Further Mathematical Methods (Linear Algebra) 2002

Lecture 2: Other Useful Concepts

In this hand-out we are going to continue to revise some of the ideas which we encountered in MA100. This will include defining terms like linear span, linear independence, basis and dimension. We shall also see how these concepts are related by proving some theorems. It should be noted that the proofs which I have included are quite 'wordy' as I want to make the reasoning involved explicit for those of you who are unfamiliar with the methods used. The Learning Objectives associated with this hand-out are given at the end.

2.1 Linear Span

Imagine that we are given a set of m vectors, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$, which are all in the vector space $\mathbb{R}^n$. This means, as we have seen, that we can take vector sums and scalar multiples of these vectors to get other vectors in $\mathbb{R}^n$. But, instead of doing these operations individually, as we have done so far, we can also take *linear combinations* of these vectors. That is, we can perform several of the operations which are defined on the vector space at the 'same time.' For example, we could find the vector formed by

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_m\mathbf{u}_m$$

where $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{R}$ and this is an example of a linear combination. More specifically,

Definition 2.1 *A vector* $\mathbf{v}$ *is a* linear combination *of the vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ *iff there are scalars* $\alpha_1, \alpha_2, \ldots, \alpha_m$ *such that*

$$\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_m\mathbf{u}_m$$

The scalars, α_i $(1 \leq i \leq m)$ *are called* coefficients.

To illustrate this, consider a given a set of vectors, how would we find the linear combination that corresponds to a certain vector? Well, all we have to do is find the coefficients.

Example: We may be given the set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$$

and then asked whether the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

are linear combinations of them.

Solution: This problem is easily solved, as we just have to see if there is a linear combination of the two vectors which gives us the required vector. That is, are there two scalars, α and β say, such that, for $\mathbf{v}_1$,

$$\begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}?$$

So, all we have to do is look at how the components behave on both sides of this expression to get three simultaneous equations in α and β, i.e.

$$\begin{aligned} 1\alpha + 1\beta &= 4 \\ 0\alpha + 2\beta &= 2 \quad \Rightarrow \beta = 1 \\ 1\alpha + 0\beta &= 3 \quad \Rightarrow \alpha = 3 \end{aligned}$$

Now, as these values give $3+1=4$ in the top equation, this is a solution and so $\mathbf{v}_1$ is a linear combination of the two vectors, i.e.

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

Notice that the coefficients of the two vectors $[1,0,1]^t$ and $[1,2,0]^t$ are 3 and 1 respectively.

For $\mathbf{v}_2$, we do much the same, and ask whether there are two scalars, α and β say, such that

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}?$$

So, again, we just have to look at how the components behave on both sides of this expression to get three simultaneous equations in α and β, i.e.

$$\begin{aligned} 1\alpha + 1\beta &= 1 \\ 0\alpha + 2\beta &= 2 \quad \Rightarrow \beta = 1 \\ 1\alpha + 0\beta &= 1 \quad \Rightarrow \alpha = 1 \end{aligned}$$

But, as these values give $1+1=2\neq 1$ in the top equation, this is not a solution and so $\mathbf{v}_2$ is not a linear combination of the two vectors. Thus, $\mathbf{v}_2$ cannot be written in the form

$$\alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

as there is no appropriate choice for the coefficients α and β (i.e. the simultaneous equations for α and β are inconsistent). ♣

This leads us onto two other useful concepts. Firstly, if we have a vector space, we can ask: which sets of vectors *span* it? That is, which sets of vectors will give us *all* of the vectors in the vector space if all we can do is take linear combinations of the vectors in the set? Or, formally:

Definition 2.2 *A set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ spans *a vector space* V *iff every vector in* V *can be expressed as a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$.

This is, again, quite easy to do in practice.

Example: Find a set of vectors which will span the vector space $S_{2,-1,-2}$. (Recall that $S_{2,-1,-2}$ is the set consisting of vectors that lie in the plane $2x-y-2z=0$, see Section 1.4.2.)

Solution: Clearly, we have one equation $2x-y-2z=0$ in three variables, and so two of them must be free. Let us take y and z, say, to be the free variables and set them equal to the free parameters $2r$ and s respectively. This means that our equation can now be written as $x=r+s$, or equivalently, we can see that vectors of the form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r+s \\ 2r \\ s \end{bmatrix} = r \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

will satisfy this equation. Thus, every vector in the vector space $S_{2,-1,-2}$ can be written as a linear combination of the vectors in the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$$

with a suitable choice of coefficients r and s, i.e. this set spans the vector space $2x - y - 2z = 0$. ♣

Secondly, given a set of vectors, we can ask: what vector space does it span? That is, which vector space contains *all* of the vectors that we can get from linear combinations of the vectors in the set? Formally, this leads to the notion of the *linear span* of a set of vectors:

Definition 2.3 *The linear span of a set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$*, denoted by* $\mathrm{Lin}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$*, where*

$$\mathrm{Lin}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \equiv \{\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_m\mathbf{u}_m \,|\, \alpha_i \text{ is a scalar and } 1 \leq i \leq m\}$$

is the set of all linear combinations of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$.

However, before we move on, we still have to justify the assertion that the set of vectors given by $\mathrm{Lin}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is a vector space. That is, we have to establish that it is a subspace of V. To do this we claim:

Theorem 2.4 *If* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$*, then* $L = \mathrm{Lin}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ *is a subspace of* V*. Further, it is the smallest subspace containing the vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ *(in the sense that every other subspace of* V *that contains the vectors* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ *must contain* L*). Incidentally, we say that* L *is the space spanned by the set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$.

Proof: To make the proof shorter, let us use, as [an obvious] shorthand for linear combinations,

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_m\mathbf{u}_m \equiv \sum_{i=1}^{m} \alpha_i\mathbf{u}_i$$

If L is to be a subspace, then by Theorem 1.4, we require that it is closed under vector addition and scalar multiplication. So, let $\mathbf{x}$ and $\mathbf{y}$ be two [general] vectors in L, i.e.

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i\mathbf{u}_i \text{ and } \mathbf{y} = \sum_{i=1}^{m} \beta_i\mathbf{u}_i$$

where, for $1 \leq i \leq m$, α_i and β_i are scalars. Now,

$$\mathbf{x} + \mathbf{y} = \sum_{i=1}^{m} \alpha_i\mathbf{u}_i + \sum_{i=1}^{m} \beta_i\mathbf{u}_i = \sum_{i=1}^{m} (\alpha_i + \beta_i)\mathbf{u}_i = \sum_{i=1}^{m} \gamma_i\mathbf{u}_i$$

where $\gamma_i = \alpha_i + \beta_i$ is a scalar too; and for any scalar η,

$$\eta\mathbf{x} = \eta\sum_{i=1}^{m} \alpha_i\mathbf{u}_i = \sum_{i=1}^{m} \eta\alpha_i\mathbf{u}_i = \sum_{i=1}^{m} \gamma_i\mathbf{u}_i$$

where $\gamma_i = \eta\alpha_i$ is, again, a scalar. Thus, $\mathbf{x}+\mathbf{y}$ and $\eta\mathbf{x}$ are both linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$, and consequently, both are in L. Hence, L is closed under vector addition and scalar multiplication (as required).

Further, each vector $\mathbf{u}_i$ in the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ can be written as a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$, since setting $\alpha_i = 1$ and $\alpha_j = 0$ $(\forall j \neq i, 1 \leq j \leq m)$ we get

$$\mathbf{u}_i = 0\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + 1\mathbf{u}_i + \cdots + 0\mathbf{u}_m.$$

The subspace L therefore contains each of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ (as one would expect). Now, let W be *any other* subspace that contains $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$. Since W is a subspace, it is closed under vector addition and scalar multiplication (by Theorem 1.4), and so must contain [at the very least] all linear combinations, $\sum_{i=1}^{m} \alpha_i\mathbf{u}_i$ of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$. That is, W must contain every vector in L, i.e. $L \subseteq W$, and consequently, L must be the smallest subspace containing the vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ (as required). ♠

Of course, we only needed the first bit of this theorem, but the second part will be useful in a moment. Let us see what all this means by looking at an example.

Example: Find the Cartesian equation of the subspace of $\mathbb{R}^3$ that is spanned by the set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$$

(Notice that we already know the answer to this!)

Solution: Clearly, the subspace that will be spanned by these vectors is given by the set

$$\text{Lin} \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$$

and so to find the Cartesian equation, we notice that this gives us vectors of the form

$$\mathbf{x} = \alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

where $\alpha, \beta \in \mathbb{R}$ (as we want a subspace of $\mathbb{R}^3$). Now, equating the components in this expression gives us three equations relating x, y and z, i.e.

$$\begin{aligned} x &= 1\alpha + 1\beta \\ y &= 0\alpha + 2\beta \quad \Rightarrow y = 2\beta \\ z &= 1\alpha + 0\beta \quad \Rightarrow z = \alpha \end{aligned}$$

and substituting for α and β in the first equation we get $2x - y - 2z = 0$ (as expected). ♣

If you are still unsure about the difference between span and linear span, let us consider another way of looking at what it means for a set of vectors, $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$. When we say that S *spans* a vector space, we are using the word 'span' as a verb, i.e. 'spanning' a vector space is what the vectors in S *do* by means of linear combinations. However, when we say that a vector space is the *linear span* of S, we are using the word 'span' as a noun, i.e. this *is* the vector space which is 'spanned by' (or 'composed of') all of the linear combinations the vectors in S give rise to. This leads us onto another useful concept, namely:

Definition 2.5 *A set of vectors* $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ *is a spanning set of a vector space* V *iff* $V = \text{Lin}(S) = \text{Lin}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$*. Further,* S *is a smallest spanning set iff no vector in* S *is a linear combination of the others.*

Notice that a smallest spanning set is *smallest* in the sense that we need all of the vectors in the set to span the space in question. That is, if we do *not* have a smallest spanning set, we could remove any vector that is a linear combination of the others (i.e. we could replace it by its linear combination) and hence get a *smaller* set (as the vectors in this linear combination are already in the set!). Let us look at an example of this in action.

For example: As we found in the previous example, the set of vectors $\{[1, 0, 1]^t, [1, 2, 0]^t\}$ is a spanning set of the vector space represented by the plane $2x - y - 2z = 0$. Further, as the vector $[1, 0, 1]^t$ is not a linear combination of the vector $[1, 2, 0]^t$ (i.e. there is no scalar α such that $[1, 0, 1]^t = \alpha[1, 2, 0]^t$) it is also a smallest spanning set of this space.

Indeed, if we take a subset of this set of vectors, say the set $\{[1, 0, 1]^t\}$ we can see that the vector space given by $\text{Lin}\{[1, 0, 1]^t\}$ is the space which contains vectors of the form $\mathbf{x} = \alpha[1, 0, 1]^t$ where $\alpha \in \mathbb{R}$. Notice that this is the smallest subspace containing the vector $[1, 0, 1]^t$ by Theorem 2.4 and further, this is a subspace of $\text{Lin}\{[1, 0, 1]^t, [1, 2, 0]^t\}$ as all of the vectors in the former are in the latter.

We can see this because linear combinations of the vectors in the set $\{[1,0,1]^t, [1,2,0]^t\}$ have the form

$$\alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

and this gives us all of the vectors in the subspace $\text{Lin}\{[1,0,1]^t\}$ when $\beta = 0$. ♣

So, to summarise, we say that a set of vectors spans a vector space if we can use this set as a spanning set for it. Indeed, given a spanning set, we say that the space is spanned by these vectors. Further, if none of the vectors in this set are a linear combination of the others, we say that we have a smallest spanning set for that space.

Lastly, notice that all of the definitions and theorems introduced above used a general vector space (i.e. the exact nature of the vectors and scalars was left unspecified[1]). Consequently, we can use them in any vector space whatsoever, including $\mathbb{R}^n$, $\mathbb{C}^n$ and $\mathbb{F}^{\mathbb{R}}$. So, to summarise some of these ideas, and to further our understanding of the space of real functions, let us consider another example:

Example: The subspace $\mathbb{P}_n^{\mathbb{R}}$ of the vector space $\mathbb{F}^{\mathbb{R}}$ is defined to be the set of all real polynomial functions of degree at most n (where n is a non-negative integer), i.e.

$$\mathbb{P}_n^{\mathbb{R}} \equiv \{a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + \cdots + a_n \cdot \mathbf{x^n} : x \to a_0 + a_1 x + \cdots + a_n x^n \,|\, a_0, a_1, \ldots, a_n \in \mathbb{R}\},$$

for all $x \in \mathbb{R}$. Let us investigate some of its properties:

1. Show that $\mathbb{P}_n^{\mathbb{R}}$ is indeed a subspace of $\mathbb{F}^{\mathbb{R}}$.
2. Further, show that if $0 \leq m \leq n$, then $\mathbb{P}_m^{\mathbb{R}}$ is a subspace of $\mathbb{P}_n^{\mathbb{R}}$.
3. Show that $\mathbf{3} + \mathbf{x^2}$ is a linear combination of the vectors in the set $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$, but that $\mathbf{x} + \mathbf{x^3}$ isn't.
4. Find a set of vectors that will span the vector space $\mathbb{P}_2^{\mathbb{R}}$.
5. What vector space is spanned by the set of vectors $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$?

Recall,[2] that the vector $\mathbf{f} \in \mathbb{F}^{\mathbb{R}}$ represents the function $f(x)$ (i.e. $\mathbf{f} : x \to f(x)$ for all $x \in \mathbb{R}$), and so we have $\mathbf{x^i} : x \to x^i$ for all $x \in \mathbb{R}$. Indeed, as a special case of this, $\mathbf{x^0} : x \to x^0$ for all $x \in \mathbb{R}$ gives us the *unit function*, $\mathbf{1} : x \to 1$ for all $x \in \mathbb{R}$ (i.e. the function which gives us 1 for all values of the variable x).

Solution: To make the answers shorter, we use as [an, again, obvious] shorthand for a real polynomial function of degree at most n,

$$a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + \cdots + a_n \cdot \mathbf{x^n} : x \to a_0 + a_1 x + \cdots + a_n x^n \equiv \sum_{i=0}^{n} a_i \cdot \mathbf{x^i} : x \to \sum_{i=0}^{n} a_i x^i$$

for all $x \in \mathbb{R}$.

(1) To show that $\mathbb{P}_n^{\mathbb{R}}$ is a subspace of $\mathbb{F}^{\mathbb{R}}$, by Theorem 1.4, we need to show that it is closed under vector addition and scalar multiplication.[3] Now, let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two general vectors in $\mathbb{P}_n^{\mathbb{R}}$, i.e.

$$\mathbf{x}_1 = \sum_{i=0}^{n} a_i \cdot \mathbf{x^i} \text{ and } \mathbf{x}_2 = \sum_{i=0}^{n} b_i \cdot \mathbf{x^i}$$

[1] Subject, of course, to the proviso of Footnote 2 in the hand-out for Lecture 1.

[2] See Section 1.3.3 of the hand-out for Lecture 1.

[3] Since, it should be clear that $\mathbb{P}_n^{\mathbb{R}} \subseteq \mathbb{F}^{\mathbb{R}}$ as polynomials of degree at most n defined over $\mathbb{R}$ are real-valued functions.

where a_0, a_1, …, a_n, b_0, b_1, …, $b_n \in \mathbb{R}$. Now,

$$\mathbf{x_1} + \mathbf{x_2} = \sum_{i=0}^{n} a_i \cdot \mathbf{x^i} + \sum_{i=0}^{n} \cdot b_i \mathbf{x^i} = \sum_{i=0}^{n} (a_i + b_i) \cdot \mathbf{x^i} = \sum_{i=0}^{n} c_i \cdot \mathbf{x^i}$$

where $c_i = a_i + b_i$ and so clearly, c_0, c_1, …, $c_n \in \mathbb{R}$. Also, for any $\eta \in \mathbb{R}$,

$$\eta \cdot \mathbf{x_1} = \eta \sum_{i=0}^{n} a_i \cdot \mathbf{x^i} = \sum_{i=0}^{n} \eta a_i \cdot \mathbf{x^i} = \sum_{i=0}^{n} c_i \cdot \mathbf{x^i}$$

where $c_i = \eta a_i$ and so again, c_0, c_1, …, $c_n \in \mathbb{R}$. Thus, $\mathbf{x_1} + \mathbf{x_2}$ and $\eta \mathbf{x_1}$ are both polynomial functions of degree at most n, and consequently, both are in $\mathbb{P}_n^{\mathbb{R}}$. Therefore, $\mathbb{P}_n^{\mathbb{R}}$ is closed under vector addition and scalar multiplication, and so $\mathbb{P}_n^{\mathbb{R}}$ is a subspace of $\mathbb{F}^{\mathbb{R}}$ (as required).

(2) This part is fairly obvious, but I shall really 'spell' it out! By definition, we can see that for all $x \in \mathbb{R}$,

$$\mathbb{P}_m^{\mathbb{R}} = \left\{ \sum_{i=0}^{m} a_i \cdot \mathbf{x^i} : x \to \sum_{i=0}^{m} a_i x^i \;\middle|\; a_0, a_1, \ldots, a_m \in \mathbb{R} \right\}$$

and $\mathbb{P}_n^{\mathbb{R}}$ can be written in a similar way. Now, a general vector in $\mathbb{P}_n^{\mathbb{R}}$ has the form

$$\sum_{i=0}^{n} a_i \cdot \mathbf{x^i}$$

where for $0 \leq i \leq n$, $a_i \in \mathbb{R}$, and we can form subsets of $\mathbb{P}_n^{\mathbb{R}}$ by restricting the values which the a_i can take. In particular, if we set $a_i = 0$ for $m+1 \leq i \leq n$, we will [obviously] get

$$\sum_{i=0}^{n} a_i \cdot \mathbf{x^i} = \sum_{i=0}^{m} a_i \cdot \mathbf{x^i} + \sum_{i=m+1}^{n} a_i \cdot \mathbf{x^i} = \sum_{i=0}^{m} a_i \cdot \mathbf{x^i} + 0 = \sum_{i=0}^{m} a_i \cdot \mathbf{x^i}$$

which is a general vector in $\mathbb{P}_m^{\mathbb{R}}$. Thus, we have shown that $\mathbb{P}_m^{\mathbb{R}} \subseteq \mathbb{P}_n^{\mathbb{R}}$. Also, from (1), we know that all sets of the form $\mathbb{P}_r^{\mathbb{R}}$ (where r is a non-negative integer) are vector spaces. Consequently, $\mathbb{P}_m^{\mathbb{R}}$ is a subspace of $\mathbb{P}_n^{\mathbb{R}}$ (as required).

(3) To show that $\mathbf{3} + \mathbf{x^2}$ is a linear combination of the vectors in the set $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$, we need to find coefficients a_0, a_1, $a_2 \in \mathbb{R}$ such that

$$\mathbf{3} + \mathbf{x^2} = a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + a_2 \cdot \mathbf{x^2}$$

Now, we note that using the functions that these vectors represent, this is equivalent to

$$3 + x^2 = a_0 + a_1 x + a_2 x^2$$

for all $x \in \mathbb{R}$. This clearly means that $a_0 = 3$, $a_1 = 0$ and $a_2 = 1$. Thus, we can see that $\mathbf{3} + \mathbf{x^2}$ is a linear combination of the three vectors in the set, i.e.

$$\mathbf{3} + \mathbf{x^2} = 3 \cdot \mathbf{1} + 0 \cdot \mathbf{x} + 1 \cdot \mathbf{x^2}$$

where the coefficients of the vectors $\mathbf{1}$, $\mathbf{x}$ and $\mathbf{x^2}$ are 3, 0 and 1 respectively.

Then, we can attempt to do the same thing for $\mathbf{x} + \mathbf{x^3}$, that is, we ask whether there are three real numbers, a_0, a_1 and a_2, such that

$$\mathbf{x} + \mathbf{x^3} = a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + a_2 \cdot \mathbf{x^2}$$

Again, we look at the problem in terms of the functions that these vectors represent, i.e. this is equivalent to

$$x + x^3 = a_0 + a_1 x + a_2 x^2$$

for all $x \in \mathbb{R}$. Now, equating the coefficient of x^3 on both sides gives us $0 = 1$, and so [as this is inconsistent,] we cannot write $\mathbf{x} + \mathbf{x^3}$ as a linear combination of the three vectors $\mathbf{1}$, $\mathbf{x}$ and $\mathbf{x^2}$.

(4) The vector space $\mathbb{P}_2^{\mathbb{R}}$ is given by the set

$$\{a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + a_2 \cdot \mathbf{x^2} : x \to a_0 + a_1 x + a_2 x^2 \,|\, a_0, a_1, a_2 \in \mathbb{R}\},$$

for all $x \in \mathbb{R}$, and so a general vector in $\mathbb{P}_2^{\mathbb{R}}$ will have the form

$$a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + a_2 \cdot \mathbf{x^2}$$

where a_0, a_1, $a_2 \in \mathbb{R}$. Thus, every vector in the vector space $\mathbb{P}_2^{\mathbb{R}}$ can be written as a linear combination of the vectors in the set $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$ with a suitable choice of coefficients a_0, a_1 and a_2. Consequently, this set spans the vector space $\mathbb{P}_2^{\mathbb{R}}$. (Notice that $\mathbb{P}_2^{\mathbb{R}}$ is the set of all real polynomials that are at most quadratic, and further, this puts (3) into context, as obviously $\mathbf{x} + \mathbf{x^3}$ is a cubic [i.e. not at most quadratic].)

(5) Clearly, the vector space that is spanned by the set of vectors $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$ is given by $\text{Lin}\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$, or writing this out in full, we get the set

$$\{a_0 \cdot \mathbf{1} + a_1 \cdot \mathbf{x} + a_2 \cdot \mathbf{x^2} \,|\, a_0, a_1, a_2 \in \mathbb{R}\}$$

and this is clearly just $\mathbb{P}_2^{\mathbb{R}}$. So, the set of vectors $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$ spans the vector space $\mathbb{P}_2^{\mathbb{R}}$. Further, we can say that $\{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$ is a spanning set of $\mathbb{P}_2^{\mathbb{R}}$. ♣

2.2 Linear Independence

We now introduce one of the most important concepts in Linear Algebra, namely linear independence. Intuitively, we are asking whether, given a set of vectors and the operations of vector addition and scalar multiplication, can we construct any of the vectors in the set from the rest? Or, more literally, whether any of the vectors in the set are *dependent* on the others, given the two operations. If the answer is no, then the set is called *linearly independent*.[4] So, formally, one possible definition of this term is:

Definition 2.6 *A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset V$ is linearly independent if the vector equation*

$$\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_m \mathbf{u}_m = \mathbf{0}$$

has $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ as its only solution [this solution is called the trivial solution]. Otherwise, the set of vectors is linearly dependent [and the solution that we thereby obtain is called a non-trivial solution].

This is not the only way of seeing if a set of vectors is linearly independent, but we shall look at other tests later. For now, let us see what this definition entails. One immediate consequence of this definition is:

Theorem 2.7 *A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is linearly dependent iff some vector $\mathbf{u}_i$ is a linear combination of the others.*[5]

[4]The reason why this dependence, or lack of it, is called *linear* is that under our two operations we can only construct *linear* combinations of the vectors.

[5]Notice that a logically equivalent form of this theorem (i.e. the contrapositive) states that: A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is linearly independent iff none of the vectors in the set is a linear combination of the others.

Proof: As this is a proof of an 'iff' statement, we have to prove it both ways:
LTR: Assume that the set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is linearly dependent. By Definition 2.6, this means that the vector equation

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_m\mathbf{u}_m = \mathbf{0}$$

has a non-trivial solution, i.e. at least one of the vectors, $\mathbf{u}_i$ say, has a coefficient, $\alpha_i \neq 0$. Thus, we can write $\mathbf{u}_i$ as a linear combination of the other vectors, i.e.

$$\mathbf{u}_i = -\frac{1}{\alpha_i}\sum_{j=1}^{m}\alpha_j\mathbf{u}_j$$

where in the summation, $j \neq i$ (as required).
RTL: Assume that there is a vector $\mathbf{u}_i$ which is a linear combination of the other vectors in the set, i.e.

$$\mathbf{u}_i = \sum_{j=1}^{m}\alpha_j\mathbf{u}_j$$

where, again, $j \neq i$ in the summation. We can re-arrange this to give

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots - \mathbf{u}_i + \cdots + \alpha_m\mathbf{u}_m = \mathbf{0}$$

which is the vector equation from Definition 2.6 with a non-trivial solution (as $\alpha_i = -1 \neq 0$) and so the set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ is linearly dependent (as required). ♠

Let us look at this in action.

Example: Are either of the sets of vectors

$$\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}\right\} \text{ and } \left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\0\\2\end{bmatrix}\right\}$$

linearly independent?

Solution: In the first case, the set of vectors gives us the vector equation

$$\alpha\begin{bmatrix}1\\0\\1\end{bmatrix} + \beta\begin{bmatrix}1\\2\\0\end{bmatrix} = \mathbf{0}$$

and looking at how the components behave on both sides of the equation, we get three simultaneous equations in α and β, i.e.

$$\begin{aligned} 1\alpha + 1\beta &= 0 \\ 0\alpha + 2\beta &= 0 \quad \Rightarrow \beta = 0 \\ 1\alpha + 0\beta &= 0 \quad \Rightarrow \alpha = 0 \end{aligned}$$

Now, as these values give $0 + 0 = 0$ in the top equation, this is a solution, and clearly, it is the only solution. Thus, the two vectors $[1, 0, 1]^t$ and $[1, 2, 0]^t$ in the set are linearly independent. Further, as we saw in an earlier example, neither of these vectors is a linear combination of the other, as expected by [the contrapositive of] Theorem 2.7.

In the second case, the set of vectors gives us the vector equation

$$\alpha\begin{bmatrix}1\\0\\1\end{bmatrix} + \beta\begin{bmatrix}2\\0\\2\end{bmatrix} = \mathbf{0}$$

and looking at how the components behave on both sides of the equation, we also get three simultaneous equations in α and β, i.e.

$$\begin{aligned} 1\alpha + 2\beta &= 0 \\ 0\alpha + 0\beta &= 0 \\ 1\alpha + 2\beta &= 0 \end{aligned}$$

However, the middle equation just gives $0 = 0$, and this is always satisfied; whereas the top and bottom equations are identical. Thus, in effect, we have just one equation, i.e. $\alpha + 2\beta = 0$, now if we let β, say, equal the free parameter $-r$, then clearly, $\alpha = 2r$ and further, all of the $[\alpha, \beta]^t$ pairs given by

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 2r \\ -r \end{bmatrix} = r \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

will be solutions of this equation. Thus, there are an infinite number of solutions to the vector equation (as $r \in \mathbb{R}$) and so the two vectors $[1, 0, 1]^t$ and $[2, 0, 2]^t$ in the set are not linearly independent, but linearly dependent.

Further, it should be clear that this set of vectors is linearly dependent because the vector $[2, 0, 2]^t$ clearly *depends* on $[1, 0, 1]^t$ as it is just the scalar multiple of $[1, 0, 1]^t$ and the scalar 2. That is, in terms of Theorem 2.7, the vector $[2, 0, 2]$ can be written as a linear combination of $[1, 0, 1]^t$ because it is just this vector with a coefficient of 2. ♣

Using Theorem 2.7, we can now see that a smallest spanning set is the *smallest* such set due to the fact that the vectors it contains are linearly independent, i.e.

Theorem 2.8 *A set of vectors is a smallest spanning set iff it is a linearly independent set.*

Proof:[6] By Definition 2.5, a set of vectors, S is a smallest spanning set iff no vector in S is a linear combination of the others. But, by [the contrapositive of] Theorem 2.7, none of the vectors in S is a linear combination of the others iff S is a linearly independent set of vectors. So, we can see that S is a smallest spanning set iff it is a linearly independent set. ♠

For example: We have seen in previous examples that the set $\{[1, 0, 1]^t, [1, 2, 0]^t\}$ is both linearly independent and a smallest spanning set, as expected by Theorem 2.8. ♣

Indeed, we can also prove another useful result which relates a linearly independent set of vectors to the vector space given by the linear span of this set, namely:

Theorem 2.9 *Let V be a vector space and let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ be a linearly independent set of vectors. If $\mathbf{v} \in V$ and $\mathbf{v} \notin \text{Lin}(S) \subseteq V$, then the set $S \cup \{\mathbf{v}\}$ (i.e. the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m, \mathbf{v}\}$) is also linearly independent.*

Proof: Assume that V is a vector space and that $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ is a linearly independent set of vectors. Further, assume that the vector $\mathbf{v}$ is in V, but not in $\text{Lin}(S) \subseteq V$. Now, to show that the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m, \mathbf{v}\}$ is linearly independent, by Definition 2.6, we must show that the only scalars which satisfy the vector equation

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_m\mathbf{u}_m + \alpha_{m+1}\mathbf{v} = \mathbf{0} \tag{2.6}$$

are $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \alpha_{m+1} = 0$.

To do this, we note that if α_{m+1} was non-zero, then we could re-arrange Equation 2.6 and get $\mathbf{v}$ as a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$. But, this contradicts our assumption that $\mathbf{v}$ is not in $\text{Lin}(S)$. Thus, we conclude that $\alpha_{m+1} = 0$. Further, this means that Equation 2.6 reduces to

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_m\mathbf{u}_m = \mathbf{0}$$

Now, as we know that the vectors in the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ are linearly independent, by Definition 2.6, it must be the case that $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ too. Thus, by Definition 2.6, the set $S \cup \{\mathbf{v}\}$ is linearly independent. ♠

We now introduce another important concept which links linear independence and the vector space which a set of vectors spans, namely the notion of a *basis* of a vector space.

[6]This proof is pretty simple, and so we shall prove both directions at once by using a 'chain' of 'iff' statements.

2.3 Bases

Intuitively, a *basis* is the *smallest* set of vectors that can span a vector space. This means that every vector in the vector space can be written as a linear combination of the vectors in the basis, and further, this linear combination will be *unique*. Or, more formally, we have:

Definition 2.10 *A set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ *is a basis of a vector space* V *iff it spans* V *and it is linearly independent.*

With this, we can justify the other two assertions about bases by using the following two theorems. Firstly, we prove the claim that a basis of a vector space is the smallest set of vectors that can span it using:

Theorem 2.11 *A set of vectors* $B \subseteq V$ *is a basis of the vector space* V *iff it is a smallest spanning set of the vector space* V.

Proof: The proof of this theorem gets quite involved if we split it up into the LTR and RTL cases and so we adopt the method that we used to prove Theorem 2.8 to make it simpler.

Assume that the set of vectors $B \subseteq V$ is a smallest spanning set of the vector space V. By Definition 2.5, this is the case iff $V = \text{Lin}(B)$ and no vector in B is a linear combination of the others. But, by Definition 2.3, this is the case iff V is the set of all linear combinations of the vectors in B (i.e. every vector in V is a linear combination of the vectors in B) and no vector in B is a linear combination of the others. So, by Definition 2.2, this is the case iff B spans V and no vector in B is a linear combination of the others. Now, by [the contrapositive of] Theorem 2.7, this is the case iff B spans V and the vectors in B are linearly independent. Consequently, by Definition 2.10, this is the case iff B is a basis of V. ♠

Secondly, we show that every vector in a vector space can be written as a unique linear combination of the vectors in the basis using:

Theorem 2.12 *A set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ *is a basis of a vector space* V *iff every* $\mathbf{v} \in V$ *is a unique linear combination of the vectors* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$.

Proof: As this is a proof of an 'iff' statement, we have to prove it both ways:
LTR: Let us assume that the set of vectors $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ is a basis of the vector space V. That is, by Definition 2.10, B spans V and the vectors in V are linearly independent. Now, as B spans V, by Definition 2.2, every vector in V can be written as a linear combination of the vectors in B. That is, by Definition 2.1, every vector $\mathbf{v} \in V$ can be written in the form

$$\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_m\mathbf{u}_m. \tag{2.7}$$

To prove uniqueness, we now assume that there is another way of writing the vector $\mathbf{v}$ as a linear combination, say,

$$\mathbf{v} = \alpha'_1\mathbf{u}_1 + \alpha'_2\mathbf{u}_2 + \cdots + \alpha'_m\mathbf{u}_m.$$

But, by equating these two expressions for $\mathbf{v}$ and rearranging, we get

$$(\alpha_1 - \alpha'_1)\mathbf{u}_1 + (\alpha_2 - \alpha'_2)\mathbf{u}_2 + \cdots + (\alpha_m - \alpha'_m)\mathbf{u}_m = 0.$$

Now, as the vectors in B are linearly independent, by Definition 2.6, this equation can only have one solution, namely $\alpha_1 - \alpha'_1 = \alpha_2 - \alpha'_2 = \cdots = \alpha_m - \alpha'_m = 0$. Thus, we find that $\alpha_1 = \alpha'_1$, $\alpha_2 = \alpha'_2$, $\ldots$, $\alpha_m = \alpha'_m$; i.e. the linear combination is unique (as required).
RTL: Let us assume that every vector in V can be written as a unique linear combination of the vectors in the set $C = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$. By Definition 2.2, this means that the set C spans V. Now, by Definition 2.10, to establish that B is a basis of V, we just have to prove that C is linearly independent. We shall do this using a proof by contradiction.[7]

[7] This is a very useful method of proof, commonly referred to as *reductio ad absurdum*. To use it, assume the opposite (strictly, the negation) of what you want to prove and show that this contradicts your *other* assumptions. This, in turn, entails that what you wanted to prove must be true. If you can't see why this works, just use it anyway — it would take too long to explain. (Also notice that this method was used in the proof of Theorem 2.9 to establish that α_{m+1} must be zero.)

Assume, for contradiction, that C is a linearly dependent set of vectors and consider the vector $\mathbf{v} \in V$, which can be written as a linear combination of the vectors in C as in Equation 2.7. Now, as the vectors in C are linearly dependent, by Theorem 2.7, this means that at least one of the vectors in C is a linear combination of the others. Let us take one of these vectors to be $\mathbf{u}_i$, and so, we can write

$$\mathbf{u}_i = \sum_{j=1}^{m} \beta_j \mathbf{u}_j$$

where in this summation, $i \neq j$. Now, on substitution into Equation 2.7, this yields

$$\mathbf{v} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_{i-1} \mathbf{u}_{i-1} + \alpha_i \left(\sum_{j=1}^{m} \beta_j \mathbf{u}_j \right) + \alpha_{i+1} \mathbf{u}_{i+1} + \cdots + \alpha_m \mathbf{u}_m.$$

where, once again, the summation is such that $i \neq j$. But, this new expression for the vector $\mathbf{v}$ is different from the one in Equation 2.7, as the latter contains the vector $\mathbf{u}_i$ but the former doesn't. Consequently, we have two distinct linear combinations that represent the vector $\mathbf{v}$, *contrary* to the assumption that such a linear combination must be unique. Thus, by contradiction, the set C must be linearly independent (as required). ♠

Let us see all of this in action.

For example: We have seen, in previous examples, that the set of vectors $\{[1,0,1]^t, [1,2,0]^t\}$ both spans the vector space $S_{2,-1,-2}$, and is linearly independent. Thus, by Definition 2.10, it is a basis of this vector space. We have also seen that it is a smallest spanning set of this vector space, as expected by Theorem 2.11.

Further, in the very first example, we considered the vector $[4,2,3]^t$ (which clearly lies in the vector space $S_{2,-1,-2}$ as its components satisfy the equation $2x - y - 2z = 0$, i.e. $8 - 2 - 6 = 0$) and we saw that it could be expressed as a linear combination of the vectors in this set, i.e.

$$\begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Glancing back at the calculation that gave us this, it is clear that this is the only linear combination that we are going to find. Thus, this vector can be expressed as a unique linear combination of the vectors in this set, as expected by Theorem 2.12. ♣

So, to summarise these two results, we could say that the fact that a basis spans a space implies that every vector in the space can be written as a linear combination of the vectors in the basis; and the fact that the vectors in the basis are linearly independent means that this linear combination will be unique. Further, taking these two properties together gives us the fact that a basis represents the smallest set of vectors that can be used to span a certain vector space. This last point leads us on to another useful concept in the study of vector spaces, namely the *dimension* of a vector space.

2.4 Dimension

Intuitively, this is the number of vectors that a set must contain in order to be a basis of V. This may not sound like a very interesting concept but, among other things, it enables us to see straight away that some sets of vectors cannot be a basis for a given vector space. Formally, we say that

Definition 2.13 *The dimension of a vector space V, denoted by $\dim(V)$, is equal to the number of vectors in a basis of V. Further, if a vector space V has $\dim(V) = m$ where $m \in \mathbb{N}$, then V is called finite dimensional; otherwise, V is called infinite dimensional.*

To justify this definition, we should convince ourselves that the dimension of a finite dimensional vector space is unique. That is, we should show that all possible bases for a given vector space contain the same number of vectors. But, before we do this, we have to prove another useful theorem about bases:

Theorem 2.14 *Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ be a basis of the finite dimensional vector space V[8] and let $S' \subseteq V$ be any set containing n vectors. If $n > m$, then the vectors in S' are linearly dependent.*[9]

Proof: Let $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any set of vectors in V where $n > m$. To prove the theorem we need to show that S' is linearly dependent. Now, since S is a basis of V, each vector $\mathbf{v}_i \in S'$ can be expressed as a linear combination of the vectors in S, i.e.

$$\begin{array}{ccccccccc}
\mathbf{v}_1 & = & a_{11}\mathbf{u}_1 & + & a_{12}\mathbf{u}_2 & + & \cdots & + & a_{1m}\mathbf{u}_m \\
\mathbf{v}_2 & = & a_{21}\mathbf{u}_1 & + & a_{22}\mathbf{u}_2 & + & \cdots & + & a_{2m}\mathbf{u}_m \\
\vdots & & \vdots & & \vdots & & & & \vdots \\
\mathbf{v}_n & = & a_{n1}\mathbf{u}_1 & + & a_{n2}\mathbf{u}_2 & + & \cdots & + & a_{nm}\mathbf{u}_m
\end{array}$$

and to show that the vectors in the set S' are linearly dependent, by Definition 2.6, we must find scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ which are not all zero, such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0} \tag{2.8}$$

So, substituting the expressions for $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ above into this vector equation and re-arranging gives

$$(\alpha_1 a_{11} + \alpha_2 a_{21} + \cdots + \alpha_n a_{n1})\mathbf{u}_1 + \cdots + (\alpha_1 a_{1m} + \alpha_2 \alpha_{2m} + \cdots + \alpha_n a_{nm})\mathbf{u}_m = \mathbf{0}$$

and as the vectors in S are a basis of the vector space V, they must be linearly independent. Thus, each of the coefficients in this new vector equation must be zero, i.e.

$$\begin{array}{ccccccccc}
a_{11}\alpha_1 & + & a_{21}\alpha_2 & + & \cdots & + & a_{n1}\alpha_n & = & 0 \\
a_{12}\alpha_1 & + & a_{22}\alpha_2 & + & \cdots & + & a_{n2}\alpha_n & = & 0 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{1m}\alpha_1 & + & a_{2m}\alpha_2 & + & \cdots & + & a_{nm}\alpha_n & = & 0
\end{array}$$

Now, to show that the vectors in S' are linearly dependent, all we have to do is show that there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ which are not all zero and satisfy these simultaneous equations. But, since $n > m$, we have more unknowns than equations[10] and so there will be $n - m \neq 0$ free variables which, in general, will be non-zero. Consequently, there will be non-trivial solutions to Equation 2.8 and hence the vectors in the set S' are linearly dependent. ♠

Now, it is easy to show that the dimension of a vector space is unique, i.e.

Theorem 2.15 *If the set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ is a basis of the finite dimensional vector space V, then every basis of V contains exactly m vectors.*

Proof: Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ be a basis for the vector space V which [clearly] contains m vectors and let us assume that $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$ is another basis for the vector space V which contains n vectors, where n may be different from m. To prove the theorem, we need to show that $m = n$.

To do this, we note that since S is a basis and S' is a linearly independent set of vectors (as it is a basis too), by [the contrapositive of] Theorem 2.14, we must have $n \leq m$. Further, since S' is a basis and S is a linearly independent set of vectors (as it is a basis too), by [the contrapositive of] Theorem 2.14 [again], we must have $m \leq n$. But, in order for these two inequalities to hold, it must be the case that $m = n$. ♠

[8]That is, V is an m-dimensional vector space.

[9]Notice that the contrapositive of this theorem states: Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subseteq V$ be a basis of the vector space V and let $S' \subseteq V$ be a set containing n vectors. If the vectors in S' are linearly independent, then $n \leq m$.

[10]Notice that we have m simultaneous equations relating the n variables $\alpha_1, \alpha_2, \ldots, \alpha_n$.

As always, we now point out [the obvious] connections between these new concepts.

For example: We have seen that the set $\{[1,0,1]^t, [1,2,0]^t\}$ is a basis of the vector space $S_{2,-1,-2}$, and so, as this set contains two vectors, this vector space must have a dimension of two. Further, this means that this vector space is finite dimensional.

Further, if we take a larger set of vectors that is still contained within this vector space, say the set $\{[1,0,1]^t, [1,2,0]^t, [0,2,1]^t\}$, we can test them for linear independence by using the vector equation from Definition 2.6, i.e.

$$\alpha \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

As before, we look at how the components behave on both sides to get three simultaneous equations in α, β and γ, that is:

$$\begin{aligned} 1\alpha + 1\beta + 2\gamma &= 0 \\ 0\alpha + 2\beta + 2\gamma &= 0 \quad \Rightarrow \beta = -\gamma \\ 1\alpha + 0\beta + 1\gamma &= 0 \quad \Rightarrow \alpha = -\gamma \end{aligned}$$

Now, as these values give $-\gamma - \gamma + 2\gamma = 0$ in the top equation, this set of simultaneous equations is consistent. Further, setting γ equal to a free parameter, say $-r$ (in $\mathbb{R}$), we can see that these equations have infinitely many solutions, as given by the $[\alpha, \beta, \gamma]^t$ triples

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} r \\ r \\ -r \end{bmatrix} = r \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

That is, these vectors are linearly dependent as expected by Theorem 2.14. ♣

It is interesting to note that we should supplement Definition 2.13 by saying that:

Definition 2.16 *The dimension of the vector space* $\{\mathbf{0}\}$ *is zero.*

This is clearly a special case because the vector space $\{\mathbf{0}\}$ does not contain any linearly independent sets[11] and so it can't have any bases either. But, it is useful to extend our definition of dimension in this way so that it corresponds to our geometrical understanding of the term (i.e. a point [represented by a vector] is defined as a geometrical entity with no dimension). Also note that $\{\mathbf{0}\}$ is the only vector space that contains just one vector.

Finally, to show that the notion of dimension is useful in our study of vector spaces, we prove a theorem which sometimes allows us to immediately assess whether certain sets of vectors are bases of a vector space, i.e.

Theorem 2.17 *If S is a set of n linearly independent vectors in an n-dimensional vector space V, then S is a basis for V.*

Proof: Let S be a set of n linearly independent vectors in an n-dimensional vector space V. Now, by Definition 2.10, to show that S is a basis for V we need to establish that S spans V (note that we already know that S is linearly independent!). We shall do this using proof by contradiction.

Assume, for contradiction, that $S \subseteq V$ does *not* span V. That is, by Definition 2.2, there is some vector $\mathbf{v} \in V$ which is *not* a linear combination of the other vectors in S. Now, by Definition 2.3, this means that $\mathbf{v} \notin \text{Lin}(S) \subseteq V$. Consequently, using Theorem 2.9, we can form the set $S \cup \{\mathbf{v}\}$ (by adding the vector $\mathbf{v}$ to S) and this set of $n+1$ vectors will be linearly independent too. But, this is contrary to Theorem 2.14 which implies that a set of $n+1$ vectors in an n-dimensional space is linearly dependent. Thus, by contradiction, S must span V. ♠

You will be asked to prove a related result in Problem Sheet 1.

[11] That is, the vector equation $\alpha\mathbf{0} = \mathbf{0}$ is satisfied by all $\alpha \in \mathbb{R}$ and consequently, we do not get a unique trivial solution as required by Definition 2.6.

2.5 Learning Objectives

At the end of this hand-out you should:

- Understand the meaning of the terms given in the Definitions and how they are related *via* the Theorems.
- Begin to understand the methods by which these Theorems are proved. (Although, detailed knowledge of these proofs is not required for this course.)
- Be able to apply the Definitions and Theorems when dealing with specific vector spaces and sets of vectors as illustrated in the Examples.

This material will be developed in Problem Sheet 1.