LECTURE 13

Quotient Spaces

In all the development above we have created examples of vector spaces primarily as subspaces of other vector spaces. Below we'll provide a construction which starts with a vector space V over a field $\mathbb{F}$ and a subspace S of V, and which furnishes with an entirely new vector space from V/S which is particularly prominent in applications. In so doing, it will be absolutely vital to think of the new vector space as a vector space in the abstract sense; a set endowed with a notion of scalar multiplication and vector addition satisfying certain axioms. For the individual "vectors" in the vector space V/S will not be vectors in V, but rather large families of vectors in V.

1. Digression: Modular Arithmetic

In the hopes of making our discussion of quotient spaces of vector spaces more digestible, I'm going to make a brief digression into modular arithmetic.

Theorem 13.1. *(The Division Algorithm) Let n be an integer > 0 and let z be any other integer. Then there exists unique integers r and q such that*

(i) $z = qn + r$
(ii) $0 \leq r < n$

Even though this is the basic fact underlying the long division algorithm you've used since elementary school, its proof is moderately sophiscated (involving the Well Ordering Axiom of the positive integers at one point). Even so, the basic idea is simple. The theorem says, for example, if you take $z = 23$ and $n = 5$, then since

$$23 = 4 \cdot 5 + 3 \qquad (*)$$

and because $0 \leq 3 < 5$, and this is the only way of writing 23 as a multiple of 5 plus an integer remainder that's between 0 and 5.

By virtue of the Division Algorithm, once we fix the divisor n, we have have exactly n possibilities for the remainder r; viz., $r = 0, 1, \ldots, n-1$. Thus, we can break up the set of integers into n distinct, disjoint families, depending on their remainders by n:

$$\begin{aligned}
[0]_n \quad &: \quad = \{z \in \mathbb{Z} \mid z = qn + 0 \quad \text{for some } q \in \mathbb{Z}\} \\
[1]_n \quad &: \quad = \{z \in \mathbb{Z} \mid z = qn + 1 \quad \text{for some } q \in \mathbb{Z}\} \\
[2]_n \quad &: \quad = \{z \in \mathbb{Z} \mid z = qn + 2 \quad \text{for some } q \in \mathbb{Z}\} \\
&\ \ \vdots \\
[n-1]_n \quad &: \quad = \{z \in \mathbb{Z} \mid z = qn + n - 1 \quad \text{for some } q \in \mathbb{Z}\}
\end{aligned}$$

Now let $\mathbb{Z}_n := \{[0]_n, [1]_n, \cdots, [n-1]_n\}$. Note each element of $\mathbb{Z}_n$ is an infinite set of integers and that we have

$$\mathbb{Z} = \coprod_{k=0}^{n-1} [k]_n \qquad \text{(disjoint union of sets)}$$

The situation when $n = 2$ is quite familiar

$$\begin{array}{rcl}[0]_2 & : & = \{z \in \mathbb{Z} \mid z = 2q + 0 \quad \text{for some } q \in \mathbb{Z}\} = \text{ the set of even integers} \\ [1]_2 & : & = \{z \in \mathbb{Z} \mid z = 2q + 1 \quad \text{for some } q \in \mathbb{Z}\} = \text{ the set of odd integers}\end{array}$$

and

$$\mathbb{Z} = [0]_2 \cup [1]_2 = \text{the set of even integers} \cup \text{ the set of odd integers}$$

Recall the following simple rules for working with even and odd numbers;

$$\begin{array}{rcl}\text{(an even integer)} + \text{(an even integer)} & = & \text{(an even integer)} \\ \text{(an even integer)} + \text{(an odd integer)} & = & \text{(an odd integer)} \\ \text{(an odd integer)} + \text{(an odd integer)} & = & \text{(an even integer)}\end{array}$$

$$\begin{array}{rcl}\text{(an even integer)} * \text{(an even integer)} & = & \text{(an even integer)} \\ \text{(an even integer)} * \text{(an odd integer)} & = & \text{(an even integer)} \\ \text{(an odd integer)} * \text{(an odd integer)} & = & \text{(an odd integer)}\end{array}$$

Since these rules do not depend on the particular even or odd integers we use on the left hand side, we can think of these rules as applying to families of even and odd integers; thus, one writes

$$\begin{array}{rcl}[0]_2 + [0]_2 & = & [0]_2 \\ [0]_2 + [1]_2 & = & [1]_2 \\ [1]_2 + [1]_2 & = & [0]_2\end{array}$$

$$\begin{array}{rcl}[0]_2 * [0]_2 & = & [0]_2 \\ [0]_2 * [1]_2 & = & [0]_2 \\ [1]_2 * [1]_2 & = & [1]_2\end{array}$$

These rules thus define a certain arithmetric for the families of even and odd integers. But I stress here we are not really adding and multiplying individual integers, rather we are developing more abstract arithmetic rules that can applied to the family of even integers and the family of odd integers.

For more general n, we would like to have corresponding "arithmetics" defined on $\mathbb{Z}_n$. For this purpose, we would like to fill in the slots of the following addition and multiplication tables

$+$	$[0]_n$	$[1]_n$	$\cdots$	$[n-1]_n$
$[0]_n$	?	?	$\cdots$	?
$[1]_n$	?	?		?
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$[n-1]_n$	?	?		?

,

$*$	$[0]_n$	$[1]_n$	$\cdots$	$[n-1]_n$
$[0]_n$	?	?	$\cdots$	?
$[1]_n$	?	?		?
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$[n-1]_n$	?	?		?

The algorithm by which we'll fill in the tables is based on the following facts

- Suppose a has remainder r_1 when divided by n and b has remainder r_2 when divided by n, then $a + b$ has same remainder as $r_1 + r_2$.
- Suppose b has remainder r_1 when divided by n and b has remainder r_2 when divided by n, then $a * b$ has the same remainder as $r_1 * r_2$.

These facts lead to the following rules of addition and multiplication in $\mathbb{Z}_n$

$$[a]_n + [b]_n = [a + b]_n \tag{1}$$

$$[a]_n * [b]_n = [a * b]_n \tag{2}$$

Thus, for example in $\mathbb{Z}_3$, we have

$$[2]_3 * [2]_3 = [2 * 2]_3 = [1]_3 \quad \text{since the remainder of } 2 * 2 = 4 \text{ when divided by 3 is 1}$$

The rules (1) and (2) above thus provide a means of filling in the addition and multiplication tables for $\mathbb{Z}_n$, and thus provide a certain *arithmetic structure* for families $[0]_n, [1]_n, \ldots, [n-1]_n$ in $\mathbb{Z}_n$.

A bit more conceptionally, we started with a nice set $\mathbb{Z}$ with its natural arithmetic and from that set up other sets $\mathbb{Z}_n$ (consisting of certain subsets of $\mathbb{Z}$) and used the arithmetic in $\mathbb{Z}$ to define an arithmetic on $\mathbb{Z}_n$.

What we are going to do next is very similar, except instead of trying to define the arithmetic operations of addition and multiplication for families of integers, we will try to develop rules of scalar multiplication and vector addition for families of vectors.

2. Quotient Spaces

Suppose S is a subspace of a finitely generated vector space V over a field $\mathbb{F}$ and a is a vector in V. Let

$$[a]_S = \{v \in V \mid v = a + s \quad \text{for some } s \in S\} \tag{3}$$

This is a subset of V, but in general it is not a subspace of V. In fact, it is only a subspace of V when v itself is in S, and in this case $S_v = S$. Note that, geometrically,

$$[a]_S = \text{hyperplane through } a \text{ generated by the directions in } S$$

For this reason, it is far more common to denote the sets $[a]_S$ as $a + S$ (as in Lecture 9). Indeed we will adopt the latter notation shortly. However, in order to stress the analogy with modular arithmetic, we'll continue with the notation (3) until we get the basic definition/construction of quotient spaces completed.

Lemma 13.2. *If $b \in [a]_S$, then $a \in [b]_S$ and in fact, $[a]_S = [b]_S$.*

Proof. If $b \in [a]_S$, then by definition there is some $s \in S$ such that $b = a + s$. But then $a = b - s$. Since S is a subspace $s \in S$ implies $-s \in S$ and so

$$a = b - s = b + (-s) \quad \Rightarrow \quad a \in [b]_S$$

Now suppose $b \in [a]_S$. Let c be an arbitary element of $[b]_S$. We have

$$c = b + s \quad \text{for some } s \in S$$

On the other hand, by hypothesis, $b \in [a]_S$ and so

$$b = a + s' \quad \text{for some } s' \in S$$

Thus,

$$c = a + s' + s = a + (s + s') \quad \Rightarrow \quad c \in [a]_S \quad .$$

Thus, if $b \in [a]_S$, then every element of $[b]_S$ is in $[a]_S$. On the other hand, by the first part of the lemma, if $b \in [a]_S$, then $a \in [b]_S$ and so, by the discussion just preceding, every element of $[a]_S$ is also in $[b]_S$. Thus,

$$[a]_S = [b]_S$$

□

Lemma 13.3. *Two vectors v_1, v_2 belong to the same $[a]_S$ if and only if $v_1 - v_2 \in S$.*

Proof.

$\Rightarrow$ Suppose $v_1, v_2 \in [a]_S$. Then by definition, there must be vectors s_1 and s_2 such that

$$\begin{aligned} v_1 &= a + s_1 \\ v_2 &= a + s_2 \end{aligned}$$

But then

$$v_1 - v_2 = (a + s_1) - (a - s_2) = s_1 - s_2 \in S \quad .$$

$\Longleftarrow$ Suppose $v_1 - v_2 \in S$; in fact, say $v_1 - v_2 = s \in S$.

$$v_1 = v_2 + s \qquad \Rightarrow \quad v_1 \in [v_2]_S$$

Then by the preceding lemma

$$[v_1]_S = [v_2]_S$$

Now suppose $v_1 \in [a]_S$, then by the preceding lemma $a \in [v_1]_S$. But $[v_1]_S = [v_2]_S$ and so $a \in [v_2]_S$. Thus, the statement is proved. □

Theorem 13.4. *Let V be a vector space over a field $\mathbb{F}$ and let S be a subspace of V. Let*

$$V/S = \{[v]_S \mid v \in V\}$$

Define operators of scalar multiplication and vector addition on V/S as follows:

$$\begin{aligned} [v]_S + [v']_S &= [v + v']_S && (5) \\ \lambda \cdot [v]_S &= [\lambda v]_S && (6) \end{aligned}$$

Then with these operations V/S has the structure of a vector space over $\mathbb{F}$.

Proof.

First let me demonstrate that the formulas fof addition and scalar multiplication are self-consistent. Recall that $[v]_S$ and $[v']_S$ are not individual vectors, but rather infinite sets of vectors. As such we can not add them directly. The rule (5) gives us, nevertheless, with a means for taking two input hyperplanes and coming up with a corresponding "sum". However, vector space addition has to be an actual function - meaning given two inputs $[v]_S$ and $[v']_S$ there has to be a unique output. On other hand, since

$$[v_1]_S = [v_2]_S \quad \text{whenever } v_1 - v_2 \in S$$

we have to make sure that, e.g.,

$$[v]_S + [v_1]_S = [v]_S + [v_2]_S \quad \text{whenever } v_1 - v_2 \in S \quad .$$

In other words, we need to make sure that the vector addition that takes place inside the brackets on the right hand side of (5) is independent of the vectors v and v' we choose to represent the subsets $[v]_S$ and $[v']_S$. Suppose instead we replaced v by $v + s_1$ and v' by $v' + s_2$ in (5), Since

$$\begin{aligned} [v]_S &= [v + s_1]_S \qquad (s_1 \in S) \\ [v']_S &= [v' + s_2]_S \qquad (s_2 \in S) \end{aligned}$$

we should get the same subset on the right hand side when we add $[v + s_1]_S$ to $[v' + s_2]_S$. According to the rule (5)

$$[v + s_1]_S + [v' + s_2]_S = [v + s_1 + v' + s_2]_S = [v + v' + (s_1 + s_2)]_S = [v_1 + v_2]_S \, .$$

And so the result of applying rule (5) doesn't depend on how we represent the vectors in $[v]_S$ and $[v']_S$ in order to carry out the calculation.

Similarly,

$$\lambda \cdot [v + s]_S := [\lambda \cdot (v + s)]_S = [\lambda v + \lambda s]_S = [\lambda v]_S \quad \text{since } \lambda s \in S$$

To see that V/S is actually a vector space over $\mathbb{F}$, we need to confirm the 8 axioms for a vector space. This is tedious but pretty routine. I'll just point out a couple sample computattions here.

- Associativity of vector addition

$$([a]_S + [b]_S) + [c]_S = [a + b]_S + [c]_S = [(a + b) + c]_S = [a + (b + c)]_S = [a]_S + [b + c]_S = [a]_S + ([b]_S + [c]_S)$$

- Distributivity of scalar multiplication over vector addition

$$\lambda \cdot ([a]_S + [b]_S) = \lambda \cdot [a + b]_S = [\lambda \cdot (a + b)]_S = [\lambda a + \lambda b]_S = [\lambda a]_S + [\lambda b]_S = \lambda \cdot [a]_S + \lambda \cdot [b]_S$$

- Existence of additive inverse. Consider the subset $[\mathbf{0}_V]_S$. We have, for any $[a]_S$ in V/S
$$[\mathbf{0}_V]_S + [a]_S = [\mathbf{0}_V + a]_S = [a]_S$$
and so $[\mathbf{0}_V]_S$ acts like an additive identity in V/S.

The other 5 axioms are verified in a similarly easy fashion. □

Let me now pull these results all together with a definition.

Definition 13.5. *Suppose V is a vector space over a field $\mathbb{F}$ and S is a subspace of V. Then*
$$V/S \equiv \{[v]_S \mid v \in V\}$$
is the vector space over $\mathbb{F}$ where the operations of scalar multiplication and vector addition are defined by
$$\begin{aligned} \lambda \cdot [v]_S &= [\lambda v]_S \\ [v]_S + [v']_S &= [v + v']_S \end{aligned}$$
and the additive identity is given by
$$\mathbf{0}_{V/S} = [\mathbf{0}_V]_S$$

Now to some of you, this construction of a vector space over $\mathbb{F}$ might appear bizarre. In some sense, this is a good thing, as I have been stressing all along the conventional view of vectors as lists of numbers is far from typical. But this construction is also pretty fundamental to what we might call *advanced linear algebra*. This notion will be particularly useful when we try to get a better understanding of the image of a linear transformations.

Lemma 13.6. *Let V be a finitely generated vector space over a field a field $\mathbb{F}$ and let S be a subspace of V. Then the map*
$$p_S : V \to V/S \quad : \quad \mathbf{v} \longmapsto [\mathbf{v}]_S$$
is a surjective vector space homomorphism with kernel S.

Proof: p_S is surjective simply by virtue of the fact that its codoman V/S is already defined as the image of the map $\mathbf{v} \longmapsto [\mathbf{v}]_S$. It is a linear transformation because

$$\begin{aligned} p_S(\alpha\mathbf{v} + \beta\mathbf{u}) &= [\alpha\mathbf{v} + \beta\mathbf{v}]_S \\ &= [\alpha\mathbf{v}]_S + [\beta\mathbf{u}]_S \qquad \text{by the defnition of vector addition in } V/S \\ &= \alpha[\mathbf{v}]_S + \beta[\mathbf{u}]_S \qquad \text{by the defnition of scalar multiplication in } V/S \\ &= \alpha p_S(\mathbf{v}) + \beta p_S(\mathbf{u}) \quad \text{by the definition of } p_S. \end{aligned}$$

Finally, $v \in \ker(p_S)$ means
$$[v]_S = [\mathbf{0}_V]_S \equiv \{v \in V \mid v = \mathbf{0}_V + s \text{ for some } s \in S\} = \{v \in V \mid v \in S\}$$
In other words, $v \in \ker(p_S) \quad \Rightarrow \quad v \in S$. □

Remark 13.7. Given a subspace S of a vector space V, the linear transformation $p_S : V \to V/S$ defined above is often referrred to as the *canonical projection* (of V onto V/S).

Lemma 13.8. *Let V be a finitely generated vector space over a field a field $\mathbb{F}$ and let S be a subspace of V. Then V/S is finitely generated and* $\dim(V/S) = \dim(V) - \dim(S)$.

Proof. Let p_S be the canonical projection of V onto S. p_S is obviously surjective by construction. Let $B = \{b_1, \ldots, b_n\}$ be a basis for V. Since p_S is surjective, each $[v]_S \in V/S$ is the image of some $v \in V$ by p_S.

$$[v]_S \ = [\alpha_1 b_1 + \cdots + \alpha_n b_n]_S = \alpha_1 [b_1]_S + \cdots + \alpha_n [b_n]_S$$

hence V/S is generated by $\{p_S [b_1], \ldots, p_S [b_n]\}$.

Next, we note that $\ker p_S = S$. To see this, suppose first that $s \in S \subset V$. Then

$$p(s) = [s]_S = [0 + s]_S = [\mathbf{0}_V]_S = \mathbf{0}_{V/S}$$

so $s \in \ker(p_s)$. On the other hand, suppose $v \in \ker(p_S)$. Then

$$\mathbf{0}_{V/S} = [\mathbf{0}_V]_S = [v]_S = p(v) \quad \Rightarrow \quad v = \mathbf{0}_V + s \text{ for some } s \in S \quad \Rightarrow \quad v \in S.$$

So now we have seen (see Corollary 11.7 of Lecture 11) that whenever we have a vector space homomorphism $f : V \to W$, $\dim(V) = \dim(\ker(f)) + \dim(\operatorname{Im}(f))$. In that case at hand we have

$$\begin{aligned} \dim(V) &= \dim(\ker(p_S)) + \dim(\operatorname{Im}(p_S)) \\ &= \dim(S) + \dim(V/S) \end{aligned}$$

and the statement follows. □

The notion of quotient spaces is particularly useful in the setting of infinite-dimensional vector spaces. Moreover, proofs of the most important properties of quotient spaces can be formulated without reference to finite bases, and so are valid even in the setting infinite-dimensional vector spaces. We'll do this below and in the next lecture. However, let me now adopt more standard notation for the elements of a quotient space; writing $v + S$ instead of $[v]_S$ for the elements of a quotient space V/S

$$v + S \equiv \{v + s \mid s \in S\} = S\text{-hyperplane through } v$$

We'll continue to denote the canonical projection from V to V/S by p_S

$$p_S : V \to V/S \quad , \quad v \longmapsto v + S \quad .$$

THEOREM 13.9. *Let T be a subspace of subspace S of a vector space V. Then S/T may be regarded as a subspace of V/S. Moreover, there is a one-to-one correspondence between the set of all subspaces of V containing T and the set of all subspaces of S/T.*

Proof. We have

$$\begin{aligned} V/T &= \{v + T \mid v \in V\} \\ S/T &= \{s + T \mid s \in S\} \end{aligned}$$

So clearly, $S/T \subset V/T$, since the condition for membership in S/T is just a restricted version of the condition for membership in V/T. Since S/T is the image of a vector space homomorphism it is naturally closed under scalar multiplication and vector addition. Moreover, scalar multiplication and vector addition in S/T are just the restrictions of the same operations applied in V/T. Hence, S/T is a subspace of $V/T/$

Now suppose X is a subspace of V/T and let $p_T : V \to V/T$ be the canonical projection of V onto V/T. Since p_T is a linear transformation

$$p_T^{-1}(X) := \{v \in V \mid p_T(v) \in X\}$$

is a subspace of V (see Theorem 10.8). Thus, p_T^{-1} maps subsets of V/T to certain subspaces of V. Now each subset X of V/T contains the additive identity of V/T :

$$\mathbf{0}_{V/T} = \mathbf{0}_V + T$$

Now, on the one hand, since $\ker p_T = T$, we have

$$p_T^{-1}\left(\mathbf{0}_{V/T}\right) = T$$

while, on the other, since each subspace X of V/T contains $\mathbf{0}_{V/T}$, we have

$$S = p_T^{-1}\left(\mathbf{0}_{V/T}\right) \subset p_T^{-1}\left(X\right)$$

Thus, p_T^{-1} maps subspaces of V/T to subspaces of X containing T.

Now we know p_T maps every subspace of V containing T to a subspace of V/T and that p_T^{-1} maps every subspace to V/T to a subspace of V containing T. Moreover, by construction,

$$\begin{aligned} p_T \circ p_T^{-1} &= \text{ identity map on subspaces of } V/T \\ p_T^{-1} \circ p_T &= \text{ identity map on subspaces of } V \text{ containing } T \end{aligned}$$

Hence, we have a bijection between the two sets of subspaces and so the one-to-one correspondence of the theorem statement. □