LECTURE I: INTRODUCTION TO VECTOR SPACES
MAT 204 - FALL 2006
PRINCETON UNIVERSITY

ALFONSO SORRENTINO

[Read also §2.1, 2.3 and appendix A in the textbook]

1. VECTOR SPACES AND SUBSPACES

Roughly speaking, a *vector space* is a set of *vectors* together with rules for vector addition and multiplication by real numbers. These operations must produce vectors in the space and must satisfy some conditions.

Let us define it more clearly and precisely.

Consider a set V endowed with two operations:

- *Addition*: given two vectors v_1 and v_2 in V, it associates a new vector in V, that we will denote $v_1 + v_2$; *i.e.*,

$$+ : V \times V \longrightarrow V$$
$$(v_1, v_2) \longmapsto v_1 + v_2 \,.$$

- *Scalar multiplication*: given a vector v in V and a real number c, it associates a new vector in V, that we will denote $c \cdot v$ (or simply cv); *i.e.*,

$$\cdot : \mathbb{R} \times V \longrightarrow V$$
$$(c, v) \longmapsto c \cdot v \,.$$

Definition. A triple $(V, +, \cdot)$ is a **(real) vector space**[1] if:

i) for any v_1, v_2, $v_3 \in V$, $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ (*associative law*);
ii) there exists a *zero vector* $0 \in V$, such that $v + 0 = 0 + v = v$ for any $v \in V$;
iii) for any $v \in V$, there exists "$-v \in V$" satisfying $-v + v = v + (-v) = 0$;
iv) for any v_1, $v_2 \in V$, $v_1 + v_2 = v_2 + v_1$ (*commutative law*);
v) for any $v \in V$ and c, $d \in \mathbb{R}$, $(c + d) \cdot v = c \cdot v + d \cdot v$ (*distributive law I*);
vi) for any v_1, $v_2 \in V$ and $c \in \mathbb{R}$, $c \cdot (v_1 + v_2) = c \cdot v_1 + c \cdot v_2$ (*distributive law II*);
vii) for any $v \in V$ and c, $d \in \mathbb{R}$, $(cd) \cdot v = c \cdot (d \cdot v)$;
viii) for any $v \in V$, $1 \cdot v = v$.

The elements in V are called *vectors* and the elements in $\mathbb{R}$ *scalars*.

Example. [2]

(1) $\mathbb{R}$ itself is a vector space;
(2) for any positive integer n, $\mathbb{R}^n$ (*i.e.*, the set of real n-ples) is a vector space;
(3) in general, if V is a vector space, V^n (*i.e.*, the set of n-ples whose components are elements of V) is still a vector space;
(4) the space of functions $f : [0, 1] \longrightarrow \mathbb{R}$ is a vector space.

From the properties above, one can deduce the following.

Proposition 1. *Let $(V, +, \cdot)$ be a vector space. Then,*

---

[1]See also problem 5, pag.74.
[2]See also examples in § 2.1.

(1) $0 \cdot v = 0$, *for any* $v \in V$;
(2) $(-c) \cdot v = c \cdot (-v)$, *for any* $c \in \mathbb{R}$ *and* $v \in V$;
(3) $c \cdot 0 = 0$, *for any* $c \in \mathbb{R}$;
(4) *if* $c \cdot v = 0$, *then either* $c = 0$ *or* $v = 0$.

Proof. (1) Using the properties viii), v) and ii), one gets

$$\begin{aligned} v + 0 \cdot v &= 1 \cdot v + 0 \cdot v = (1 + 0) \cdot v = 1 \cdot v = \\ &= v = v + 0 \, ; \end{aligned}$$

adding $-v$ to both sides of the equality:

$$-v + v + 0 \cdot v = -v + v + 0 \qquad \Longrightarrow \qquad 0 \cdot v = 0 \, .$$

(2) By definition, $-(c \cdot v)$ is an element of V that satisfies iii) (with v replaced by $c \cdot v$). Let us show that iii) holds, if $-(c \cdot v) = (-c) \cdot v$. Using v) and the previous property:

$$(-c) \cdot v + c \cdot v = (-c + c) \cdot v = 0 \cdot v = 0 \, .$$

(3) In fact, for any $v \in V$:

$$\begin{aligned} c \cdot 0 &= c \cdot (v - v) = c \cdot v + c \cdot (-v) = \\ &= c \cdot v + c \cdot [(-1) \cdot v] = c \cdot v + (-c) \cdot v = \\ &= c \cdot v + -c \cdot v = 0 \end{aligned}$$

(we have used properties iii), v), vii), viii) and the previous results in this proposition).

(4) If $c \neq 0$, then it makes sense to consider $\frac{1}{c}$; therefore:

$$\begin{aligned} v &= 1 \cdot v = \left(\frac{1}{c} c\right) \cdot v = \frac{1}{c} (c \cdot v) = \\ &= \frac{1}{c} \cdot 0 = 0 \, . \end{aligned}$$

$\square$

Another important concept is the concept of *subspace.* In few words, a subspace of a vector space is a subset of V that is "closed" under addition and scalar multiplication (roughly speaking, the results of the operations remain in this subset). These operations follow the rule of the host space, keeping us inside the subspace.

Definition. Let $(V, +, \cdot)$ be a vector space and let $W \subseteq V$ a nonempty subspace of V. W is said to be a vector subspace of V, if it is a vector space with the induced operations; *i.e.* $(W, +, \cdot)$ still satisfies i)-viii), where $+$ and $\cdot$ are the same operations of V, restricted to W.

A useful characterization is the following.

Proposition 2. *Let W be a nonempty subset of a vector space V. W is a vector subspace if and only if the following two conditions are satisfied:*

a) $w_1 + w_2 \in W$, *for any* $w_1, w_2 \in W$;
b) $c \cdot w \in W$, *for any* $c \in \mathbb{R}$ *and* $w \in W$.

Proof. The proof of ($\Rightarrow$) is obvious (it comes directly from the definition above and the one of vector space).
Let us prove the other implication ($\Leftarrow$).
From a) and b) it follows that the restricted addition and scalar multiplication are well defined on W (*i.e.*, the result vector is still in W). Moreover, from b), it is easy to deduce that for any $w \in W$:

$$0 = 0 \cdot w \in W \qquad \text{and} \qquad -w = (-1) \cdot w \in W \, .$$

Therefore, i)-iv) hold. In particular, since axioms v)-viii) hold in V, they keep on being satisfied also in V. $\qquad\square$

Remark.
- One can easily show that properties i) and ii) above, are equivalent to the unique condition:

$$c \cdot w_1 + d \cdot w_2 \in W, \qquad \text{for any } w_1, w_2 \in W \quad \text{and } c, d \in \mathbb{R}.$$

- The zero vector will belong to every subspace. That comes from rule ii): choose the scalar to be $c = 0$.

Here there are some examples of vector subspaces for a vector space $(V, +, \cdot)$.

Example.
(1) The subset containing only the zero vector, $Z = \{0\}$ and the whole space V are trivial subspaces (Z is the smallest possible subspace and V the biggest one).

(2) For any $v_0 \in V$, the set

$$\langle v_0 \rangle := \{c \cdot v_0, \quad \forall c \in \mathbb{R}\}$$

is a vector subspace. In fact, for any $c \cdot v_0, d \cdot v_0 \in \langle v_0 \rangle$ and $a, b \in \mathbb{R}$:

$$a \cdot (c \cdot v_0) + b \cdot (c \cdot v_0) = (ac + bd) \cdot v_0 \in \langle v_0 \rangle.$$

This subspace is called *subspace generated by v_0*. Observe that $Z = \langle 0 \rangle$. More generally, if $v_1, \ldots, v_k$ are k fixed vectors in V, one can consider the set

$$\langle v_1, \ldots, v_k \rangle := \{c_1 \cdot v_1 + \ldots + c_k \cdot v_k, \quad \forall c_1, \ldots, c_k \in \mathbb{R}\}.$$

This is a vector subspace and is called *subspace generated by $v_1, \ldots, v_k$*.

(3) (Exercise) If W_1 and W_2 are two vector subspaces, then the *intersection space* $W_1 \cap W_2$ is still a subspace (the intersection is the subset containing all the elements that are both in W_1 and W_2).

(4) (Exercise) If W_1 and W_2 are two vector subspaces, then the *sum space*

$$W_1 + W_2 := \{w_1 + w_2, \quad w_1 \in W_1, w_2 \in W_2\},$$

is a vector subspace.
Observe that $W_1 + W_2 \supseteq W_1 \cup W_2$. In fact, in general the union might not be a subspace. One can show that $W_1 + W_2$ is the smallest subspace containing the union $W_1 \cup W_2$ [Exercise].

Proposition 3. *Let $(V, +, \cdot)$ be a vector space and W_1, W_2 two subspaces. The following two conditions are equivalent:*

i) $W_1 \cap W_2 = \langle 0 \rangle$;

ii) *for any $w \in W_1 + W_2$, there exists a unique couple $(w_1, w_2) \in W_1 \times W_2$, such that $w = w_1 + w_2$ (i.e., any vector in the sum space can be written uniquely as the sum of vectors in W_1 and W_2).*

Proof.
i) $\Rightarrow$ ii) Suppose that a vector $w \in W_1 + W_2$ can be written in two ways:

$$w = u_1 + u_2 = v_1 + v_2, \qquad \text{with } u_1, v_1 \in W_1 \quad \text{and } u_2, v_2 \in W_2.$$

Hence, $u_1 - v_1 = v_2 - u_2 \in W_1 \cap W_2 = \langle 0 \rangle$, and consequently $u_1 = v_1$ and $u_2 = v_2$.

i) $\Leftarrow$ ii) Suppose by contradiction that there exists $0 \neq w \in W_1 \cap W_2$. Therefore,

$$w = 0 + w = w + 0 \in W_1 + W_2;$$

this leads to a contradiction, since it implies that w can be decomposed in two different ways, as a vector of $W_1 + W_2$ $\qquad\square$

Definition. Let W_1, W_2 be two vector subspaces of a vector space V.

- W_1 and W_2 are said *supplementary subspaces* if
$$W_1 + W_2 = V \qquad \text{and} \qquad W_1 \cap W_2 = \langle 0 \rangle \,.$$

  From the proposition above, it follows immediately that this definition is equivalent to say that every vector in V can be decomposed uniquely as the sum of a vector in W_1 and one in W_2.
- More generally, if $W_1 \cap W_2 = \langle 0 \rangle$, then the sum subspace $W_1 + W_2$ is called *direct sum* and it is denoted $W_1 \oplus W_2$.

2. Linear independence, set of generators and basis

Definition. Let $(V, +, \cdot)$ be a vector space, $v_1, \ldots, v_k \in V$ and $c_1, \ldots, c_k \in \mathbb{R}$. The vector
$$c_1 \cdot v_1 + \ldots + c_k \cdot v_k$$
is called *linear combination of $v_1, \ldots, v_k$ with coefficients $c_1, \ldots, c_k$.*

Remark. Observe that for any k vectors $v_1, \ldots, v_k \in V$, obviously
$$0 \cdot v_1 + \ldots + 0 \cdot v_k = 0 \,,$$
i.e., the zero vector is linear combination of the vectors $v_1, \ldots, v_k$ (with coefficients all equal to zero). The linear combination $0 \cdot v_1 + \ldots + 0 \cdot v_k$ is called *trivial linear combination of $v_1, \ldots, v_k$.*

A natural question is: given any k vectors $v_1, \ldots, v_k \in V$, is it possible to obtain the zero vector as a non-trivial linear combination of them (*i.e.*, with at least one of the coefficients different from zero)?
Consider the following examples.

Example.   a) Let $0 \neq v \in V$ and $0 \neq c \in \mathbb{R}$. Consider the two vectors v and $c \cdot v$; one has:
$$0 = c \cdot v + (-1) \cdot (c \cdot v) \,,$$
consequently these two vectors admit a non trivial linear combination of the zero vector (with coefficients c and -1).
   b) In $\mathbb{R}^2$ consider the vectors $v_1 = (1, 2)$ and $v_2 = (0, 1)$. Suppose that $0 = c_1 \cdot v_1 + c_2 \cdot v_2 = c_1 \cdot (1, 2) + c_2 \cdot (0, 1) = (c_1, 2c_1 + c_2)$. It follows that $c_1 = 0$ and $2c_1 + c_2 = 0$, and consequently $c_1 = c_2 = 0$.
   Hence, v_1 and v_2 do not admit a non trivial combination of 0.

These examples justify the following definition.

Definition. Let V be a vector space and let $v_1, \ldots, v_k \in V$. These vectors are said to be *linearly dependent* if they admit a trivial linear combination of the zero vector (*i.e.*, there exist $c_1, \ldots, c_k \in \mathbb{R}$, with at least one different from zero, such that $c_1 \cdot v_1 + \ldots + c_k \cdot v_k = 0$).
Otherwise, they are said to be *linearly independent*, *i.e.*, the only linear combination that gives the zero vector is the trivial one:
$$c_1 \cdot v_1 + \ldots + c_k \cdot v_k = 0 \qquad \Longrightarrow \qquad c_1 = \ldots = c_k = 0 \,.$$

Remark.   i) A single vector $v \in V$ is linearly dependent if and only if $v = 0$.
     In fact, if v is linearly dependent, then there exists $c \neq 0$ such that $c \cdot v = 0$. From prop. 1 (iv), it follows that $v = 0$. Conversely, if $v = 0$ then $1 \cdot v = 0$, hence 0 is linearly dependent.
   ii) Given k linearly independent vectors, any their subset still consists of linearly independent vectors.
     Showing this claim is equivalent to show (for $0 < s < k$):
$$v_1, \ldots, v_s \quad \text{linearly dependent} \quad \Longrightarrow \quad v_1, \ldots, v_k \quad \text{linearly dependent} \,.$$

In fact, if there exists a linear combination

$$c_1 \cdot v_1 + \ldots + c_s \cdot v_s = 0$$

with some of the c_i's different from zero, then it suffices to consider

$$c_1 \cdot v_1 + \ldots + c_s \cdot v_s + 0 \cdot v_{s+1} + \ldots 0 \cdot v_k = 0 \,.$$

iii) From a geometrical point of view, we can think of the linear dependence in the following way:
   - two vectors v_1, v_2 are linearly dependent if and only if they are *parallel*;
   - three vectors v_1, v_2, v_3 are linearly dependent if and only if they are *coplanar* (*i.e.*, contained in a same plane).

Proposition 4. *Let $v_1, \ldots, v_k \in V$, with $k \geq 2$. They are linearly dependent if and only if at least one of them can be written as a linear combination of the remainings.*

Proof. ($\Rightarrow$) Suppose that there exist $c_1, \ldots, c_k \in \mathbb{R}$ (not all zeros) such that $c_1 \cdot v_1 + \ldots + c_k \cdot v_k = 0$. Without any loss of generality, we can assume $c_1 \neq 0$. Therefore,

$$c_1 \cdot v_1 = -c_2 \cdot v_2 - \ldots - c_k \cdot v_k$$

and multiplying by $\frac{1}{c_1}$:

$$v_1 = -\frac{c_2}{c_1} \cdot v_2 - \ldots - \frac{c_k}{c_1} \cdot v_k \,.$$

($\Leftarrow$) If, for instance, v_1 is a linear combination of $v_2, \ldots, v_k$, then (for suitable $a_2, \ldots, a_k \in \mathbb{R}$)

$$v_1 = a_2 \cdot v_2 + \ldots + a_k \cdot v_k \,,$$

and it follows that:

$$1 \cdot v_1 - a_2 \cdot v_2 - \ldots - a_k \cdot v_k = 0 \,.$$

$\square$

Proposition 5. *Let $v_1, \ldots, v_k \in V$. They are linearly independent if and only if the following condition holds:*

$$if \quad \sum_{i=1}^{k} c_i \cdot v_i = \sum_{i=1}^{k} d_i \cdot v_i \quad \Longrightarrow \quad c_1 = d_1, \ldots, c_k = d_k \,.$$

Proof. ($\Rightarrow$) From the hypothesis, it follows that

$$\sum_{i=1}^{k} (c_i - d_i) \cdot v_i = 0 \,;$$

therefore, since the vectors are linearly independent:

$$c_1 - d_1 = 0, \ldots, c_k - d_k = 0 \,.$$

($\Leftarrow$) Suppose that $\sum_{i=1}^{k} a_i \cdot v_i = 0$. Since $0 = \sum_{i=1}^{k} 0 \cdot v_i$, it follows from the assumption that $a_1 = \ldots = a_k = 0$ and consequently the vectors are linearly independent.

$\square$

Let us know introduce the concept of *set of generators*, or what means for a set of vectors to *span* a space.

We have defined in the previous section the space generated by k vectors. If $v_1, \ldots, v_k$ are k fixed vectors in V, one can consider the set

$$\langle v_1, \ldots, v_k \rangle := \{ c_1 \cdot v_1 + \ldots + c_k \cdot v_k, \quad \forall c_1, \ldots, c_k \in \mathbb{R} \} \,.$$

This is a vector subspace and is called *subspace generated by* $v_1, \ldots, v_k$.

Observe that $\langle v_1, \ldots, v_k \rangle$ is the smallest vector subspace containing $v_1, \ldots, v_k$; in fact, any other vector subspace W that contained $v_1, \ldots, v_k$, should contain any their linear combination, consequently $W \supseteq \langle v_1, \ldots, v_k \rangle$.

Moreover, $\langle v_1, \ldots, v_k \rangle$ coincides with the intersection of all vector subspaces containing $v_1, \ldots, v_k$:

$$\langle v_1, \ldots, v_k \rangle = \bigcap_{W \ni v_1, \ldots, v_k} W \, .$$

In fact, the intersection of all vector subspace is still a vector subspace. Then,

$$\bigcap_{W \ni v_1, \ldots, v_k} W \supseteq \langle v_1, \ldots, v_k \rangle \, ;$$

on the other hand, $\langle v_1, \ldots, v_k \rangle$ is one of the subspaces $W's$ in the intersection, therefore

$$\bigcap_{W \ni v_1, \ldots, v_k} W \subseteq \langle v_1, \ldots, v_k \rangle \, .$$

Definition. Let V be a vector space and $v_1, \ldots, v_k \in V$. If $\langle v_1, \ldots, v_k \rangle = V$, we will say that $\{v_1, \ldots, v_k\}$ is a *set of generators* of V (or *spanning set* of V).

More generally, let W be a subspace of V and let $w_1, \ldots, w_s \in W$. If $\langle w_1, \ldots, w_s \rangle = W$, we will say that $\{w_1, \ldots, w_s\}$ is a *set of generators* of W (or *spanning set* of W).

We have introduced all the ingredients that we need to define a *basis* of a vector space. In few words, a basis of a vector space is a set of generators, made of linearly independent vectors. More formally:

Definition. Let V be a vector space and let $v_1, \ldots, v_n \in V$. We will say that $\{v_1, \ldots, v_n\}$ is a *basis* of V, if:

    i) $v_1, \ldots, v_n$ are linearly independent;
    ii) $\langle v_1, \ldots, v_n \rangle = V$.

Why should one interested in considering a basis on a vector space?

A basis allows us to associate to each vector its coordinates (with respect to the chosen basis). The following proposition will make it more clear.

Proposition 6. *Let V be a vector space and $v_1, \ldots, v_n \in V$. The following conditions are equivalent:*

    i) *$\{v_1, \ldots, v_n\}$ is a basis of V;*
    ii) *every vector $v \in V$ can be written uniquely as a linear combination of $v_1, \ldots, v_n$.*

Proof. i) $\Rightarrow$ ii) Since $\{v_1, \ldots, v_n\}$ are a set of generators of V, then $v = \sum_{i=1}^{n} c_i \cdot v_i$ (for suitable $c_1, \ldots, c_n \in \mathbb{R}$). Using the fact that the vectors are linearly independent and prop. 5 it follows that this linear combination is unique.

i) $\Leftarrow$ ii) By hypothesis, every vector $v \in V$ can be written as $v = \sum_{i=1}^{n} c_i \cdot v_i$. Therefore, $\{v_1, \ldots, v_n\}$ is a set of generators of V. We need to show that they are linearly independent. In fact, if $\sum_{i=1}^{n} c_i \cdot v_i = 0$, then $\sum_{i=1}^{n} c_i \cdot v_i = \sum_{i=1}^{n} 0 \cdot v_i$, and by assumption $c_1 = \ldots = c_n = 0$. This completes the proof. $\square$

In other words, the above proposition shows that there exists a bijection

$$
\begin{array}{ccc}
V & \longrightarrow & \mathbb{R}^n \\
v & \longmapsto & (c_1, \ldots, c_n)
\end{array}
$$

such that $v = \sum_{i=1}^{n} c_i \cdot v_i$.

This bijection is clearly dependent on the choice of the basis $\{v_1, \ldots, v_n\}$ of V and allows to associate to every vector a n-ple of real numbers, that we will call *coordinates*.

Definition. Let V a vector space and $\{v_1, \ldots, v_n\}$ a basis. For any $v = \sum_{i=1}^{n} c_i \cdot v_i$, the n-ple $(c_1, \ldots, c_n) \in \mathbb{R}^n$ is called *coordinates of v with respect to the basis* $\{v_1, \ldots, v_n\}$.

Remark. In $V = \mathbb{R}^n$ we can consider n particular vectors:

$$e_1 = (1, 0, 0, \ldots, 0), \; e_2 = (0, 1, 0, \ldots, 0), \; \ldots, \; e_n = (0, 0, 0, \ldots, 0, 1).$$

[Exercise] One can show that:

(1) $\{e_1, \ldots, e_n\}$ is a set of generators of $\mathbb{R}^n$;
(2) they are linearly independent.

Therefore, $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{R}^n$. In this case, the coordinates of a vector $a = (a_1, \ldots, a_n)$ with respect to this basis, are exactly its n components. This fact makes this basis the "simplest" possible, and in some sense the "most natural" among the other basis. For this reason this basis is called *canonical basis* or *standard basis* of $\mathbb{R}^n$.

3. Dimension of a vector space

In this last section, we will show four important results about vector spaces. In few words:

- Every basis consists of the same number of vectors (*theorem of dimension*);
- it is possible to complete any set of linearly independent vectors, in order to obtain a basis for the space (*theorem of completition*);
- conversely, given any set of generators, it is possible to extract a basis for the space (*theorem of reduction*);
- the dimensions of two vector subspaces are related, via a precise formula, to the dimensions of the sum and the intersection spaces (*Grassmann's dimension formula*).

In other words, the number of vectors in a basis is a property of the space itself. Moreover, a basis is a maximal independent set (it cannot be made larger without losing the linear independence of the vectors) and a minimal spanning set (it cannot be made smaller and still span the space).

Theorem 1 (Theorem of dimension). *Let V be a vector space and $\{v_1, \ldots, v_n\}$ a basis. Every other basis of V consists of n vectors.*

In order to show this theorem, we need a lemma.

Lemma 1. *Let $\{v_1, \ldots, v_n\}$ be a set of generators of V. If $w_1, \ldots, w_m \in V$, with $m > n$, then they are linearly dependent.*

Proof. [Lemma 1] Consider the first n vectors $w_1, \ldots, w_n$. There are two possibilities: either these vectors are linearly dependent (and in this case, also the bigger set is so and the claim is proved) or they are independent. We consider the second alternative. To show the lemma, it suffices to show that:

(1) $$V = \langle w_1, \ldots, w_n \rangle.$$

In fact, if this is true, the vector w_{n+1} is a linear combination of the first n and the vectors $\{w_1, \ldots, w_n, w_{n+1}\}$ are linearly dependent (and the bigger set too).
Let us show (1).

Consider the vector w_1. Obviously $w_1 \neq 0$ (otherwise $0, w_2, \ldots, w_n$ would be linearly dependent). Since by hypothesis $V = \langle v_1, \ldots, v_n \rangle$, then:

$$w_1 = a_1 \cdot v_1 + \ldots + a_n \cdot v_n \,.$$

The coefficients $a_1, \ldots, a_n$ are not all zeros (since $w_1 \neq 0$), and without any loss of generality we can assume that $a_1 \neq 0$. Hence,

$$v_1 = \frac{1}{a_1} \cdot w_1 - \frac{a_2}{a_1} \cdot v_2 - \ldots - \frac{a_n}{a_1} \cdot v_n$$

and consequently $v_1 \in \langle w_1, v_2, \ldots, v_n \rangle$. Since $v_1, v_2, \ldots, v_n \in \langle w_1, v_2, \ldots, v_n \rangle$, then:

$$\langle v_1, v_2, \ldots, v_n \rangle \subseteq \langle w_1, v_2, \ldots, v_n \rangle \,;$$

but $\langle v_1, v_2, \ldots, v_n \rangle = V$, therefore:

$$\langle w_1, v_2, \ldots, v_n \rangle = V \,.$$

If $n = 1$, then the claim (1) is proved. Suppose $n \geq 2$ and consider the vector w_2. Obviously, $w_2 \in \langle w_1, v_2, \ldots, v_n \rangle = V$ and consequently

$$w_2 = b_1 \cdot w_1 + b_2 \cdot v_2 + \ldots + b_n \cdot v_n \,.$$

The coefficients $b_2, \ldots, b_n$ are not all zeros (otherwise $w_2 = b_2 \cdot w_1$ and the vectors $w_1, \ldots, w_n$ would be linearly dependent). Without any loss of generality, we can assume that $b_2 \neq 0$. Hence,

$$v_2 = \frac{1}{b_2} \cdot w_2 - \frac{b_1}{b_2} \cdot w_1 - \frac{b_3}{b_2} \cdot v_3 - \ldots - \frac{b_n}{b_2} \cdot v_n$$

and consequently $v_2 \in \langle w_2, w_1, v_3, \ldots, v_n \rangle$. Since $v_2, w_1, v_3, \ldots, v_n \in \langle w_2, w_1, v_3, \ldots, v_n \rangle$, then:

$$V = \langle v_2, w_1, v_3, \ldots, v_n \rangle \subseteq \langle w_2, w_1, v_3, \ldots, v_n \rangle \,;$$

therefore,

$$\langle w_2, w_1, v_3, \ldots, v_n \rangle = V \,.$$

Iterating this procedure a finite number of times, one shows the claim and completes the proof.   □

We can now show the theorem.

Proof. [Theorem of dimension] Suppose that $\{w_1, \ldots, w_m\}$ is another basis of V. We need to show that $n = m$.

Observe that $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$ are spanning sets; from the previous lemma it follows that if $n > m$, the vectors $v_1, \ldots, v_n$ would be linearly dependent (it is a contradiction). On the other hand, if $m > n$ the vectors $w_1, \ldots, w_m$ would be dependent (it is also a contradiction). Hence, $n = m$.   □

Definition. Let V be a vector space. The number of vectors in a basis (and consequently in any other basis) is called *dimension* of V and will be denoted $\dim(V)$.

Example.   • $\dim(\mathbb{R}^n) = n$. In fact the standard basis $\{e_1, \ldots, e_n\}$ is a basis (see previous section).
          • the trivial subspace $Z = \langle 0 \rangle$, has no basis. In fact, 0 is its only vector and it is linearly dependent. We set $\dim(Z) = 0$.
          • If $V = \langle v \rangle$, with $v \neq 0$, then $\dim(V) = 1$ (in fact $\{v\}$ is a basis).

Remark. There are some vector spaces for which it is not possible to find a finite basis (for instance, the spaces of functions, of polynomials or sequences). In this case we set $\dim(V) = +\infty$.

For the sake of the results hereafter, we will consider just finite dimensional vector spaces.

The following corollary to lemma 1 comes in handy for proving that a set of vector is a basis.

Corollary 1. *Let V be a vector space, with $\dim(V) = n$. One has:*

    i) *n linearly independent vectors form a basis;*

    ii) *a spanning set of V, consisting of n vectors, is a basis.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a basis of V.

    i) Let $w_1, \ldots, w_n$ be n linearly independent vectors. We want to show that they are also a spanning set, *i.e.*, $\langle w_1, \ldots, w_n \rangle = V$.

For any vector $v \in V$, consider $v, w_1, \ldots, w_n$. They are linearly dependent (from lemma 1), therefore there exists a non-trivial linear combination of 0:

$$a \cdot v + b_1 \cdot w_1 + \ldots + b_n \cdot w_n = 0\,.$$

We have: $a \neq 0$ (otherwise some among $b_1, \ldots, b_n$ would be different from 0 and the vectors $w_1, \ldots, w_n$ would be dependent). Therefore:

$$v = -\frac{b_1}{a} \cdot w_1 - \ldots - \frac{b_n}{a} \cdot w_n\,,$$

and $v \in \langle w_1, \ldots, w_n \rangle$. This shows that $\langle w_1, \ldots, w_n \rangle = V$.

    ii) Suppose now that $\langle w_1, \ldots, w_n \rangle = V$. We need to show that the vectors $w_1, \ldots, w_n$ are linearly independent. By contradiction, if this were not true, at least one of them (and we can assume w_1) would be linear combination of the others:

$$w_1 = c_2 \cdot w_2 + \ldots + c_n \cdot w_n\,.$$

But, $\langle w_1, \ldots, w_n \rangle = \langle w_2, \ldots, w_n \rangle$ and $\{w_2, \ldots, w_n\}$ would be a spanning set of V. If this were the case, lemma 1 would imply that $\{v_1, \ldots, v_n\}$ are linearly dependent: contradiction!

$\square$

Corollary 2. *Let V be a vector space, with $\dim(V) = n$, and let W, W_1, W_2 be three vector subspaces. One has:*

    i) *$\dim(W) \leq \dim(V)$. Moreover, $\dim(W)$ is the maximum number of linearly independent vectors in W.*

    ii) *If $W_1 \subseteq W_2$, then $\dim(W_1) \leq \dim(W_2)$. Moreover,*

$$\dim(W_1) = \dim(W_2) \qquad \Longleftrightarrow \qquad W_1 = W_2\,.$$

Proof.   i) Observe that linearly independent vectors in W are linearly independent also in V. Let us denote by m the maximum number of independent vectors in W. Since $\dim(V) = n$, $n+1$ vectors in V are always dependent; then $m \leq n$.

Let us fix m linearly independent vectors in W: $w_1, \ldots, w_m$. We want to show that $\langle w_1, \ldots, w_m \rangle = W$ (this implies that $\dim(W) = m$ and shows i)). Choose arbitrarily $w \in W$; the vectors $w, w_1, \ldots, w_m$ are linearly dependent (because they are $m+1$), therefore there exist $a, b_1, \ldots, b_m \in \mathbb{R}$ (not all zeros) such that:

$$a \cdot w + b_1 \cdot w_1 + \ldots + b_m \cdot w_m = 0\,.$$

We have that $a \neq 0$ (otherwise $w_1, \ldots, w_m$ would be linearly dependent). It follows that $w \in \langle w_1, \ldots w_m \rangle$. Hence, $\langle w_1, \ldots, w_m \rangle = W$.

    ii) Since W_1 is a subspace of W_2, from i) it follows that $\dim(W_1) \leq \dim(W_2)$. It is obvious that if $W_1 = W_2$, then $\dim(W_1) = \dim(W_2)$. Conversely, assume that

$$W_1 \subset W_2 \qquad \text{and} \qquad \dim(W_1) = \dim(W_2)\,.$$

Fix $t = \dim(W_1)$ and choose a basis $\{w_1, \ldots, w_t\}$ of W_1. The vectors $w_1, \ldots, w_t$ are linearly independent in W_1 and therefore in W_2. Because of corollary 1, $\{w_1, \ldots, w_t\}$ is a basis of W_2. Therefore,

$$W_2 = \langle w_1, \ldots, w_t \rangle = W_1 \,.$$

$\square$

Remark. Let $\dim(V) = n$. From the previous corollary, it follows that V has a unique subspace of dimension n, namely V itself. Moreover, any non-zero vector space has dimension ≥ 1. Therefore, V has also a unique subspace of dimension zero (*i.e.*, $Z = \langle 0 \rangle$).

Theorem 2 (Theorem of completition). *Let V be a vector space of dimension n and $w_1, \ldots, w_t \in V$ t linearly independent vectors, with $t < n$. Then, there exist $n - t$ vectors $w_{t+1}, \ldots, w_n \in V$ such that $\{w_1, \ldots, w_n\}$ is a basis of V.*

Roughly speaking, any linearly independent set in V can be extended to a basis, by adding more vectors if necessary.

Proof. $\{w_1, \ldots, w_t\}$ is not a spanning set of V (otherwise it would be a basis of V and $t = n$). Therefore $V \supset \langle w_1, \ldots, w_t \rangle$.
We can choose a vector $w_{t+1} \in V \setminus \langle w_1, \ldots, w_t \rangle$. Let us verify that the vectors $w_1, \ldots, w_t, w_{t+1}$ are linearly independent. Suppose, in fact, that

$$c_1 \cdot w_1 + \ldots + c_t \cdot w_t + c_{t+1} \cdot w_{t+1} = 0 \,.$$

If $c_{t+1} \neq 0$, then $w_{t+1} \in \langle w_1, \ldots, w_t \rangle$, contradicting our assumption. Therefore, $c_{t+1} = 0$; this implies that $c_1 \cdot w_1 + \ldots + c_t \cdot w_t = 0$, but since these vectors are linearly independent, it follows that also $c_1 = \ldots = c_t = 0$. This shows the linear independence.
If $n = t + 1$, the theorem is proved; otherwise, the proof proceeds by iteration of the same argument. $\square$

Now, we want to show that any spanning set in V can be reduced to a basis, by discarding vectors if necessary.

Theorem 3 (Theorem of reduction). *Let V be a n-dimensional vector space and $\{v_1, \ldots, v_m\}$ a spanning set of V. There exist n vectors*

$$v_{i_1}, \ldots, v_{i_n} \in \{v_1, \ldots, v_m\}$$

that form a basis of V.

Proof. First observe that $m \geq n$ (because of lemma 1). Moreover, we can assume that $v_1, \ldots, v_m$ are all different from zero (otherwise I can discard the zero ones and the remaining ones still form a set of generators).
Consider the non-zero vector v_1 and settle $v_{i_1} := v_1$. Let us examine $v_2, \ldots, v_m$ (in this prescribed order).
Denote by v_{i_2}, the first among these vectors such that:

$$\langle v_1 \rangle \subset \langle v_1, v_2, \ldots, v_{i_2} \rangle \,;$$

in other words:

$$\langle v_1 \rangle = \langle v_1, v_2 \rangle = \ldots = \langle v_1, v_2, \ldots, v_{i_2-1} \rangle \subset \langle v_1, v_2, \ldots, v_{i_2} \rangle = \langle v_1, v_{i_2} \rangle \,.$$

Observe that a similar vector might not exist; in such a case $\langle v_1 \rangle = V$ and $\{v_1\}$ is the basis for V we were looking for. While, if such a vector exists, we will keep on iterating the same argument and denote by v_{i_3} the first vector (if it exists) among $v_{i_2+1}, \ldots, v_m$ such that

$$\langle v_1, v_{i_2} \rangle \subset \langle v_1, v_{i_2}, \ldots, v_{i_3} \rangle \,.$$

Iterating the same idea, as long as the algorithm works (it will die in a finite number of steps), we end with a set of vectors:

$$v_1 = v_{i_1}, v_{i_2}, \ldots, v_{i_s}$$

with $1 = i_1 < i_2 < \ldots < i_s \le t$ and

$$\langle v_1 \rangle \subset \langle v_1, v_{i_2} \rangle \subset \ldots \subset \langle v_1, v_{i_2}, \ldots, v_{i_s} \rangle = V \,.$$

Obviously $\{v_{i_1}, v_{i_2}, \ldots, v_{i_s}\}$ is a set of generators. We need to show that they are linearly independent. Suppose that:

$$a_1 \cdot v_{i_1} + \ldots + a_s \cdot v_{i_s} = 0 \,.$$

We want to show that $a_1 = \ldots = a_s = 0$. Suppose by contradiction, that they are not all zeros, and denote by a_r, the last one different from zero (it is evident that at least another coefficient a_k must be different from zero and therefore $2 \le r \le s$). One has

$$a_r \cdot v_r = -\sum_{k=1}^{r-1} a_k \cdot v_{i_k}$$

and consequently $v_{i_r} \in \langle v_{i_1}, v_{i_2}, \ldots, v_{i_{r-1}} \rangle$; therefore:

$$\langle v_{i_1}, v_{i_2}, \ldots, v_{i_{r-1}} \rangle = \langle v_{i_1}, v_{i_2}, \ldots, v_{i_r} \rangle$$

contradicting the strict inclusions above. $\qquad\square$

Corollary 3. *Let* $\dim(V) = n$ *and let* W *be a vector subspace of* V, *spanned by* $\{w_1, \ldots, w_t\}$. *The dimension of* W *is the maximum number of linearly independent vectors among* $\{w_1, \ldots, w_t\}$.

Proof. It suffices to apply theorem 3 to W and keep on mind corollary 2(ii).

$\qquad\square$

Definition. Let $\dim(V) = n$ and let $w_1, \ldots, w_t \in V$. We define the *rank* of $w_1, \ldots, w_t$ [it will be denoted $\mathrm{rank}(w_1, \ldots, w_t)$] as the dimension of the vector subspace spanned by these vectors; *i.e.*,

$$\mathrm{rank}(w_1, \ldots, w_t) = \dim(\langle w_1, \ldots, w_t \rangle).$$

Remark. For what said above, this is the maximum number of linearly independent vectors among $w_1, \ldots, w_t$. Obviously,

$$\mathrm{rank}(w_1, \ldots, w_t) \le \min\{t, \dim(V)\} \,.$$

Finally, we want to show how the dimensions of two vector subspaces are related to the dimensions of the sum and the intersection spaces; this is known as *Grassmann's dimension formula.*

Theorem 4 (Grassmann's dimension formula). *Let* V *be a vector space and* U_1, U_2 *two vector subspaces of* V. *We have:*

$$\dim(U_1) + \dim(U_2) = \dim(U_1 + U_2) + \dim(U_1 \cap U_2) \,.$$

Proof. Denote $n_1 = \dim(U_1)$ and $n_2 = \dim(U_2)$. Let us observe that $U_1 \cap U_2$ has finite dimension $\le \min\{n_1, n_2\}$. Denote $i = \dim(U_1 \cap U_2)$ and choose one basis $\{z_1, \ldots, z_i\}$. Because of theorem 2, we can complete this basis to a basis of U_1 and one of U_2:

$$\{z_1, \ldots, z_i, u_1, \ldots, u_t\} \qquad \text{basis of } U_1 \qquad (t + i = n_1) \,,$$
$$\{z_1, \ldots, z_i, w_1, \ldots, w_s\} \qquad \text{basis of } U_2 \qquad (s + i = n_2) \,.$$

Let now consider all vectors together (they are $i + t + s = n_1 + s = n_1 + n_2 - i$):

$$z_1, \ldots, z_i, u_1, \ldots, u_t, w_1, \ldots, w_s \,.$$

If we manage to show that these vectors are a basis for $U_1 + U_2$, we are done, since:

$$\dim(U_1 + U_2) = n_1 + n_2 - i = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

We need to show:

i) They are a spanning set of $U_1 + U_2$.

Let $v_1 + v_2$ a vector in $U_1 + U_2$. We have (for suitably chosen coefficients):

$$\begin{aligned}
v_1 &= a_1 \cdot z_1 + \ldots + a_i \cdot z_i + b_1 \cdot u_1 + \ldots b_t \cdot u_t \\
v_2 &= a_1' \cdot z_1 + \ldots + a_i' \cdot z_i + c_1 \cdot w_1 + \ldots c_s \cdot w_s.
\end{aligned}$$

Hence:

$$\begin{aligned}
v_1 + v_2 &= (a_1 + a_1') \cdot z_1 + \ldots + (a_i + a_i') \cdot z_i + b_1 \cdot u_1 + \ldots b_t \cdot u_t + \\
&\quad + c_1 \cdot w_1 + \ldots c_s \cdot w_s,
\end{aligned}$$

and consequently

$$v_1 + v_2 \in \langle z_1, \ldots, z_i, u_1, \ldots, u_t, w_1, \ldots, w_s \rangle.$$

ii) They are a linearly independent.

$$\begin{aligned}
&a_1 \cdot z_1 + \ldots + a_i \cdot z_i + b_1 \cdot u_1 + \ldots b_t \cdot u_t + \\
&\quad + c_1 \cdot w_1 + \ldots c_s \cdot w_s = 0.
\end{aligned}$$

For simplyfing the notation, let us denote:

$$\begin{aligned}
z &= a_1 \cdot z_1 + \ldots + a_i \cdot z_i, \\
u &= b_1 \cdot u_1 + \ldots b_t \cdot u_t, \\
w &= c_1 \cdot w_1 + \ldots c_s \cdot w_s.
\end{aligned}$$

Therefore, $z + u + w = 0$ and

$$w = -(z + u) \in U_1 \cap U_2$$

(in fact $w \in U_2$ and $(z + u) \in U_1$). Hence, w can be expressed with respect to the basis $\{z_1, \ldots, z_i\}$:

$$w = -(z + u) = d_1 \cdot z_1 + \ldots + d_i \cdot z_i,$$

that implies

$$z + u + d_1 \cdot z_1 + \ldots + d_i \cdot z_i = 0,$$

or, equivalently,

$$(a_1 + d_1) \cdot z_1 + \ldots + (a_i + d_i) \cdot z_i + b_1 \cdot u_1 + \ldots b_t \cdot u_t = 0.$$

Since $z_1, \ldots, z_i, u_1, \ldots, u_t$ are linearly independent, the above coefficients must be zero. In particular, $b_1 = \ldots = b_t = 0$. Substituting these in our initial linear combination, we get:

$$a_1 \cdot z_1 + \ldots + a_i \cdot z_i + c_1 \cdot w_1 + \ldots c_s \cdot w_s = 0.$$

Using the linear independence of $z_1, \ldots, z_i, w_1, \ldots, w_s$, we conclude that also

$$a_1 = \ldots = a_i = c_1 = \ldots = c_s = 0.$$

Therefore, all coefficients in the above linear combination are zero, and this means that the vectors are linearly independent.

$\square$

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY
E-mail address: `asorrent@math.princeton.edu`