

SUPPLEMENT TO CHAPTER 3

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 . Similarly, in $\mathbb{M}(2, 2)$ consider the matrices:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

For all $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}(2, 2)$ we have that

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= aE_{11} + bE_{12} + cE_{21} + dE_{22}. \end{aligned}$$

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 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ a sequence of vectors in V . A vector $\mathbf{w} \in V$ is called a *linear combination of vectors* $\mathbf{v}_1, \dots, \mathbf{v}_n$ if there are scalars $r_1, \dots, r_n \in \mathbb{R}$ so that

$$\mathbf{w} = r_1\mathbf{v}_1 + \dots + r_n\mathbf{v}_n = \sum_{i=1}^n r_i\mathbf{v}_i.$$

The scalars r_1, \dots, 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 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and that every matrix of $\mathbb{M}(2, 2)$ is a linear combination of the matrices $E_{11}, E_{12}, E_{21}, E_{22}$.

Given any n vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ 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 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ vectors in V . The *span of the vectors* $\mathbf{v}_1, \dots, \mathbf{v}_n$ is the set of all linear combinations of these n vectors:

$$\begin{aligned} S = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} &= \{r_1\mathbf{v}_1 + \dots + r_n\mathbf{v}_n \mid r_1, \dots, r_n \in \mathbb{R}\} \\ &= \left\{ \sum_{i=1}^n r_i\mathbf{v}_i \mid r_1, \dots, r_n \in \mathbb{R} \right\}. \end{aligned}$$

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

Whenever we define certain subsets of a vector space the first question which arises is if these subsets relate in some ways to the vector space structure on the whole vector space? In almost all cases we want to know if these subsets are subspaces. Here is the answer for spanning sets:

Theorem 1.1. (Theorem 3.3 of the book) Let V be a vector space and $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ vectors in V . The spanning set of $\mathbf{v}_1, \dots, \mathbf{v}_n$

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

is a subspace of V .

Proof. In order to prove this theorem remember that by Theorem 1.11 (of the book) we have to show:

- (0) $S \neq \emptyset$
 - (i) S is closed under addition
 - (ii) S is closed under scalar multiplication
- (0)

$$\mathbf{0} = \sum_{i=1}^n 0\mathbf{v}_i = 0\mathbf{v}_1 + \dots + 0\mathbf{v}_n \in S$$

Thus S contains (at least) the zero vector and is not empty.

(i) Next we want to show that S is closed under vector addition. Let \mathbf{v}, \mathbf{w} be vectors in S . Since S is the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_n$, vectors \mathbf{v} and \mathbf{w} are linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_n$, that is, there are scalars $r_1, \dots, r_n \in \mathbb{R}$ and $s_1, \dots, s_n \in \mathbb{R}$ with

$$\mathbf{v} = \sum_{i=1}^n r_i \mathbf{v}_i \quad \text{and} \quad \mathbf{w} = \sum_{i=1}^n s_i \mathbf{v}_i.$$

Then

$$\begin{aligned} \mathbf{v} + \mathbf{w} &= \sum_{i=1}^n r_i \mathbf{v}_i + \sum_{i=1}^n s_i \mathbf{v}_i \\ &= \sum_{i=1}^n (r_i \mathbf{v}_i + s_i \mathbf{v}_i) \\ &= \sum_{i=1}^n (r_i + s_i) \mathbf{v}_i \end{aligned}$$

and the vector $\mathbf{v} + \mathbf{w}$ is a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Hence $\mathbf{v} + \mathbf{w} \in S$.

(ii) In order to show that S is closed under scalar multiplication, let $\mathbf{v} \in S$ and $c \in \mathbb{R}$. Then

$$\mathbf{v} = \sum_{i=1}^n r_i \mathbf{v}_i$$

where $r_i \in \mathbb{R}$ and

$$\begin{aligned} c\mathbf{v} &= c \sum_{i=1}^n r_i \mathbf{v}_i \\ &= \sum_{i=1}^n c(r_i \mathbf{v}_i) \\ &= \sum_{i=1}^n (cr_i) \mathbf{v}_i \end{aligned}$$

and $c\mathbf{v} \in S = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. This proves that $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a subspace of V .

The examples at the beginning of this section show that $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ span the vector space \mathbb{R}^3 and that $E_{11}, E_{12}, E_{21}, E_{22}$ span $\mathbb{M}(2, 2)$:

$$\mathbb{R}^3 = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \quad \text{and} \quad \mathbb{M}(2, 2) = \text{span}\{E_{11}, E_{12}, E_{21}, E_{22}\}.$$

You can easily convince yourself that when you leave out one of the vectors in the sequence $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (in $E_{11}, E_{12}, E_{21}, E_{22}$ in case of $\mathbb{M}(2, 2)$) the vector (resp. matrix) left out is not in the spanning set of the remaining vectors, for example:

$$\mathbf{e}_3 = (0, 0, 1) \notin \text{span}\{\mathbf{e}_1, \mathbf{e}_2\} \quad \text{and} \quad E_{21} \notin \text{span}\{E_{11}, E_{12}, E_{22}\}.$$

We say that $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ (resp. $\{E_{11}, E_{12}, E_{21}, E_{22}\}$) are *minimal* spanning sets of \mathbb{R}^n ($\mathbb{M}(2, 2)$, resp.).

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 $\mathbf{v} \neq \mathbf{w}$. Similarly, every matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}(2, 2)$ can be written in a unique way as a linear combination of the matrices $E_{11}, E_{12}, E_{21}, E_{22}$.

On the other hand, the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{u} = (2, 0, 1)$ also span \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:

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

Similarly, the matrices $E_{11}, E_{12}, E_{21}, E_{22}, D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ span $\mathbb{M}(2, 2)$ and the ma-

trix $A = \begin{bmatrix} -1 & 2 \\ 4 & 0 \end{bmatrix}$ can be written as a linear combination of matrices $E_{11}, E_{12}, E_{21}, E_{22}, D$ in at least two different ways.

In the following sections we are going to address the questions:

(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 vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$,

say:

$$\mathbf{v} = \sum_{i=1}^n r_i \mathbf{v}_i$$

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

The best tool to study these questions is the notion of linear independence which will be introduced next.

LINEAR INDEPENDENCE

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

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

implies that $r_1 = r_2 = \dots = r_n = 0$. 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 sequence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ the zero vector can *always* be written as

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

The definition of linear independence states that the sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent if $(*)$ is *the one and only* way the zero vector can be written as a linear combination of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Thus the sequence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent if and only if the zero vector can be written *in a unique way* (namely $(*)$) as a linear combination of the sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$.

The following question seems natural: Suppose that the sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent. We know then that the zero vector can be written in a unique way as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_n$. What about other linear combinations? Can other vectors in $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ also be written as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_n$ with unique coefficients? The answer is yes and even more can be shown:

Theorem 1.2. *Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a sequence of vectors in a vector space V . The following are equivalent:*

(1) $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.

(2) Every vector in $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ has a unique expression as a linear combination of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$.

(3) No vector \mathbf{v}_i is a linear combination of the other vectors $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n$.

(Compare with Theorem 3.5 of the book.)

Proof. (1) \Rightarrow (2): Let $\mathbf{w} \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and assume that

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

where $r_i, s_i \in \mathbb{R}$. We have to show that $r_i = s_i$ for all $1 \leq i \leq n$. By subtracting \mathbf{w} from itself we obtain:

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

Since the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent:

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

(2) \Rightarrow (3) : Proof by contrapositive: We show that if a vector \mathbf{v}_i is a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n$, then there is a vector in $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ which can be written in at least two different ways as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_n$. Suppose that

$$\mathbf{v}_i \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n\}.$$

Then there are scalars $r_i \in \mathbb{R}$ so that:

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

On the other hand:

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

which shows that the vector $\mathbf{v}_i \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ can be written in two different ways as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_n$. Thus if (3) is false, then (2) is false.

(3) \Rightarrow (1) : Proof by contrapositive: Suppose that vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly dependent and that

$$\mathbf{0} = r_1 \mathbf{v}_1 + \dots + r_n \mathbf{v}_n = \sum_{i=1}^n r_i \mathbf{v}_i$$

with $r_i \neq 0$ for some $1 \leq i \leq n$. Then

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

and therefore

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

and $\mathbf{v}_i \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n\}$. This shows that if (1) is false then so is (3).

Theorem 1.4. *Let V be a vector space and $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ a sequence of linearly independent vectors. If $\mathbf{w} \notin \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ then the sequence $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}$ is linearly independent.*

(This is the expansion lemma 3.12 of the book.)

Proof. Let $r_i, t \in \mathbb{R}$ be scalars with

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

Then

$$t \mathbf{w} = (-r_1) \mathbf{v}_1 + \dots + (-r_n) \mathbf{v}_n$$

and if $t \neq 0$ then

$$\mathbf{w} = (-r_1/t) \mathbf{v}_1 + \dots + (-r_n/t) \mathbf{v}_n \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\},$$

a contradiction. Thus we must have that $t = 0$. But then

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

and also $r_1 = r_2 = \dots = r_n = 0$ since the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.

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

$$\mathbf{u} = (2, 0, 1) \in \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \mathbb{R}^3.$$

A similar argument shows that the sequence of matrices $E_{11}, E_{12}, E_{21}, E_{22}, D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ from section 1 is linearly dependent while $E_{11}, E_{12}, E_{21}, E_{22}$ is linearly independent.

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

Linear Dependence Lemma 1.5. *If $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ is a sequence of linearly dependent vectors in a vector space V with $\mathbf{v}_1 \neq \mathbf{0}$, then there is a $j \in \{2, \dots, n\}$ so that*

- (a) $\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$
- (b) $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$.

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

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

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

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

which shows that $\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$.

(b) By (a) we know that there is an integer $j \in \{2, \dots, n\}$ so that $\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$. Obviously, $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\} \subseteq \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. In order to show the other inclusion we use the fact that $\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$ and write

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

with $r_i \in \mathbb{R}$. Let $\mathbf{v} \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ then

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

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

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

and $\mathbf{v} \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$. This shows (b).

Corollary 1.6. *If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a set of vectors in V with $\mathbf{v}_1 \neq \mathbf{0}$ and $\mathbf{v}_j \notin \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}\}$ for all $2 \leq j \leq n$ then $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a set of linearly independent vectors in V .*

BASES

Definition. Let V be a vector space and $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ 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 $V = \text{span}(B) = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Definition. A vector space V is called *finite dimensional* if there is a finite sequence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ so that $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

In this course most of our focus will be on the study of finite dimensional vector spaces, like the vector spaces $\mathbb{R}^n, \mathbb{P}_n$ and $\mathbb{M}(m, n)$. We also know a few examples of vector spaces which are not finite dimensional, for example, \mathbb{P} , the vector space of all polynomials or $\mathbb{C}([0, 1])$, the vector space of all continuous functions from the interval $[0, 1]$ into \mathbb{R} .

This first theorem on finite dimensional vector spaces is the foundation for rest of the course:

Theorem 1.7. *Every finite dimensional vector space has a basis.*

Proof. Let V be a vector space and $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ a sequence of vectors of V with $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

If the sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent, then the set $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of V and we are done.

If vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly dependent we need to distinguish two more cases.

In the first case we have that $\mathbf{v}_1 = \mathbf{0}$, then

$$V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = \text{span}\{\mathbf{v}_2, \dots, \mathbf{v}_n\}.$$

Thus we may replace the spanning sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$ by the shorter spanning sequence $\mathbf{v}_2, \dots, \mathbf{v}_n$.

In the second case where $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly dependent and the first vector $\mathbf{v}_1 \neq \mathbf{0}$ we apply the Linear Dependence Lemma 1.5. According to Lemma 1.5 there is an integer $2 \leq j \leq n$ so that

$$V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}.$$

Again we can replace the spanning sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$ by the shorter spanning sequence $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n$.

In summary, we have just proven that if the spanning sequence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly dependent, then there is a shorter spanning sequence, say $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ (after renumbering the vectors).

We now apply the same arguments as before to the new and shorter spanning sequence $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ and obtain (as before):

If the sequence $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ is linearly independent then $B' = \{\mathbf{v}_1, \dots, \mathbf{v}_{n-1}\}$ is a basis of V . If the spanning sequence $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ is linearly dependent we use the same arguments as before to show that it can be shortened to a spanning sequence of $n-2$ vectors, say (after renumbering) $\mathbf{v}_1, \dots, \mathbf{v}_{n-2}$ is a spanning sequence of V .

This process has to stop since there are only finitely many vectors in any spanning sequence. It stops when we have found a spanning sequence $\mathbf{v}_1, \dots, \mathbf{v}_r$ of linearly independent vectors. This means that the set $B'' = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ is a basis of the vector space V .

Corollary 1.8. (*Contraction theorem 3.11 of the book*) Suppose that the sequence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ spans the vector space V . Then some subset of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of V .

Proof. The proof is exactly the same as the proof of Theorem 1.7.

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 . Similarly, the set $B' = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ is a basis of the vector space $\mathbb{M}(2, 2)$.

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 i th *standard (basis) vector* of \mathbb{R}^n is the vector \mathbf{e}_i that has 0's in all coordinate positions except the i th, where it has 1. Thus

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

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

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

Theorem 1.9. Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 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 $r_1, \dots, r_n \in \mathbb{R}$ so that $\mathbf{v} = r_1\mathbf{v}_1 + \dots + r_n\mathbf{v}_n$.
- (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 $r_1, \dots, r_n \in \mathbb{R}$ so that

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

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

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

then

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

Since the sequence $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent, we have that $r_1 = s_1, r_2 = s_2, \dots, r_n = s_n$.

(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 $A = \{\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$. Then $V = \text{span}(A)$, $\mathbf{v}_1 \in \text{Span}(A)$, and there are scalars $r_2, \dots, r_n \in \mathbb{R}$ so that

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

This is one way to write \mathbf{v}_1 as a linear combination of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Another is $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \dots + 0\mathbf{v}_n$, 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 $r_1, \dots, r_n \in \mathbb{R}$ so that

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

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

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

and the sequence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{v}$ 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 $i = 1, 2, \dots, n$ 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 $t, r_1, \dots, r_n \in \mathbb{R}$, not all 0, so that

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

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

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

DIMENSION

It is not hard to show that every nonzero vector space which has a basis must have infinitely many different bases. Here are some examples for bases in \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. Before we can prove this theorem we need:

Theorem 1.10. (see the Comparison Theorem 3.11 of the book) *Let V be a finite vector space with $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. If $\mathbf{u}_1, \dots, \mathbf{u}_m$ is a sequence of linearly independent vectors then $m \leq n$.*

Proof. The idea of the proof is to show that we can replace j vectors in the set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ by vectors $\mathbf{u}_1, \dots, \mathbf{u}_j$, so that the new set of n vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_j, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$ 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, $m \leq n$. One difficulty in the proof is that at every stage we need to renumber vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in order to avoid multiple index sets or renaming remaining vectors from $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

We start by distinguishing two cases:

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

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

Since $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a spanning set of V , by Theorem 1.2 the set of vectors $\{\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly dependent. Moreover, since $\mathbf{u}_1, \dots, \mathbf{u}_m$ are linearly independent, $\mathbf{u}_1 \neq \mathbf{0}$, and by the Linear Dependence Lemma we can remove one of the \mathbf{v} 's, say \mathbf{v}_i , from the spanning set $\{\mathbf{u}_1, \mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V . After renumbering vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ so that \mathbf{v}_i becomes the first vector on the list, we obtain a new spanning set with n vectors, namely, $\{\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$.

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

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

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

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

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

In this case we repeat the argument from above. The set $\{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}, \mathbf{u}_j, \mathbf{v}_j, \dots, \mathbf{v}_n\}$ is a spanning set of V with $n+1$ elements where \mathbf{u}_j is a linear combination of vectors $\mathbf{u}_1, \dots, \mathbf{u}_{j-1}, \mathbf{v}_j, \dots, \mathbf{v}_n$. By Theorem 1.2 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 linearly 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 $\mathbf{v}_i \in \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_j, \mathbf{v}_j, \dots, \mathbf{v}_{i-1}\}$. Renumbering the remaining 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 $\mathbf{u}_{n+1} \in \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$, a contradiction to the linear independence of the \mathbf{u} 's. Thus in this case $n = m$.

Corollary 1.11. (see Theorem 3.14 of the book) *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 $U = \text{span}\{\mathbf{u}_1\}$ we are done. If $U \neq \text{span}\{\mathbf{u}_1\}$ take $\mathbf{u}_2 \in U - \text{span}\{\mathbf{u}_1\}$. Again, if $U = \text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ we are done. If $U \neq \text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ take $\mathbf{u}_3 \in U - \text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ etc. This way we obtain a set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_n\} \subseteq U$ with $\mathbf{u}_j \notin \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}\}$. Since $\mathbf{u}_1 \neq \mathbf{0}$ by the Theorem 1.3 the sequence $\mathbf{u}_1, \dots, \mathbf{u}_n$ is linearly independent. Theorem 1.10 tells us that the process must stop after finitely many steps. Thus U has a finite spanning set.

Theorem 1.12. *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 .*

(see the Contraction Theorem 3.11 and the Expansion Theorem 3.13 of the book.)

Proof. (a) Let $\{\mathbf{u}_1, \dots, \mathbf{u}_m\} \subseteq V$ be a set of linearly independent vectors in V . If $V = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ we are done. If not we expand the list by a vector

$\mathbf{u}_{m+1} \in V - \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$. Again if $V = \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{m+1}\}$ we are done. If not take $\mathbf{u}_{m+2} \in V - \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{m+1}\}$ etc. This way we create a set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ with $\mathbf{u}_j \notin \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_{j-1}\}$ for $2 \leq j \leq n$. By Theorem 1.3 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 $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. 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 $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n\}$. 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 1.13. (see Theorem 3.10 of the book) *Any two bases of a finite dimensional vector space contain the same number of vectors.*

Proof. Let $B_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $B_2 = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ 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 1.11 $n \leq m$. 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 1.11 $m \leq n$. 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 1.14. (see Theorem 3.15 of the book) *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 $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ be a linearly independent subset of V . By Theorem 1.12(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 $T = \{\mathbf{u}_1, \dots, \mathbf{u}_n\} \subseteq V$ is a spanning set of V , by Theorem 1.12(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 .