Our basic example of a vector space is $\mathbb{R}^n$. We have distinguished a sequence of vectors $e_1, e_2, \ldots, e_n$ in $\mathbb{R}^n$ and we have seen that every other vector in $\mathbb{R}^n$ can be expressed as $a_1e_1 + \ldots + a_ne_n$ for some (in fact unique) choice of numbers $a_1, \ldots, a_n$. There is no fundamental reason why the vectors e_i should be distinguished (they just appear naturally to us because of the way we write elements of $\mathbb{R}^n$). In fact, there are situations when a different choice is more convenient. It is our next goal to make this more precise and extend this idea to arbitrary vector spaces. The following concept will be fundamental to achieve this goal.

Definition. Let $v_1, v_2, \ldots, v_n$ be a sequence of vectors from a vector space V. We say that a vector $v \in V$ is a **linear combination** of the vectors $v_1, \ldots, v_n$ if $v = a_1v_1 + a_2v_2 + \ldots + a_nv_n$ for some choice of the numbers $a_1, \ldots, a_n$. The set of all vectors which are linear combinations of $v_1, \ldots, v_n$ is denoted by $\mathrm{span}\{v_1, \ldots, v_n\}$.

There is a very useful way of thinking about linear combinations and span via the concept of linear transformation.

Let V be a vector space and let $v_1, v_2, \ldots, v_n$ be a sequence of vectors from V. The function $L : \mathbb{R}^n \longrightarrow V$ defined by

$$L(a_1, \ldots, a_n) = a_1v_1 + a_2v_2 + \ldots + a_nv_n$$

is a linear transformation. It is the unique linear transformation which sends e_1 to v_1, e_2 to v_2, ..., e_n to v_n. The image of L is equal to $\mathrm{span}\{v_1, \ldots, v_n\}$.

Indeed, note that

$$L((a_1, \ldots, a_n) + (b_1, \ldots, b_n)) = L(a_1 + b_1, \ldots, a_n + b_n) = (a_1 + b_1)v_1 + \ldots + (a_n + b_n)v_n =$$

$$= a_1v_1 + b_1v_1 + a_2v_2 + b_2v_2 + \ldots + a_nv_n + b_nv_n = a_1v_1 + a_2v_2 + \ldots + a_nv_n + b_1v_1 + b_2v_2 + \ldots + b_nv_n =$$

$$= L(a_1, \ldots, a_n) + L(b_1, \ldots, b_n)$$

and

$$L(c(a_1, \ldots, a_n)) = L(ca_1, \ldots, ca_n) = (ca_1)v_1 + (ca_2)v_2 + \ldots + (ca_n)v_n = c(a_1v_1 + a_2v_2 + \ldots + a_nv_n) = cL(a_1, \ldots, a_n).$$

These prove that L is indeed a linear transformation. It is clear from the definition that $L(e_i) = v_i$ for $i = 1, 2, \ldots, n$. Conversely, suppose that $T : \mathbb{R}^n \longrightarrow V$ is a linear transformation such that $T(e_i) = v_i$ for $i = 1, 2, \ldots, n$. Then

$$T(a_1, \ldots, a_n) = T(a_1e_1 + \ldots + a_ne_n) = a_1T(e_1) + \ldots + a_nT(e_n) = a_1v_1 + a_2v_2 + \ldots + a_nv_n = L(a_1, \ldots, a_n).$$

This shows that $T = L$. Finally, v is in the image of L if and only if

$$v = L(a_1, \ldots, a_n) = a_1v_1 + a_2v_2 + \ldots + a_nv_n$$

for some $(a_1, \ldots, a_n) \in \mathbb{R}^n$, which is the same as saying that v is in $\mathrm{span}\{v_1, \ldots, v_n\}$. Thus the image of L coincides with $\mathrm{span}\{v_1, \ldots, v_n\}$.

Since the image of any linear transformation is a subspace, we get

Corollary. $\mathrm{span}\{v_1, \ldots, v_n\}$ is a subspace of V.

Exercise. Give a direct argument that $\mathrm{span}\{v_1, \ldots, v_n\}$ is a subspace of V.

We say that the vectors $v_1, \ldots, v_n$ **span** the vector space V if $V = \mathrm{span}\{v_1, \ldots, v_n\}$.

The following rather obvious observation is often useful: if U is a subspace of V and $v_i \in U$ for $i = 1, \ldots, n$ then $\operatorname{span}\{v_1, \ldots, v_n\}$ is contained in U. This is an immediate consequence of the fact that U is closed under addition and scalar multiplication. This observation leads us to the following simple but important result.

Theorem (Going Down). Let $v_1, v_2, \ldots, v_n$ be a sequence of vectors from a vector space V. If v_i is a linear combination of the other vectors then

$$\operatorname{span}\{v_1, \ldots, v_n\} = \operatorname{span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$$

i.e. we can remove v_i from the sequence and the remaining vectors span the same subspace.

Indeed, since every vector in the spanning sequence on the right belongs to the subspace $\operatorname{span}\{v_1, \ldots, v_n\}$, we get the inclusion

$$\operatorname{span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\} \subseteq \operatorname{span}\{v_1, \ldots, v_n\}.$$

On the other hand, since v_i is a linear combination of the other vectors, it belongs to the subspace $\operatorname{span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ and so do the vectors $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$. Thus we get the opposite inclusion

$$\operatorname{span}\{v_1, \ldots, v_n\} \subseteq \operatorname{span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$$

proving that both spans coincide.

Starting with any sequence of vectors and using the Going Down Theorem we can keep removing vectors until we obtain a sequence which spans the same subspace as the original vectors and no vector in this new sequence is a linear combination of the the other vectors. This prompts the following definition.

Definition. Vectors $v_1, v_2, \ldots, v_n$ are called **linearly independent** if they are all non-zero and none of the vectors is a linear combination of the other vectors.

We can now state our observation above as follows: every sequence of vectors, not all of which are zero, contains a subsequence which is linearly independent and spans the same subspace as the original sequence. Moreover, linearly independent vectors play analogous role for the subspace they span as the vectors $e_1, \ldots, e_n$ do for $\mathbb{R}^n$ (it is very easy to see that $e_1, \ldots, e_n$ are linearly independent). To make this sentence more clear, we note the following characterization of linearly independent vectors.

Proposition. Vectors $v_1, v_2, \ldots, v_n$ are linearly independent if and only if the only numbers $a_1, \ldots, a_n$ such that $a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0$ are $a_1 = a_2 = \ldots = a_n = 0$.

Indeed, if $v_1, v_2, \ldots, v_n$ are not linearly independent, then either $v_i = 0$ for some i and then $a_i = 1$ and $a_j = 0$ for $j \neq i$ are not all zero and satisfy $a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0$, or $v_i = b_1 v_1 + \ldots + b_{i-1} v_{i-1} + b_{i+1} v_{i+1} + \ldots + b_n v_n$ for some choice of the numbers b_i. Taking $a_i = -1$ and $a_j = b_j$ for $i \neq j$, we see that $a_1, \ldots, a_n$ are not all zero and $a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0$. This proves that if $v_1, v_2, \ldots, v_n$ are not linearly independent then there exist $a_1, \ldots, a_n$ which are not all zero and such that $a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0$. Conversely, suppose that $a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0$ and $a_i \neq 0$ for some i. If $n = 1$, this means that $v_1 = 0$. If $n > 1$, we can write

$$v_i = \frac{a_1}{-a_i} v_1 + \ldots + \frac{a_{i-1}}{-a_i} v_{i-1} + \frac{a_{i+1}}{-a_i} v_{i+1} + \ldots + \frac{a_n}{-a_i} v_n$$

so v_i is a linear combination of the other vectors. In both cases, $v_1, v_2, \ldots, v_n$ are not linearly independent.

It is convenient to make the following definition.

Definition. Vectors $v_1, v_2, \ldots, v_n$ are called **linearly dependent** if they are not linearly independent. In other words, $v_1, v_2, \ldots, v_n$ are linearly dependent if and only if there exist numbers $a_1, \ldots, a_n$, not all zero, such that $a_1v_1 + a_2v_2 + \ldots + a_nv_n = 0$.

It is clear now that if $v_1, \ldots, v_n$ are linearly independent and $a_1v_1 + \ldots a_nv_n = b_1v_1 + \ldots + b_nv_n$ then $a_1 = b_1$, ..., $a_n = b_n$. Thus every vector in $\mathrm{span}\{v_1, \ldots, v_n\}$ can be expressed in a **unique** way as a linear combination $a_1v_1 + a_2v_2 + \ldots + a_nv_n$. This is why the vectors $v_1, \ldots, v_n$ play analogous role for $\mathrm{span}\{v_1, \ldots, v_n\}$ as the vectors $e_1, \ldots, e_n$ do for $\mathbb{R}^n$

We restate the last proposition in a slightly different form.

Let V be a vector space and let $v_1, v_2, \ldots, v_n$ be a sequence of vectors from V. Consider the linear transformation $L : \mathbb{R}^n \longrightarrow V$ defined by

$$L(a_1, \ldots, a_n) = a_1v_1 + a_2v_2 + \ldots + a_nv_n.$$

1. L is onto if and only if $V = \mathrm{span}\{v_1, \ldots, v_n\}$;
2. L is one-to-one if and only if $v_1, v_2, \ldots, v_n$ are linearly independent;
3. L is a bijection if and only if $v_1, v_2, \ldots, v_n$ are linearly independent and span V.

This result prompts the following definitions.

Definition. We say that $v_1, v_2, \ldots, v_n$ is a **basis** of a vector space V if $v_1, v_2, \ldots, v_n$ are linearly independent and $V = \mathrm{span}\{v_1, \ldots, v_n\}$.

Definition. We say that a vector space V is **finite dimensional** if there is a sequence $v_1, v_2, \ldots, v_n$ such that $V = \mathrm{span}\{v_1, \ldots, v_n\}$.

Restating some of our earlier observations we get the following theorem.

Theorem. Every non-trivial finite dimensional vector space V has a basis. Moreover, every sequence which spans V contains a subsequence which is a basis.

Remark. We say that a vector space V is **trivial** if $V = \{0\}$, i.e. V consists of only the zero vector. It is convenient to define the empty sequence to be a basis of a trivial vector space.

The name "finite dimensional" suggests that there is a notion of dimension. Indeed such notion exists and it captures our intuitive idea that $\mathbb{R}^n$ should have dimension n. It should be clear that $e_1, \ldots, e_n$ is a basis of $\mathbb{R}^n$. This suggests that dimension should be related to the size of a basis. For this to make sense it would be ideal to know that any two bases of a vector space have the same number of elements. As we will see, this is indeed true.

Theorem. Any two bases of a finite dimensional vector space V have the same number of elements. This number is called the **dimension** of V and it is denoted by $\dim(V)$.

To justify this result, note that if $v_1, \ldots, v_n$ is a basis of V then there is a linear transformation $L : \mathbb{R}^n \longrightarrow V$ which is a bijection. Similarly, if $w_1, \ldots, w_k$ is a basis of V then there is a linear transformation $T : \mathbb{R}^k \longrightarrow V$ which is a bijection. Now $T^{-1}L : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ is also a linear transformation which is a bijection. However, we know that this is only possible when $n = k$.

Remark. We define the dimension of the trivial vector space to be 0. It is consistent with the convention in our previous remark that the empty sequence is a basis of a trivial vector space, as empty sequence has 0 elements.

Suppose now that V is a finite dimensional vector space and U is a subspace of V. Our intuition suggests that U should be also finite dimensional and $\dim(U) \leq \dim(V)$. We are going to show that this is indeed true. The key observation is the following result.

Theorem (Going Up). Let $v_1, v_2, \ldots, v_n$ be a linearly independent sequence of vectors from a vector space V. If $v \in V$ is a vector not in $\operatorname{span}\{v_1, \ldots, v_n\}$ then $v_1, v_2, \ldots, v_n, v$ are linearly independent.

Indeed, suppose that $v_1, v_2, \ldots, v_n, v$ are not linearly independent. Then there exists numbers $a_1, \ldots, a_n, a$, not all zero, such that $a_1v_1 + \ldots a_nv_n + av = 0$. If $a = 0$, then $v_1, v_2, \ldots, v_n$ are not linearly independent contrary to our assumption. Thus $a \neq 0$. However this means that

$$v = \frac{a_1}{-a}v_1 + \ldots + \frac{a_n}{-a}v_n$$

i.e. v is a linear combination of $v_1, \ldots, v_n$, again contrary to our assumptions.

We also need the following observation.

Theorem. Let V be a vector space of dimension n. If $v_1, \ldots, v_k$ is a linearly independent sequence in V then $k \leq n$. Moreover, this sequence can be extended to a basis of V.

Indeed, since V has dimension n there is a linear transformation $L : \mathbb{R}^n \longrightarrow V$ which is a bijection. Since $v_1, \ldots, v_k$ is a linearly independent sequence in V, we have an injective linear transformation $T : \mathbb{R}^k \longrightarrow V$. It follows that $L^{-1} \circ T : \mathbb{R}^k \longrightarrow \mathbb{R}^n$ is an injective linear transformation. This can only happen if $k \leq n$.

Now, using the Going Up Theorem and starting with $v_1, \ldots, v_k$, we can produce longer and longer linearly independent sequences as long as they do not span V. However we already proved that there is no linearly independent sequence in V longer than n. Thus, we must arrive at a a basis of V.

We are now ready to justify the result we stated earlier.

Theorem. Let V be a vector space of dimension n. If U is a proper subspace of V (i.e. U is not equal to V) then U has finite dimension which is smaller than n.

Indeed, the result is obvious when U is the trivial subspace. Assume that U is not trivial. We proved that any linearly independent sequence of vectors from V has length at most n. Thus we may choose longest linearly independent sequence which consists of elements from U, say $u_1, \dots, u_k$. If this sequence does not span U, we could choose $u \in U$ which is not in the span$\{u_1, \dots, u_k\}$ and we would get a longer linearly independent sequence $u_1, \dots, u_k, u$ by the Going Up Theorem, which contradicts our choice of k. Thus $U = \text{span}\{u_1, \dots, u_k\}$ and $\dim(U) = k$. Since $U \neq V$, we can chose $v \in V$ which is not in U and then, again by the Going Up Theorem, the sequence $u_1, \dots, u_k, v$ is linearly independent. It follows that $k + 1 \leq n$, so $k < n$.

Remarks. Not every vector space is finite dimensional. For example, consider the vector space P of all polynomial functions. If $p_1, \dots, p_k$ is a finite sequence of polynomials then all these polynomials have degree bounded above by some integer k and then any linear combination of these polynomials has also degree bounded above by k. Thus the span of $p_1, \dots, p_k$ does not contain polynomials of degree larger than k, hence it can not be equal to the whole P. This shows that P is not finite dimensional.

The concept of span and basis can be extended to vector spaces which are not finite dimensional. First, as it is done in many texts on linear algebra, we define $\text{span}(X)$ for any subset X of a vector space V. When X is finite, we can put the elements of X in a finite sequence $x_1, \dots, x_n$ and define $\text{span}(X) = \text{span}\{x_1, \dots, x_n\}$. Furthermore, we say that X is linearly independent if $x_1, \dots, x_n$ is linearly independent. It is easy to see that these notions do not depend on the way we list the elements. When X is infinite, we define $\text{span}(X)$ as the union of the subspaces $\text{span}(Y)$, where Y runs through all finite subsets of X. It is easy to see that $\text{span}(X)$ is a subspace consisting of all vectors which are linear combinations of some finite sequence of vectors from X. We say that X is linearly independent if every finite subset of X is linearly independent. We say that X is a basis of V if X is linearly independent and spans V. It can be proved that every vector space has a basis, but it is significantly harder than the finite dimensional case. In our course we will study only finite dimensional vector spaces.