LINEAR ALGEBRA: DIMENSION AND THE STEINITZ EXCHANGE TRICK

LEO GOLDMAKHER

In class we made the following

Definition. The *dimension* of a vector space V is the number of elements in a basis of V.

There's something fishy about this definition. We've seen that a vector space might have more than one basis. Thus, for the notion of *dimension* to be well-defined, we need to know that any two bases of a given vector space have the same number of elements. This motivates us to prove:

Theorem 1. *Any two bases of a vector space V have the same number of elements.*

We will deduce this from the following

Lemma 2. *Any spanning set of a vector space is at least as large as any linearly independent set.*

The proof of this lemma is quite clever. We postpone it temporarily, and instead show how to apply it to give a short proof of the theorem.

Proof of Theorem 1. Let V be a vector space, and pick any two bases $\mathcal{B}_1$ and $\mathcal{B}_2$; it suffices to prove these two have the same number of elements. By the Fundamental Property of Bases, both $\mathcal{B}_i$'s span V and are linearly independent. In particular, $\mathcal{B}_1$ spans V and $\mathcal{B}_2$ is linearly independent, and Lemma 2 implies that

$$\#\mathcal{B}_1 \geq \#\mathcal{B}_2.$$

On the other hand, since $\mathcal{B}_2$ spans V and $\mathcal{B}_1$ is linearly independent, Lemma 2 implies

$$\#\mathcal{B}_2 \geq \#\mathcal{B}_1.$$

Thus $\#\mathcal{B}_1 = \#\mathcal{B}_2$ as claimed. □

Having dispensed with Theorem 1, we can now focus on Lemma 2. The proof rests on a trick which is usually called the *Steinitz Exchange Trick*: given a spanning set $\mathcal{S}$ and a linearly independent set $\mathcal{L}$, we substitute vectors from $\mathcal{L}$ into $\mathcal{S}$ one at a time without affecting the spanning property of $\mathcal{S}$. Either we run out of vectors from $\mathcal{L}$ to sub in before exhausting $\mathcal{S}$, or else we completely replace $\mathcal{S}$ by elements from $\mathcal{L}$. But then we have a set of linearly independent vectors which span the space, meaning there can't be any more vectors left over in $\mathcal{L}$! Let's make this argument more precise.

Proof of Lemma 2. Let V be a vector space. Suppose

$$\mathcal{S} := \{\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_m\} \subseteq V$$

spans V, and that

$$\mathcal{L} := \{\vec{\ell}_1, \vec{\ell}_2, \ldots, \vec{\ell}_k\} \subseteq V$$

is a set of linearly independent vectors. Our goal is to prove that $\mathcal{S}$ is at least as large as $\mathcal{L}$, i.e., that $m \geq k$.

Since $\mathcal{S}$ spans V, we can express $\vec{\ell}_1$ as a linear combination of the elements of $\mathcal{S}$:

$$\vec{\ell}_1 = \alpha_1\vec{s}_1 + \alpha_2\vec{s}_2 + \cdots + \alpha_m\vec{s}_m.$$

Note that at least one of the α_i's must be nonzero. (Why is this?) Without loss of generality, we may assume $\alpha_1 \neq 0$. It follows that

$$\vec{s}_1 = \frac{1}{\alpha_1}\vec{\ell}_1 - \frac{\alpha_2}{\alpha_1}\vec{s}_2 - \cdots - \frac{\alpha_m}{\alpha_1}\vec{s}_m. \tag{1}$$

Date: Discussed in Lecture 28 (April 25, 2016).

We conclude that $\{\vec{\ell}_1, \vec{s}_2, \ldots, \vec{s}_m\}$ spans V: any vector can be written as a linear combination of the $\vec{s}_i$'s, and (1) tells us that we can replace $\vec{s}_1$ by a linear combination of $\vec{\ell}_1$ and the remaining $\vec{s}_i$'s.

Now we iterate this procedure. Since $\{\vec{\ell}_1, \vec{s}_2, \ldots, \vec{s}_m\}$ spans V, we must be able to write $\vec{\ell}_2$ as a linear combination of $\{\vec{\ell}_1, \vec{s}_2, \ldots, \vec{s}_m\}$:

$$\vec{\ell}_2 = \beta_1\vec{\ell}_1 + \beta_2\vec{s}_2 + \beta_3\vec{s}_3 + \cdots + \beta_m\vec{s}_m.$$

One of the coefficients $\beta_2, \beta_3, \ldots, \beta_m$ must be nonzero (why?). WLOG say $\beta_2 \neq 0$. Then we can solve for $\vec{s}_2$:

$$\vec{s}_2 = -\frac{\beta_1}{\beta_2}\vec{\ell}_1 + \frac{1}{\beta_2}\vec{\ell}_2 - \frac{\beta_3}{\beta_2}\vec{s}_3 - \cdots - \frac{\beta_m}{\beta_2}\vec{s}_m$$

which shows that $\{\vec{\ell}_1, \vec{\ell}_2, \vec{s}_3, \ldots, \vec{s}_m\}$ spans V.

Now we have two possibilities: either $k < m$, or $k \geq m$. In the former case, we're done with the proof! If the latter holds, then (continuing the above process) we deduce that $\{\vec{\ell}_1, \vec{\ell}_2, \ldots, \vec{\ell}_m\}$ spans V. But this tells us that any $\vec{\ell}_j$ with $j > m$ could be written as a linear combination of $\{\vec{\ell}_1, \vec{\ell}_2, \ldots, \vec{\ell}_m\}$, which contradicts the linear independence of the set $\mathcal{L}$. Thus, we conclude that in this case $k = m$, and we're done with the proof in this case as well. $\square$