace U of all functions that were "almost everywhere zero".

To work with this space $L^2(I)$ in analysis, we want to be able to integrate functions on it! Say we fix some function like $\sin(2\pi x)$, and we want to consider the linear functional of "integrating against it":

$$f \mapsto \int_0^1 f(x)\sin(2\pi x)dx.$$

This makes perfect sense for any f in the actual space of functions $\mathcal{L}^2(I)$, and gives us a linear transformation $\mathcal{L}^2(I) \to \mathbb{R}$. But what we'd really like is a linear transformation $L^2(I) \to \mathbb{R}$. Fortunately, the universal mapping property lets us do this! If f is in the subspace U of "almost everywhere zero" functions, then $f(x)\sin(2\pi x)$ is also "almost everywhere zero", so its integral is zero. Thus $f \mapsto \int_0^1 f(x)\sin(2\pi x)dx$ is trivial on the subspace U, and the Universal Mapping Property tells us that we actually get a homomorphism $L^2(I) \to \mathbb{R}$ given by

$$[f] = f + U \mapsto \int_0^1 f(x)\sin(2\pi x)dx.$$

This is the functional that recovers one of the Fourier coefficients of f (which, again, makes sense: we've said that Fourier series only make sense up to "equality almost everywhere"!)

The other isomorphism theorems. From the name "the first isomorphism theorem", you can probably guess that there's a few more "isomorphism theorems" to go along with it. (The universal mapping property can sometimes be grouped in with them as well). These other isomorphism theorems are a bit less important to us in this class, but they're indispensable if you're going to be seriously working with quotient spaces.

The "second isomorphism theorem" concerns what happens when you have a vector space V and two subspaces U, W, and you take a quotient $(U+W)/W$. Your first thought might be that you can "cancel out the Ws" and just be left with something isomorphic to U - this is close to correct, but you need to compensate for any overlap between U and W.

Theorem 5 (Second isomorphism theorem)**.** *Let V be a vector space and $U, W \subseteq V$ two subspaces. Then there's an isomorphism of quotient spaces*

$$\frac{U}{U \cap W} \cong \frac{U+W}{W}$$

given by $u + (U \cap W) \mapsto u + W$.

The "third isomorphism theorem" is about quotient spaces of quotient spaces, which are pretty unpleasant to think about if you're not already really comfortable with quotient spaces.

Theorem 6 (Third isomorphism theorem)**.** *Let V be a vector space, W a subspace of V, and U a subspace of W. Then the quotient space W/U is itself a subspace of the quotient space V/U, and we have a canonical* *isomorphism*

$$\frac{V/U}{W/U} \cong V/W$$

by mapping $(v+U)+W/U$ *(a coset in* V/U *by the subspace* W/U*!) to* $v+W$*.*

There's one last theorem usually grouped with these, which is usually called the "correspondence theorem" or "lattice isomorphism theorem" and tells you about all of the subspaces in a quotient. We take the notation that $\mathrm{Sub}(V)$ denotes the collection of all subspaces of V.

Theorem 7 (Corresponcence theorem)**.** *Let* V *be a vector space and* W *a subspace of* V*. Then there is a bijective correspondence*

$$\mathrm{Sub}(V/W) \leftrightarrow \{U \in \mathrm{Sub}(V) : W \subseteq U \subseteq V\},$$

given by taking a subspace U *with* $W \subseteq U \subseteq V$ *to the subspace* U/W *to* V/W*. This correspondence preserves sums and intersections: if we add or intersect two subspaces* U_1/W *and* U_2/W *of* V/W *we get*

$$\frac{U_1}{W} + \frac{U_2}{W} = \frac{U_1+U_2}{W} \qquad \frac{U_1}{W} \cap \frac{U_2}{W} = \frac{U_1 \cap U_2}{W}.$$

This characterizes $\mathrm{Sub}(V/W)$ in terms of a subset of $\mathrm{Sub}(V)$. Actually, the set $\mathrm{Sub}(V)$ of subspaces naturally has a partial order (by inclusion), and it's a *lattice* with respect to this partial order: any two subspaces U_1, U_2 have a *join* $U_1 + U_2$ (a "least upper bound") and a *join* $U_1 \cap U_2$ (a "greatest lower bound"). The last part of the theorem tells us that the lattice structure $\mathrm{Sub}(V/W)$ is compatible with the lattice structure on the sublattice $\{U : W \subseteq U \subseteq V\}$ of $\mathrm{Sub}(V)$, hence the name "lattice isomorphism theorem".

I'm omitting the proofs of these theorems in this section; trying to prove them yourself might be a good way to get in better practice with quotient spaces! (In all of the cases I've told you exactly what the function you need to look at is; what's left to check is that it's actually an isomorphism).