The Steinitz replacement lemma and its consequences

We wish to talk about the dimension of a vector space, as a measure of its size. However before we can do this we will need some technical results which will show that our definition makes sense. As the proofs of these can be a little bit confusing, we have collected them together in this handout.

First we consider a result from the last lecture:

Lemma 1.21 *(i) If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V then we can find a subset of them which form a basis (i.e. every finite spanning set contains a basis).*
(ii) Every non-zero finite-dimensional vector space has a basis.

Proof: (i) If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent then we are done. If not then there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ not all zero with $\sum \lambda_i \mathbf{v}_i = \mathbf{0}$. After renumbering we may assume that $\lambda_n \neq 0$. Then

$$(-\lambda_n)\mathbf{v}_n = \sum_{i=1}^{n-1} \lambda_i \mathbf{v}_i \quad \text{and so} \quad \mathbf{v}_n = \sum_{i=1}^{n-1} \left(\frac{-\lambda_i}{\lambda_n}\right) \mathbf{v}_i.$$

Now I claim that $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ span V. To see this note that for all $\mathbf{x} \in V$ we know that there exist $\mu_i \in \mathbb{F}$ such that $\mathbf{x} = \sum_{i=1}^n \mu_i \mathbf{v}_i$, as $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V. Therefore

$$\begin{aligned} \mathbf{x} &= \textstyle\sum_{i=1}^{n-1} \mu_i \mathbf{v}_i + \mu_n \sum_{i=1}^{n-1} \left(\frac{-\lambda_i}{\lambda_n}\right) \mathbf{v}_i \\ &= \textstyle\sum_{i=1}^{n-1} \left(\mu_i + \mu_n \left(\frac{-\lambda_i}{\lambda_n}\right)\right) \mathbf{v}_i \end{aligned}$$

And hence $\mathbf{x} \in \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$.
Now repeat this process until after $n - t$ steps we have t vectors $\mathbf{v}_1, \ldots, \mathbf{v}_t$ (possibly after renumbering) which span and are linearly independent. These vectors are thus a basis of V.
(ii) This follows from (i) and the definition of a finite-dimensional vector space. □

Next recall the final lemma from last time, which will be the key result in what follows.

Lemma 1.24 (Steinitz replacement lemma) *Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent vectors in some vector space V over $\mathbb{F}$, and that for some $0 \leq r \leq m-1$ and vectors $\mathbf{w}_i$ we have that $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{w}_{r+1}, \ldots, \mathbf{w}_n$ span V. (If $r = 0$ then there are no $\mathbf{v}_i$ in this list.) Then, possibly after renumbering the $\mathbf{w}_j$, we have that $\mathbf{v}_1, \ldots \mathbf{v}_{r+1}, \mathbf{w}_{r+2}, \ldots \mathbf{w}_n$ span V.*

Proof: Since $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{w}_{r+1}, \ldots, \mathbf{w}_n$ span V we must have that

$$\mathbf{v}_{r+1} = \sum_{i=1}^{r} \lambda_i \mathbf{v}_i + \sum_{i=r+1}^{n} \lambda_i \mathbf{w}_i. \tag{1}$$

If $\lambda_{r+1}, \ldots, \lambda_n$ are all zero then $\mathbf{v}_{r+1}$ would be a linear combination of $\mathbf{v}_1, \ldots \mathbf{v}_r$ which is not possible as $\mathbf{v}_1, \ldots \mathbf{v}_{r+1}$ are linearly independent. Therefore (after renumbering) we may assume that $\lambda_{r+1} \neq 0$. Then

$$\mathbf{w}_{r+1} = \lambda_{r+1}^{-1} \mathbf{v}_{r+1} + \sum_{i=1}^{r} (-\lambda_{r+1}^{-1} \lambda_i) \mathbf{v}_i + \sum_{i=r+2}^{n} (-\lambda_{r+1}^{-1} \lambda_i) \mathbf{w}_i$$

using (1).

Now for any $\mathbf{x} \in V$ we have that there exist $\mu_i \in \mathbb{F}$ with

$$\begin{aligned} \mathbf{x} &= \textstyle\sum_{i=1}^r \mu_i \mathbf{v}_i + \sum_{i=r+1}^n \mu_i \mathbf{w}_i \\ &= \textstyle\sum_{i=1}^r \mu_i \mathbf{v}_i + \sum_{i=r+2}^n \mu_i \mathbf{w}_i \\ &\qquad \textstyle+\mu_{r+1}(\lambda_{r+1}^{-1} \mathbf{v}_{r+1} + \sum_{i=1}^r (-\lambda_{r+1}^{-1} \lambda_i) \mathbf{v}_i + \sum_{i=r+2}^n (-\lambda_{r+1}^{-1} \lambda_i) \mathbf{w}_i) \\ &= \textstyle\sum_{i=1}^r (\mu_i - \mu_{r+1} \lambda_{r+1}^{-1} \lambda_i) \mathbf{v}_i + \mu_{r+1} \lambda_{r+1}^{-1} \mathbf{v}_{r+1} + \sum_{i=r+2}^n (\mu_i - \mu_{r+1} \lambda_{r+1}^{-1} \lambda_i) \mathbf{w}_i. \end{aligned}$$

Thus we can write any $\mathbf{x} \in V$ as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{r+1}, \mathbf{w}_{r+2}, \ldots, \mathbf{w}_n$, and hence these elements must span V. □

The Steinitz replacement lemma looks rather complicated. However in the remainder of this handout we will list some important consequences of it which are easily applied (for example, in examination questions!). These will often allow you to simplify various calculations. We begin with

Corollary 1.25 *If $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent in V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ span V then $n \geq m$.*

Proof: Suppose for a contradiction that $n < m$ and apply Lemma 1.24 n times. We start with $V = \text{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_n)$. We see that

$$\begin{aligned} V &= \text{Span}(\mathbf{v}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n) \\ &= \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{w}_n) \\ &= \cdots \\ &= \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \end{aligned}.$$

Therefore $\mathbf{v}_{n+1} = \sum_{i=1}^n \lambda_i \mathbf{v}_i$. But this contradicts the fact that $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$ are linearly independent (any subset of a linearly independent set is also linearly independent). Therefore $n \geq m$. □

From this we get

Corollary 1.26 *If V is a finite-dimensional vector space then every basis of V has the same number of elements.*

Proof: Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be two bases of V. As $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span and $\mathbf{w}_1, \ldots, \mathbf{w}_m$ are linearly independent we deduce from Corollary 1.25 that $n \geq m$. Reversing the roles of the two sets we can also deduce from the Corollary that $m \geq n$. Therefore $n = m$. □

We would like to be able to say that if $W \leq V$ then the dimension of W is no greater than the dimension of V. To prove this we will need

Lemma 1.30 *If $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are linearly independent and $\mathbf{v}$ is not in the span of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ then $\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{v}$ are linearly independent.*

Proof: Suppose that $\sum_{i=1}^{n} \lambda_i \mathbf{e}_i + \lambda_{n+1}\mathbf{v} = \mathbf{0}$ with not all λ_i zero. First note that $\lambda_{n+1} \neq 0$ as otherwise the $\mathbf{e}_i$ would be linearly dependent. Therefore we can write

$$\mathbf{v} = \sum_{i=1}^{n} (\frac{-\lambda_i}{\lambda_n})\mathbf{e}_i$$

and hence $\mathbf{v}$ is in the span of the $\mathbf{e}_i$, a contradiction. □

Corollary 1.31 *Any subspace W of a finite-dimensional vector space V is finite-dimensional and has dimension no greater than that of V. Further, if $\dim V = \dim W$ then $V = W$.*

Proof: Let V be finite-dimensional with basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $W \leq V$. If $W = \{\mathbf{0}\}$ then we are done. Otherwise take $\mathbf{e}_1 \in W \backslash \{\mathbf{0}\}$. If $W = \mathrm{Span}(\mathbf{e}_1)$ then we are done. If not then choose $\mathbf{e}_2$ not in the span of $\mathbf{e}_1$ in W: by Lemma 1.30 these two vectors are agin linearly independent. Repeating this process (which must terminate in at most n steps by Corollary 1.25), we eventually get a series $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$ which span W and are linearly independent (by Lemma 1.30). Therefore the set must be a basis of W and we know that $r \leq n$ by Corollary 1.25. Hence we deduce that $\dim W \leq \dim V$.

If $\dim W = \dim V$ but $W \neq V$ then there exists $\mathbf{v} \in V$ with $\mathbf{v} \notin W$. Another application of Lemma 1.30 implies that $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{v}$ are linearly independent in V, which contradicts Corollary 1.25. □

The following result is extremely useful for exam questions.

Corollary 1.32 *If $\dim V = n$ then*
(i) any n linearly independent vectors in V form a basis.
(ii) any n vectors which span V form a basis.

Proof: (i) Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be our set of independent vectors, and let $\mathbf{z}_1, \ldots, \mathbf{z}_n$ be a basis of V. Applying Lemma 1.24 n times we deduce that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V.
(ii) Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be our spanning set. If they do not form a basis of V then they must be linearly dependent, and there exist μ_i not all zero with $\sum \mu_i \mathbf{v}_i = \mathbf{0}$. We may assume that $\mu_n \neq 0$. But then $\mathbf{v}_n = \sum_{i=1}^{n-1}(-\mu_n^{-1}\mu_i)\mathbf{v}_i$ from which we may deduce that $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ span V. But this contradicts Corollary 1.25. □

Corollary 1.33 *Any linearly independent set of r vectors in a finite-dimensional vector space V of dimension $n > r$ can be extended to a basis of V.*

Proof: Exercise. □

The zero vector space

Finally, we include for completeness a proof that $\{\mathbf{0}\}$ is a vector space over $\mathbb{F}$ (this was Exercise 5 on the first question sheet). Recall that we defined a vector space structure on $V = \{\mathbf{0}\}$ by setting $\mathbf{u} + \mathbf{v} = \mathbf{0}$ for all $\mathbf{u}, \mathbf{v} \in V$ and $\lambda\mathbf{u} = \mathbf{0}$ for all $\mathbf{u} \in V$ and $\lambda \in \mathbb{F}$.

We need to check the 8 axioms for a vector space. In this case these are particularly simple.

(V1): $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{0}+\mathbf{w} = \mathbf{0} = \mathbf{u}+\mathbf{0} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(V2): $\mathbf{u}+\mathbf{v} = \mathbf{0} = \mathbf{v}+\mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in V$.

(V3): Take the zero element to be $\mathbf{0}$. Then $\mathbf{v}+\mathbf{0} = \mathbf{0}$ for all $\mathbf{v} \in V$.

(V4): Given $\mathbf{v} \in V$ take $(-\mathbf{v}) = \mathbf{0}$. Then $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$.

(V5): $\lambda(\mathbf{u}+\mathbf{v}) = \lambda\mathbf{0} = \mathbf{0} = \mathbf{0}+\mathbf{0} = \lambda\mathbf{0}+\lambda\mathbf{0} = (\lambda\mathbf{u})+(\lambda\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ and $\lambda \in \mathbb{F}$.

(V6): $(\lambda+\mu)\mathbf{v} = (\lambda+\mu)\mathbf{0} = \mathbf{0} = \mathbf{0}+\mathbf{0} = (\lambda\mathbf{0})+(\mu\mathbf{0}) = (\lambda\mathbf{v})+(\mu\mathbf{v})$ for all $\mathbf{v} \in V$ and $\lambda, \mu \in \mathbb{F}$.

(V7): $\lambda(\mu\mathbf{v}) = \lambda(\mathbf{0}) = \mathbf{0} = (\lambda\mu)\mathbf{0} = (\lambda\mu)\mathbf{v}$ for all $\mathbf{v} \in V$ and $\lambda, \mu \in \mathbb{F}$.

(V8): $1\mathbf{v} = \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.

This space has *no* basis, as we observed in Note 1.15(ii) that $\mathbf{0}$ can never be an element of a linearly independent set. We *define* the dimension of this vector space to be 0. Although this may look like a stupid vector space, we need it as it can occur in examples. For example, we saw in Example 1.8 that the set of solutions to a system of linear equations (when the RHS of each equals 0) is a vector space. This includes the possibility that the set of solutions is empty!