Axler Slick or Sloppy

John McCuan

January 25, 2022

Lemma 1 *Given a finite collection of vectors* $\{v_1, v_2, \ldots, v_k\}$ *in a vector space* V*, the following are equivalent:*

1. $\{v_1, v_2, \ldots, v_k\}$ *is a linearly independent set.*

2. $v_1 \neq \mathbf{0}$ *and for each* $\ell = 2, \ldots, k$,

$$v_\ell \notin \operatorname{span}\{v_1, \ldots, v_{\ell-1}\}. \tag{1}$$

This is a restatement of Axler's Lemma 2.21 which he calls the **Linear Dependence Lemma**. I would like to make some comments about this lemma and give a careful proof of it.

1 comments

First of all, I really like this lemma. I've known about the concept of liner independence for over thirty years, and I've never really thought of this "formulation" before. You have two concepts concerning, or more properly two conditions on, a collection of vectors $\{v_1, v_2, \ldots, v_k\}$ in a vector space. The first is one with which I'm familiar:

> If the zero vector $\mathbf{0}$ is expressed as a linear combination of the vectors in your set:
>
> $$\mathbf{0} = \sum_{j=1}^{k} a_j v_j,$$
>
> then all the scalars $a_j \in F$ (the field) must vanish, i.e., must be $0 \in F$.

The second one is given in the lemma:

Each vector in the list is not in the span of the vectors before it.

The two conditions, seemingly, are quite different, but the lemma asserts that they are saying precisely the same thing. I think it's pretty cool, and Axler uses his condition as a kind of cornerstone for his discussion of vector spaces and the dimension of finite dimensional vector spaces in particular. It should perhaps be emphasized, therefore, that we are in the second or third week of the course, and if a student is attempting to follow Axler's presentation of linear algebra, but doesn't fully understand this lemma, then that student won't fully understand much of the rest of the course. That's sort of scary, but maybe not uncommon.

I do not know for sure, but since I've read many linear algebra books, and many of the same ones Axler has read, I'll guess that the formulation of this lemma is his own—he probably came up with it by himself. And that's also cool. I'm glad to have come across it in my old age. In view of these warm sentiments, I suggest that from now on we elevate this result (and Herr Doctor Professor Axler himself) by using the name **Axler's Lemma**.

I can also point out, though maybe it would make more sense to do so after giving the proof, that the equivalence of these two formulations of linear independence is, though quite simple, a bit subtle. So the new formulation really seems to give you something new—a new way of looking at linear independence.

Axler's statement is not quite the same as mine. And there is something to nitpick about in his statement. You'll notice that he doesn't say anything about a vector (the first vector) being nonzero. Of course, he takes linear **dependence** as the base of his statement rather than linear independence. As a result he ends up considering **some** ℓ for which

$$v_\ell \in \operatorname{span}\{v_1, \ldots, v_{\ell-1}\}.$$

Axler might have been trying to get around this irritating business with the first vector by this "for some" assertion as opposed to the "for all" assertion (1). But it really doesn't help. To see this, consider first what happens if $v_1 = \mathbf{0}$. You can even extend the condition (1) to $\ell = 1$ and it sort of becomes vacuously true, but it still doesn't get you an equivalence. The basic problem is that there are no vectors preceeding the first vector, and condition (1) is really only saying something if the collection $\{v_1, \ldots, v_{\ell-1}\}$ makes sense, meaning in this case that it is also a nonempty collection. Notice that it's also sort of an assumption of the lemma that $\{v_1, v_2, \ldots, v_k\}$ is a nonempty set. There may actually be no vector v_2, but we've just written that to be suggestive. Also, it may be that $k = 1$. We could also have said explicitly that $\{v_1, v_2, \ldots, v_k\}$ is a **nonempty** collection of vectors. That wouldn't be much trouble, but sometimes all the little details required to be precisely precise are a bit distracting and will drive normal people crazy. Of course, Mathematicians typically love that sort of thing—which explains a lot.

On the other hand, it is not uncommon that attention to those nitpicky details actually reveals a fundamental error in thinking. That's probably why the normal people keep the mathematicians around. Returning to Axler's "reverse" formulation, if the first vector $v_1 = \mathbf{0}$ is the zero vector, then certainly $\{v_1, v_2, \ldots, v_k\}$ is a linearly dependent set, but Axler's condition in this case $\ell = 1$,

$$v_\ell \in \operatorname{span}\{v_1, \ldots, v_{\ell-1}\},$$

doesn't really make sense. There is probably a natural way to interpret it: The set on the right should be interpreted to be the empty set. And Axler may try to get around this by declaring the span of the empty set to be the subspace containing only the zero vector $\{\mathbf{0}\}$. But I declare this to be just silly. A **span** in general is a collection of linear combinations, and you can't get anything by taking linear combinations of vectors from the empty set. The zero vector is not "nothing." It's a vector, so you can't get it. I refuse.

Exercise 1 *Write down carefully Axler's* **Linear Dependence Lemma***, and fix what he has by adding a condition concerning the first vector.*

Exercise 2 *Go look through Axler's book and see if he has defined the span of the empty set to be* $\{\mathbf{0}\}$*. (I don't have time for such nonsense.)*

2 Proof

Okay, I'm going to prove two assertions. First I'm going to assume $\{v_1, v_2, \ldots, v_k\}$ is linearly independent and show that statement **2** holds. First of all $v_1 \neq \mathbf{0}$ or else $7v_1 = 2v_1 = \mathbf{0}$ and there are all kinds of ways to write the zero vector as a linear combination of the vectors in $\{v_1, v_2, \ldots, v_k\}$. Second, if

$$v_\ell \in \operatorname{span}\{v_1, \ldots, v_{\ell-1}\}$$

for some $\ell > 1$, then we can write

$$v_\ell = \sum_{j=1}^{\ell-1} a_j v_j$$

and

$$\mathbf{0} = \sum_{j=1}^{\ell} a_j v_j \qquad \text{with} \qquad a_\ell = -1 \neq 0 \in F.$$

This contradicts linear independence, so (1) must hold. This finishes the proof of the first assertion.

For the second assertion, we assume statement **2** holds. Now, if $\{v_1, v_2, \ldots, v_k\}$ is not linearly independent, there are some coefficients $a_1, a_2, \ldots, a_k$, not all zero for which

$$\mathbf{0} = \sum_{j-1}^{k} a_j v_j.$$

Here is where the subtle part comes in: We can take a_ℓ to be the nonzero coefficient with the **highest index**. This means $a_{\ell+1} = a_{\ell+2} = \ldots = a_k = 0$, and

$$\mathbf{0} = \sum_{j=1}^{\ell} a_j v_j \qquad \text{with} \qquad a_\ell \neq 0 \in F.$$

Therefore, we find

$$v_\ell = \sum_{j=1}^{\ell-1} \left(-\frac{a_j}{a_\ell} \right) v_j \in \operatorname{span}\{v_1, \ldots, v_\ell\},$$

unless $\ell = 1$. In this case, we get $a_1 v_1 = \mathbf{0}$ with $a_1 \neq 0 \in F$. And we know this means $v_1 = \mathbf{0}$. Either way then, we get a contradiction of statement **2**. $\square$