Notes

Here is a review of the "exchange" proof that two finite bases of a vector space must have the same size.

This proof is important because it can be done directly from the basic definition of a vector space instead of requiring facts about matrices.

If you are asked to do this proof, it is Lemma 2 that is expected.

Let V be a vector space.

Lemma 1. If $v_1, \ldots, v_m$ are linearly dependent, then some v_i is in the span of the preceding vectors $v_1, \ldots, v_{i-1}$.

Proof. If $r_1 v_1 + \cdots + r_m v_m = \mathbf{0}$, with $r_1, \ldots, r_m$ not all 0, let i be the highest index with $r_i \neq 0$. Then

$$v_i = \left(-\frac{r_1}{r_i}\right) v_1 + \cdots + \left(-\frac{r_{i-1}}{r_i}\right) v_{i-1} + \mathbf{0} + \cdots + \mathbf{0}.$$

Remark. This was simple, but in the theory of vector spaces this is where we need the the field property that every nonzero field element has a multiplicative inverse. If we tried to use $\mathbb{Z}$ for scalars, Lemma 1 would fail.

Corollary. [not needed for the proof of the theorem] If none of $v_1, \ldots, v_n$ is in the span of the preceding vectors, then this list is linearly independent

Proof. This is just the contrapositive of the statement of Lemma 1.

Note. To say that v_1 is not in the span of the preceding vectors means that $v_1 \neq \mathbf{0}$, because the span of no vectors is $\{\mathbf{0}\}$.

Lemma 2. If V is spanned by $v_1, \ldots, v_n$ and if $w_1, \ldots, w_m$ are linearly independent in V, then $m \leq n$.

Proof. We keep making new spanning sets by "exchanging" $w_1, \ldots, w_m$ for $v_1, \ldots, v_n$ one vector at a time, as follows.

Start with

$v_1, \ldots, v_n,$

which span V. Now consider the list

$w_1, v_1, \ldots, v_n.$

Since $w_1 \in \operatorname{span}(v_1, \ldots, v_n)$ this new list is linearly dependent. By Lemma 1, some vector is in the span of the preceding vectors. It can't be w_1 since

$w_1 \neq \mathbf{0}$ (being in a linearly independent set of vectors). So it is v_i for some i, say i_1. Then we can remove v_{i_1} from the list and still have a spanning set:

$$w_1, \underbrace{v_1, \ldots, v_n}_{\text{omit } v_{i_1}}.$$

Now consider the list

$$w_2, w_1, \underbrace{v_1, \ldots, v_n}_{\text{omit } v_{i_1}}.$$

Again, since w_2 was in the span of the preceding list, the new list is linearly dependent and some vector is in the span of the preceding. Again, it can't be w_2 or w_1 because w_2, w_1 are linearly independent. Therefore we can remove another vector v_{i_2} from the list and still have a spanning set:

$$w_2, w_1, \underbrace{v_1, \ldots, v_n}_{\text{omit } v_{i_1}, v_{i_2}}.$$

We continue in this way, putting in w's at the left and taking out v's. If $m > n$ then we will run out of v's and have a spanning set consisting of some but not all of $w_1, \ldots, w_m$, which is impossible because no w_j is in the span of the rest. This is a contradiction, so $m \leq n$.

Remarks. (1) It might be neater to throw in w_m first instead of w_1, then w_{m-1}, etc., so $w_1, \ldots, w_m$ appear in the original order.

(2) This is really an informal proof by induction. We may discuss proof by induction later.

(3) Notation for this proof is a special problem. In describing the list of $w_1, v_1, \ldots, v_n$ with v_i omitted, we would usually say $w_1, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots v_n$. That's fine, but for omitting *two* vectors it doesn't work because we don't know which of v_{i-1}, v_{i-2} occurs first.

Still another approach is to say after each step that since the list still spans in any order, by re-indexing we can assume that it's v_n that is to be deleted, then v_{n-1}, etc. That's actually a neater way to say it, but re-indexing after every step seems harder to think about.

Theorem. ("invariance of dimension") If V has a basis of m elements and a basis of n elements, then $m = n$.

Proof. Since a basis is both linearly independent and a spanning set, Lemma 2 applies both ways around to show $m \geq n$ and $m \leq n$. Therefore $m = n$.

Definition. If V has a basis of n elements then we say V has dimension n.

The Theorem tells us that this definition is not ambiguous; there is only one possible dimension.