SUPPLEMENT 1: LINEAR INDEPENDENCE AND BASES

1.1 Linear combinations and spanning sets

Consider the vector space $\mathbb{R}^3$ with the unit vectors $\mathbf{e}_1 = (1,0,0), \mathbf{e}_2 = (0,1,0), \mathbf{e}_3 = (0,0,1)$. Every vector $\mathbf{v} = (a,b,c) \in \mathbb{R}^3$ can be expressed in terms of vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, namely:

$$\mathbf{v} = (a,b,c) = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 = a(1,0,0) + b(0,1,0) + c(0,0,1),$$

this means that every vector of $\mathbb{R}^3$ is a sum of (scalar) multiples of $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$. In the following we want to study this property more closely. To do so we first need a definition:

Definition. Let V be a vector space and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ a finite set of vectors in V. A vector $\mathbf{w} \in V$ is called *a linear combination of vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ if there are scalars $r_1, \ldots, r_n \in \mathbb{R}$ so that

$$\mathbf{w} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n.$$

The scalars $r_1, \ldots, r_n$ are called *the coefficients* of this linear combination.

We have just seen that every vector $\mathbf{v} \in \mathbb{R}^3$ is a linear combination of the 3 vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

Given any set of n vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in a vector space V we want to investigate the set of all linear combinations of these n vectors. Thus we make the definition:

Definition. Let V be a vector space and $C = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ a finite set of vectors in V. The *span of the set of vectors* $C = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is the set of all linear combinations of these n vectors:

$$S = \mathrm{span}(C) = \mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \{r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n \,|\, r_1, \ldots, r_n \in \mathbb{R}\}.$$

We also say that the set S *is spanned by the set of the vectors* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

Whenever we define certain subsets of a vector space the first question which arises is if these subsets are 'interesting', that is, do they relate in some ways to the vector space structure on the whole vector space? In case of spanning sets, this means that to ask if these subsets are subspaces. Here is the answer:

Theorem S1.1. *Let V be a vector space and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ a finite set of vectors in V. The spanning set of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$*

$$S = span\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

is a subspace of V.

Proof. In order to prove this theorem remember that

$$S = \mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \{r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n \,|\, r_1, \ldots, r_n \in \mathbb{R}\}.$$

Obviously,

$$\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_n \in S,$$

and the set S is nonempty. Next we want to show that S is closed under vector addition. So, let $\mathbf{v}, \mathbf{w} \in S$ be vectors in S. Since S is the set of all linear combinations of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, each $\mathbf{v}$ and $\mathbf{w}$ is a linear combination of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, that is, there are scalars $r_1, \ldots, r_n \in \mathbb{R}$ and $s_1, \ldots, s_n \in \mathbb{R}$ with

$$\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n \text{ and } \mathbf{w} = s_1\mathbf{v}_1 + \ldots + s_n\mathbf{v}_n.$$

Then

$$\begin{aligned}\mathbf{v} + \mathbf{w} &= (r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n) + (s_1\mathbf{v}_1 + \ldots + s_n\mathbf{v}_n)\\ &= (r_1 + s_1)\mathbf{v}_1 + \ldots + (r_n + s_n)\mathbf{v}_n\end{aligned}$$

which is another linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and therefore $\mathbf{v}+\mathbf{w} \in S$.

In order to show that S is closed under scalar multiplication, let $\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n \in S$ where $r_1, \ldots, r_n \in \mathbb{R}$ and $c \in \mathbb{R}$. Then

$$\begin{aligned}c\mathbf{v} &= c(r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n)\\ &= (cr_1)\mathbf{v}_1 + \ldots + (cr_n)\mathbf{v}_n\end{aligned}$$

which is again a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $c\mathbf{v} \in S$. This shows that S is a subspace of V.

Our example $\mathbb{R}^3$ is the spanning set of the set of the unit vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, that is,

$$\mathbb{R}^3 = \mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}.$$

You can easily convince yourself that no fewer vectors than the 3 vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ span $\mathbb{R}^3$, for example,

$$\mathbf{e}_3 = (0, 0, 1) \notin \mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\}.$$

Another obvious fact is that whenever any vector $\mathbf{v} = (a, b, c)$ is written as a linear combination of the unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

$$\mathbf{v} = (a, b, c) = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3,$$

then the scalars a, b, c are unique, which means that if

$$\mathbf{w} = (a', b', c') = a'\mathbf{e}_1 + b'\mathbf{e}_2 + c'\mathbf{e}_3$$

with $a \neq a'$ or $b \neq b'$ or $c \neq c'$ then $\mathbf{v} \neq \mathbf{w}$. On the other hand, the set $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{u} = (2, 0, 1)\}$ is another spanning set of $\mathbb{R}^3$ (why?). Here the vector

$\mathbf{v} = (1,0,1)$ can be written as a linear combination of these 4 vectors in at least two different ways:

$$\begin{aligned}\mathbf{v} &= 1\mathbf{e}_1 + 0\mathbf{e}_2 + 1\mathbf{e}_3 + 0\mathbf{u}\\ &= (-1)\mathbf{e}_1 + 0\mathbf{e}_2 + 0\mathbf{e}_3 + 1\mathbf{u}.\end{aligned}$$

More generally the following questions arise:

(1) Are there shortest spanning sets and when is a spanning set shortest?

(2) Suppose we can write a vector $\mathbf{v}$ as a linear combination of the set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, say:

$$\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$$

where $r_i \in \mathbb{R}$. When are the coefficients $r_1, \ldots, r_n$ unique?

In order to study these questions more closely we need the notion of linear independence which will be discussed in the next section.

Linear independence

Definition. In a vector space V a finite set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ is called *linearly independent* if and only if the equation

$$r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n = \mathbf{0}$$

implies that $r_1 = r_2 = \ldots = r_n = 0$. If it is possible for the equation to hold when one or more of the coefficients are nonzero, the set is *linearly dependent.*

Remark. For any set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ the zero vector can *always* be written as

$$(*) \quad \mathbf{0} = 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_n.$$

The definition states that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent if $(*)$ is *the one and only* way the zero vector can be written as a linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. So, the set of vectors is linearly independent if and only of the zero vector can be written *in a unique way* (namely $(*)$) as a linear combination of the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. A natural question to ask here is if the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent implying that the zero vector can be written in a unique way as a linear combination of those vectors, then what about the other vectors in $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, can they also be written as a linear combination with unique coefficients? The answer is yes and even more can be shown:

Theorem S1.2. *Let $C = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a finite set of vectors in V. The following are equivalent:*

(1) *C is linearly independent.*
(2) *Every vector in span(C) has a unique expression as a linear combination of vectors in C.*
(3) *No vector in C is a linear combination of the other vectors in C.*

Proof. (1) $\Rightarrow$ (2) : Let $\mathbf{v} \in \text{span}(C)$ and assume that

$$\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n = s_1\mathbf{v}_1 + \ldots + s_n\mathbf{v}_n$$

where $r_i, s_i \in \mathbb{R}$. Then

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = (r_1 - s_1)\mathbf{v}_1 + \ldots + (r_n - s_n)\mathbf{v}_n.$$

Since the set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent:

$$r_1 - s_1 = r_2 - s_2 = \ldots r_n - s_n = 0 \quad \Rightarrow \quad r_1 = s_1, r_2 = s_2, \ldots, r_n = s_n.$$

(2) $\Rightarrow$ (3) : Suppose that $\mathbf{v}_i \in \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$. Then we can write $\mathbf{v}_i$ in two different ways as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, namely:

$$\begin{aligned}\mathbf{v}_i &= r_1\mathbf{v}_1 + \ldots + r_{i-1}\mathbf{v}_{i-1} + r_{i+1}\mathbf{v}_{i+1} + \ldots + r_n\mathbf{v}_n \\ &= r_1\mathbf{v}_1 + \ldots + r_{i-1}\mathbf{v}_{i-1} + 0\mathbf{v}_i + r_{i+1}\mathbf{v}_{i+1} + \ldots + r_n\mathbf{v}_n\end{aligned}$$

and

$$\mathbf{v}_i = 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \ldots + 0\mathbf{v}_n,$$

a contradiction

(3) $\Rightarrow$ (1) : Suppose that $\mathbf{0} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$. If $r_i \neq 0$ then

$$r_i\mathbf{v}_i = (-r_1)\mathbf{v}_1 + \ldots + (-r_{i-1})\mathbf{v}_{i-1} + (-r_{i+1})\mathbf{v}_{i+1} + \ldots + (-r_n)\mathbf{v}_n.$$

Thus

$$\mathbf{v}_i = (-r_1/r_i)\mathbf{v}_1 + \ldots + (-r_{i-1}/r_i)\mathbf{v}_{i-1} + (-r_{i+1}/r_i)\mathbf{v}_{i+1} + \ldots + (-r_n/r_i)\mathbf{v}_n$$

and $\mathbf{v}_i \in \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$, a contradiction

Example. The set of unit vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ in $\mathbb{R}^3$ is linearly independent, since

$$\mathbf{0} = (0,0,0) = a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 = (a,b,c)$$

implies that $a = b = c = 0$. On the other hand, the set

$$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{u} = (2,0,1)\}$$

is linearly dependent since

$$\mathbf{0} = (0,0,0) = 2\mathbf{e}_1 + 0\mathbf{e}_2 + 1\mathbf{e}_3 + (-1)\mathbf{u}.$$

The following Lemma is very useful in the next section when we discuss bases of vector spaces.

Linear Dependence Lemma. *If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ is a set of linearly dependent vectors in V with $\mathbf{v}_1 \neq \mathbf{0}$, then there is a $j \in \{2, \ldots, n\}$ so that*

(a) $\mathbf{v}_j \in \mathit{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\}$

(b) $\mathit{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \mathit{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n\}$.

Proof. (a) By assumption the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. Thus there are $r_1, \ldots, r_n \in \mathbb{R}$, not all 0, so that:

$$\mathbf{0} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n.$$

Since $\mathbf{v}_1 \neq 0$, not all $r_2, \ldots, r_n$ are 0. Let $j \geq 2$ be the largest integer with $r_j \neq 0$. Then

$$\mathbf{v}_j = (-r_1/r_j)\mathbf{v}_1 + \ldots + (-r_{j-1}/r_j)\mathbf{v}_{j-1}.$$

This shows that $\mathbf{v}_j \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\}$.

(b) As has been shown in (a) there is an integer $j \in \{2, \ldots, n\}$ so that $\mathbf{v}_j \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\}$. Then, obviously, $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n\} \subseteq \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. In order to show the other inclusion we use the fact that $\mathbf{v}_j \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n\}$ and write

$$\mathbf{v}_j = r_1\mathbf{v}_1 + \ldots r_{j-1}\mathbf{v}_{j-1} + r_{j+1}\mathbf{v}_{j+1} + \ldots + r_n\mathbf{v}_n$$

for some $r_i \in \mathbb{R}$. Let $\mathbf{v} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ then

$$\mathbf{v} = t_1\mathbf{v}_1 + \ldots + t_n\mathbf{v}_n$$

for some $t_i \in \mathbb{R}$. Substituting the first equation for $\mathbf{v}_j$ into the second equation yields:

$$\mathbf{v} = (t_1 + t_j r_1)\mathbf{v}_1 + \ldots + (t_{j-1} + t_j r_{j-1})\mathbf{v}_{j-1} + (t_{j+1} + t_j r_{j+1})\mathbf{v}_{j+1} + \ldots + (t_n + t_j r_n)\mathbf{v}_n$$

and $\mathbf{v} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n\}$.

Corollary S1.3. *If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of vectors in V with $\mathbf{v}_1 \neq \mathbf{0}$ and $\mathbf{v}_j \notin$ span$\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\}$ for all $2 \leq j \leq n$ then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of linearly independent vectors in V.*

Bases

In the following we call a subset A a *proper* subset of a set B if A is a subset of B (i.e. $A \subseteq B$) and $A \neq B$.

Definition. Let V be a vector space and $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ a finite set of vectors in V. B is called a *basis* of V if B is linearly independent and spans V, i.e., B is linearly independent and $V = \operatorname{span}(B) = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

With this definition we see that the set of unit vectors $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a basis of $\mathbb{R}^3$ while the sets $C = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{u} = (2, 0, 1)\}$ and $D = \{\mathbf{e}_1, \mathbf{e}_2\}$ are not bases of $\mathbb{R}^3$.

Theorem S1.4. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in V. The following are equivalent:*

1. *S is linearly independent and spans V.*
2. *For every $\mathbf{v} \in V$ there are unique scalars $r_1, \ldots, r_n \in \mathbb{R}$ so that $\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$.*
3. *S is a minimal spanning set, that is, S spans V and no proper subset of S spans V.*
4. *S is a maximal linear independent set, that is, S is linearly independent and any subset T of V that properly contains S is linearly dependent.*

Proof. (1) $\Rightarrow$ (2) : Let $\mathbf{v} \in V$. Since S spans V, there are scalars $r_1, \ldots, r_n \in \mathbb{R}$ so that

$$\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n.$$

If there is another list of scalars $s_1, \ldots, s_n \in \mathbb{R}$ with

$$\mathbf{v} = s_1\mathbf{v}_1 + \ldots + s_n\mathbf{v}_n$$

then

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = (r_1 - s_1)\mathbf{v}_1 + \ldots + (r_n - s_n)\mathbf{v}_n.$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, $r_1 = s_1, r_2 = s_2, \ldots, r_n = s_n$.

(2) $\Rightarrow$ (3) : By contradiction: Suppose that $A \subseteq S$ is a subset with $A \neq S$ and suppose that $V = \text{span}(A)$. Since A is properly contained in S there is at least one $\mathbf{v}_i \in S$ with $\mathbf{v}_i \notin A$. After renumbering the $\mathbf{v}'s$ - if necessary - we may assume that $\mathbf{v}_1 \notin A$. Since A spans V, any set containing A spans V. Thus we may assume that $A = \{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$. Then $V = \text{span}(A)$, $\mathbf{v}_1 \in \text{Span}(A)$, and there are scalars $r_2, \ldots, r_n \in \mathbb{R}$ so that

$$\mathbf{v}_1 = r_2\mathbf{v}_2 + \ldots + r_n\mathbf{v}_n = 0\mathbf{v}_1 + r_2\mathbf{v}_2 + \ldots + r_n\mathbf{v}_n.$$

This is one way to write $\mathbf{v}_1$ as a linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Another is $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_n$, a contradiction to assumption (2).

(3) $\Rightarrow$ (4) : Let $T \subseteq V$ be a subset with $S \subseteq T$ and $S \neq T$ and let $\mathbf{v} \in T - S$. We know by (3) that $V = \text{span}(S)$. Thus there are scalars $r_1, \ldots, r_n \in \mathbb{R}$ so that

$$\mathbf{v} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n.$$

This gives a nontrivial linear combination of $\mathbf{0}$:

$$\mathbf{0} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n + (-1)\mathbf{v}$$

and the set of vectors $\{\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. Thus T is linearly dependent.

(4) $\Rightarrow$ (1) : By assumption (4) the set S is linearly independent. We have to show that S spans V. Let $\mathbf{v} \in V$. If $\mathbf{v} \in S$ then $\mathbf{v} = \mathbf{v}_i$ for some $i = 1, 2, \ldots, n$ and, in particular, $\mathbf{v} = \mathbf{v}_i \in \text{span}(S)$. Let $\mathbf{v} \notin S$. By assumption the set $S \cup \{\mathbf{v}\}$ is linearly dependent and there are scalars $t, r_1, \ldots, r_n \in \mathbb{R}$, not all 0, so that

$$\mathbf{0} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n + t\mathbf{v}.$$

If $t = 0$ then not all of the r_i are 0 and $\mathbf{0} = r_1\mathbf{v}_1 + \ldots + r_n\mathbf{v}_n$, a contradiction to S a linearly independent set. Thus $t \neq 0$ and

$$\mathbf{v} = (-r_1/t)\mathbf{v}_1 + \ldots + (-r_n/t)\mathbf{v}_n \in \text{span}(S).$$

In the following we call the set of unit vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ the standard basis of $\mathbb{R}^3$. More generally, if n is any positive integer and $1 \leq i \leq n$, then the ith *standard (basis) vector* of $\mathbb{R}^n$ is the vector $\mathbf{e}_i$ that has 0's in all coordinate positions except the ith, where it has 1. Thus

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, \ldots, 0, 1).$$

The set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is called the *standard basis* of $\mathbb{R}^n$. (Note that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ *is* a basis of $\mathbb{R}^n$.)

Dimension

It is not hard to show that every nonzero vector space with a basis has infinitely many different bases. In this section we want to show that every vector space that can be spanned by finite many vectors has a basis. Before we prove this theorem we need another theorem which tells us something about the number of elements in a basis of vector space. Let's start by listing more bases for $\mathbb{R}^3$:

$$B_1 = \{(1,1,0),(1,0,1),(0,1,1)\}$$
$$B_2 = \{(1,2,3),(4,5,6),(7,8,0)\}$$
$$B_3 = \{(1,2,4),(1,3,9),(1,4,16)\}$$

It turns out that which ever basis of $\mathbb{R}^3$ we choose to construct, every such basis has exactly 3 elements. This is part of a general theorem which states that whenever a vector space has a basis all the bases in this vector space have the same number of vectors.

Theorem S1.5. *Let V be a finite vector space with $V = span\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$. If $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ is a set of linearly independent vectors then $m \le n$.*

Proof. The idea of the proof is to show that we can replace j vectors in the set $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ by vectors $\mathbf{u}_1,\ldots,\mathbf{u}_j$, so that the new set of n vectors $\{\mathbf{u}_1,\ldots,\mathbf{u}_j,\mathbf{v}_{j+1},\ldots,\mathbf{v}_n\}$ is again a spanning set of V. Then we show that we must exhaust the set of the $\mathbf{u}'s$ before we have exhausted the set of $\mathbf{v}'s$, that is, $m \le n$. One difficulty in the proof is that at every stage we need to renumber vectors $\mathbf{v}_1,\ldots,\mathbf{v}_n$ in order to avoid multiple index sets or renaming remaining vectors from $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$.

We start by distinguishing two cases:

Case 1: $\mathbf{u}_1 \in \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ Then $\mathbf{u}_1 = \mathbf{v}_i$ for some $1 \le i \le n$. We renumber vectors $\mathbf{v}_1,\ldots,\mathbf{v}_n$ so that $\mathbf{u}_1 = \mathbf{v}_1$ to obtain that $\{\mathbf{u}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\} = \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ is a spanning set with n elements.

Case 2: $\mathbf{u}_1 \notin \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$

Since $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ is a spanning set of V, by Theorem S1.2 the set of vectors $\{\mathbf{u}_1,\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ is linearly dependent. Moreover, since $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ is linearly independent, $\mathbf{u}_1 \ne \mathbf{0}$, and by the Linear Dependence Lemma we can remove one of the $\mathbf{v}'s$, say $\mathbf{v}_i$, from the spanning set $\{\mathbf{u}_1,\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ of V. After renumbering vectors $\mathbf{v}_1,\ldots,\mathbf{v}_n$ so that $\mathbf{v}_i$ becomes the first vector on the list, we obtain a new spanning set with n vectors, namely, $\{\mathbf{u}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$.

Now we suppose that we have added $\mathbf{u}'s$ and removed $\mathbf{v}'s$ so that the set

$$(*) \qquad \{\mathbf{u}_1,\ldots,\mathbf{u}_{j-1},\mathbf{v}_j,\ldots,\mathbf{v}_n\}$$

is a spanning set of V of length n. In order to add $\mathbf{u}_j$ and remove one of the vectors $\mathbf{v}_j,\ldots,\mathbf{v}_n$ we again need to distinguish two cases:

Case 1: $\mathbf{u}_j \in \{\mathbf{v}_j,\ldots,\mathbf{v}_n\}$ If $\mathbf{u}_j = \mathbf{v}_i$ for some $j \le i \le n$, we again renumber vectors $\mathbf{v}_j,\ldots,\mathbf{v}_n$, so that $\mathbf{u}_j = \mathbf{v}_j$ and obtain a spanning set $\{\mathbf{u}_1,\ldots,\mathbf{u}_j,\mathbf{v}_{j+1},\ldots,\mathbf{v}_n\}$ of n vectors.

Case 2: $\mathbf{u}_j \notin \{\mathbf{v}_j,\ldots,\mathbf{v}_n\}$

In this case we repeat the argument from above. The set $\{\mathbf{u}_1,\ldots,\mathbf{u}_{j-1},\mathbf{u}_j,\mathbf{v}_j,\ldots,\mathbf{v}_n\}$ is a spanning set of V with $n+1$ elements where $\mathbf{u}_j$ is a linear combination of vectors $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1},\mathbf{v}_j,\ldots,\mathbf{v}_n$. By Theorem S1.2 the set of $n+1$ vectors

$\{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}, \mathbf{u}_j, \mathbf{v}_j, \dots, \mathbf{v}_n\}$ is linear independent with $\mathbf{u}_1 \neq \mathbf{0}$. Since the set $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ is linearly independent, by the Linear Dependence Lemma there is a vector $\mathbf{v}_i$ where $j \leq i \leq n$ with $\mathbf{v}_i \in \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_j, \mathbf{u}_j, \mathbf{v}_j, \dots, \mathbf{v}_{i-1}\}$. Renumbering the renaming vectors $\mathbf{v}_j, \dots, \mathbf{v}_n$ so that $\mathbf{u}_j = \mathbf{v}_j$ the second part of the Linear Dependence Lemma yields that the set of n vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_j, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$ spans V.

The process stops when there are no $\mathbf{u}'s$ or no $\mathbf{v}'s$ left. If there are no $\mathbf{u}'s$ left then $m \leq n$, as desired. If there are no $\mathbf{v}'s$ left then $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a linearly independent spanning set of V. If $m > n$ then $\mathbf{u}_{n+1} \in \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$, a contradiction to the linear independence of the $\mathbf{u}'s$. Thus in this case $n = m$.

Definition. A vector space V is called *finite dimensional* if there is a finite subset $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ so that $V = \operatorname{span}(S) = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Corollary S1.6. *Every subspace of a finite dimensional vector space V is finite dimensional.*

Proof. Let $U \subseteq V$ be a subspace of V. If $U = \{\mathbf{0}\}$ we are done. Suppose $U \neq \{\mathbf{0}\}$ and take $\mathbf{u}_1 \in U$ with $\mathbf{u}_1 \neq \mathbf{0}$. If $U = \operatorname{span}\{\mathbf{u}_1\}$ we are done. If $U \neq \operatorname{span}\{\mathbf{u}_1\}$ take $\mathbf{u}_2 \in U - \operatorname{span}\{\mathbf{u}_1\}$. Again, if $U = \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2)$ we are done. If $U \neq \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ take $\mathbf{u}_3 \in U - \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ etc. This way we obtain a set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_n\} \subseteq U$ with $\mathbf{u}_j \notin \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}\}$. Since $\mathbf{u}_1 \neq \mathbf{0}$ by the Corollary S1.3 the set $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is linearly independent. Theorem S1.5 tells us that the process must stop after finitely many steps. Thus U has a finite spanning set.

Theorem S1.7. *Let V be a finite dimensional nonzero vector space. Then:*

(a) *Any linearly independent set in V is contained in a basis of V.*
(b) *Any spanning set of V contains a basis of V.*

Proof. (a) Let $\{\mathbf{u}_1, \dots, \mathbf{u}_m\} \subseteq V$ be a set of linearly independent vectors in V. If $V = \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ we are done. If not we expand the list by a vector $\mathbf{u}_{m+1} \in V - \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$. Again if $V = \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{m+1}\}$ we are done. If not take $\mathbf{u}_{m+2} \in V - \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{m+1}\}$ etc. This way we create a set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ with $\mathbf{u}_j \notin \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}\}$ for $2 \leq j \leq n$. By Corollary S1.3 this set of vectors is linearly independent. The process must stop after finitely many steps since any finite spanning set of V provides an upper bound to the length of a linearly independent set of vectors of V.

(b) Let $V = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. We may remove any $\mathbf{v}_i = \mathbf{0}$ from the spanning set and still have a spanning set. Thus we may assume that $\mathbf{v}_1 \neq \mathbf{0}$. If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a set of linearly independent vectors we are done. If not apply the Linear Dependence Lemma and remove one of the $\mathbf{v}_j$ where $2 \leq j \leq n$ so that $V = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$. Apply the same argument to the spanning set of $n-1$ vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$ and so on. The process stops when the reduced spanning set of vectors is linearly independent.

Theorem S1.8. *Every finite dimensional vector space has a basis.*

Proof. This is actually a corollary from Theorem S1.7. Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ be a spanning set of V. By Theorem S1.7(b) this set S contains a basis of V.

TheoremS1.9. *Any two bases of a finite dimensional vector space contain the same number of vectors.*

Proof. Let $B_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $B_2 = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be two bases of V. Then B_1 is a linearly independent set of V while B_2 is a spanning set of V. Thus by Theorem S1.5 $n \leq m$. Since B_1 and B_2 are each linearly independent spanning sets of V we can switch the role of B_1 and B_2 in the above argument, that is, B_2 is a linearly independent set and B_1 is a spanning set of V. Thus again by Theorem S1.5 $m \leq n$. Hence we have $n \leq m$ and $m \leq n$ which implies that $n = m$.

Definition. Let V be a finite dimensional vector space. If a basis of V consists of n vectors we say that V is a vector space of *dimension* n which is denoted by $\dim V = n$.

Corollary S1.10. *Let V be a vector space of dimension n. Then:*

(a) *Any linearly independent set of n vectors is a basis of V.*
(b) *Any spanning set of V with exactly n vectors is a basis of V.*

Proof. (a) Let $S = S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ be a linearly independent subset of V. By Theorem S1.7(a) this set can be extended to a basis of V. On the other hand any basis of V contains exactly n vectors. Thus S must be a basis of V
(b) If $T = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\} \subseteq V$ is a spanning set of V, by Theorem S.7(b) a subset of T is a basis of V. Any proper subset of T has fewer than n vectors. Thus T is a basis of V.