

Linear Span and Bases

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling
(January 23, 2007)

Intuition probably tells you that the plane \mathbb{R}^2 is of dimension two and the space we live in \mathbb{R}^3 is of dimension three. You have probably also learned in physics that space-time has dimension four, and that string theories are models that can live in ten dimensions. In these lectures we will give a mathematical definition of what the dimension of a vector space is. For this we will first need the notion of linear spans, linear independence and the basis of a vector space.

1 Linear span

As before, let V denote a vector space over \mathbb{F} . Given vectors $v_1, v_2, \dots, v_m \in V$, a vector $v \in V$ is a **linear combination** of (v_1, \dots, v_m) if there exist scalars $a_1, \dots, a_m \in \mathbb{F}$ such that

$$v = a_1v_1 + a_2v_2 + \dots + a_mv_m.$$

Definition 1. The **linear span** or simply **span** of (v_1, \dots, v_m) is defined as

$$\text{span}(v_1, \dots, v_m) := \{a_1v_1 + \dots + a_mv_m \mid a_1, \dots, a_m \in \mathbb{F}\}.$$

Lemma 1. Let V be a vector space and $v_1, v_2, \dots, v_m \in V$

1. $v_j \in \text{span}(v_1, v_2, \dots, v_m)$.
2. $\text{span}(v_1, v_2, \dots, v_m)$ is a subspace of V .
3. If $U \subset V$ is a subspace such that $v_1, v_2, \dots, v_m \in U$, then $\text{span}(v_1, v_2, \dots, v_m) \subset U$.

Proof. 1 is obvious. For 2 note that $0 \in \text{span}(v_1, v_2, \dots, v_m)$ and that $\text{span}(v_1, v_2, \dots, v_m)$ is closed under addition and scalar multiplication. For 3 note that a subspace U of a vector space V is closed under addition and scalar multiplication. Hence if $v_1, \dots, v_m \in U$, then any linear combination $a_1v_1 + \dots + a_mv_m$ must also be in U . \square

Lemma 1 implies that $\text{span}(v_1, v_2, \dots, v_m)$ is the smallest subspace of V containing all v_1, v_2, \dots, v_m .

Definition 2. If $\text{span}(v_1, \dots, v_m) = V$, we say that (v_1, \dots, v_m) spans V . The vector space V is called **finite-dimensional**, if it is spanned by a finite list of vectors. A vector space V that is not finite-dimensional is called **infinite-dimensional**.

Example 1. The vectors $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, \dots, 0, 1)$ span \mathbb{F}^n . Hence \mathbb{F}^n is finite-dimensional.

Example 2. If $p(z) = a_m z^m + a_{m-1} z^{m-1} + \dots + a_1 z + a_0 \in \mathcal{P}(\mathbb{F})$ is a polynomial with coefficients in \mathbb{F} such that $a_m \neq 0$ we say that $p(z)$ has degree m . By convention the degree of $p(z) = 0$ is $-\infty$. The **degree** of $p(z)$ is denoted by $\deg p(z)$. Define

$$\mathcal{P}_m(\mathbb{F}) = \text{set of all polynomials in } \mathcal{P}(\mathbb{F}) \text{ of degree at most } m.$$

Then $\mathcal{P}_m(\mathbb{F}) \subset \mathcal{P}(\mathbb{F})$ is a subspace since it contains the zero polynomial and is closed under addition and scalar multiplication. In fact

$$\mathcal{P}_m(\mathbb{F}) = \text{span}(1, z, z^2, \dots, z^m).$$

Example 3. We showed that $\mathcal{P}(\mathbb{F})$ is a vector space. In fact, $\mathcal{P}(\mathbb{F})$ is infinite-dimensional. To see this, assume the contrary, namely that

$$\mathcal{P}(\mathbb{F}) = \text{span}(p_1(z), \dots, p_k(z))$$

for a finite set of k polynomials $p_1(z), \dots, p_k(z)$. Let $m = \max(\deg p_1(z), \dots, \deg p_k(z))$. Then $z^{m+1} \in \mathcal{P}(\mathbb{F})$, but $z^{m+1} \notin \text{span}(p_1(z), \dots, p_k(z))$.

2 Linear independence

We are now going to define the notion of linear independence of a list of vectors. This concept will be extremely important in the following, especially when we introduce bases and the dimension of a vector space.

Definition 3. A list of vectors (v_1, \dots, v_m) is called **linearly independent** if the only solution for $a_1, \dots, a_m \in \mathbb{F}$ to the equation

$$a_1 v_1 + \dots + a_m v_m = 0$$

is $a_1 = \dots = a_m = 0$. In other words the zero vector can only be trivially written as the linear combination of (v_1, \dots, v_m) .

Definition 4. A list of vectors (v_1, \dots, v_m) is called **linearly dependent** if it is not linearly independent. That is, there exist $a_1, \dots, a_m \in \mathbb{F}$ not all being zero such that

$$a_1v_1 + \dots + a_mv_m = 0.$$

Example 4. The vectors (e_1, \dots, e_m) of Example 1 are linearly independent. The only solution to

$$0 = a_1e_1 + \dots + a_me_m = (a_1, \dots, a_m)$$

is $a_1 = \dots = a_m = 0$.

Example 5. The vectors $(1, z, \dots, z^m)$ in the vector space $\mathcal{P}_m(\mathbb{F})$ are linearly independent. Requiring that

$$a_01 + a_1z + \dots + a_mz^m = 0$$

means that the polynomial on the left should be zero for all $z \in \mathbb{F}$. This is only possible for $a_0 = a_1 = \dots = a_m = 0$.

An important consequence of the notion of linear independence is the fact that any vector in the span of a given list of linearly independent vectors can be uniquely written as a linear combination.

Lemma 2. *The list of vectors (v_1, \dots, v_m) is linearly independent if and only if every $v \in \text{span}(v_1, \dots, v_m)$ can be uniquely written as a linear combination of (v_1, \dots, v_m) .*

Proof.

" \implies " Assume that (v_1, \dots, v_m) is a linearly independent list of vectors. Suppose there are two ways of writing $v \in \text{span}(v_1, \dots, v_m)$ as a linear combination of the v_i :

$$\begin{aligned} v &= a_1v_1 + \dots + a_mv_m, \\ v &= a'_1v_1 + \dots + a'_mv_m. \end{aligned}$$

Subtracting the two equations yields $0 = (a_1 - a'_1)v_1 + \dots + (a_m - a'_m)v_m$. Since (v_1, \dots, v_m) are linearly independent the only solution to this equation is $a_1 - a'_1 = 0, \dots, a_m - a'_m = 0$, or equivalently $a_1 = a'_1, \dots, a_m = a'_m$.

" \impliedby " Now assume that for every $v \in \text{span}(v_1, \dots, v_m)$ there are unique $a_1, \dots, a_m \in \mathbb{F}$ such that

$$v = a_1v_1 + \dots + a_mv_m.$$

This implies in particular that the only way the zero vector $v = 0$ can be written as a linear combination of v_1, \dots, v_m is with $a_1 = \dots = a_m = 0$. This shows that (v_1, \dots, v_m) are linearly independent. \square

It is clear that if (v_1, \dots, v_m) is a list of linearly independent vectors then the list (v_1, \dots, v_{m-1}) is also linearly independent.

For the next lemma we introduce the following notation. If we want to drop a vector v_j from a given list (v_1, \dots, v_m) of vectors, we indicate the dropped vector by a hat $(v_1, \dots, \hat{v}_j, \dots, v_m)$.

Lemma 3 (Linear Dependence Lemma). *If (v_1, \dots, v_m) is linearly dependent and $v_1 \neq 0$, there exists an index $j \in \{2, \dots, m\}$ such that:*

1. $v_j \in \text{span}(v_1, \dots, v_{j-1})$.
2. If v_j is removed from (v_1, \dots, v_m) then $\text{span}(v_1, \dots, \hat{v}_j, \dots, v_m) = \text{span}(v_1, \dots, v_m)$.

Proof. Since (v_1, \dots, v_m) is linearly dependent there exist $a_1, \dots, a_m \in \mathbb{F}$ not all zero such that $a_1 v_1 + \dots + a_m v_m = 0$. Since by assumption $v_1 \neq 0$, not all of a_2, \dots, a_m can be zero. Let $j \in \{2, \dots, m\}$ be largest such that $a_j \neq 0$. Then we have

$$v_j = -\frac{a_1}{a_j} v_1 - \dots - \frac{a_{j-1}}{a_j} v_{j-1}, \quad (1)$$

which implies part 1.

Let $v \in \text{span}(v_1, \dots, v_m)$. By definition this means that there exist scalars $b_1, \dots, b_m \in \mathbb{F}$ such that

$$v = b_1 v_1 + \dots + b_m v_m.$$

The vector v_j that we determined in part 1 can be replaced by (1), so that v is written as a linear combination of $(v_1, \dots, \hat{v}_j, \dots, v_m)$. Hence $\text{span}(v_1, \dots, \hat{v}_j, \dots, v_m) = \text{span}(v_1, \dots, v_m)$. \square

Example 6. Take the list $(v_1, v_2, v_3) = ((1, 1), (1, 2), (1, 0))$ of vectors in \mathbb{R}^2 . They span \mathbb{R}^2 . To see this, take any vector $v = (x, y) \in \mathbb{R}^2$. We want to show that v can be written as a linear combination of $(1, 1), (1, 2), (1, 0)$

$$v = a_1(1, 1) + a_2(1, 2) + a_3(1, 0)$$

or equivalently

$$(x, y) = (a_1 + a_2 + a_3, a_1 + 2a_2).$$

Taking $a_1 = y, a_2 = 0, a_3 = x - y$ is a solution for given $x, y \in \mathbb{R}$. Hence indeed $\mathbb{R}^2 = \text{span}((1, 1), (1, 2), (1, 0))$. Note that

$$2(1, 1) - (1, 2) - (1, 0) = (0, 0), \quad (2)$$

which shows that the list $((1, 1), (1, 2), (1, 0))$ is linearly dependent. The Linear Dependence Lemma 3 states that one of the vectors can be dropped from $((1, 1), (1, 2), (1, 0))$ and still

span \mathbb{R}^2 . Indeed by (2)

$$v_3 = (1, 0) = 2(1, 1) - (1, 2) = 2v_1 - v_2,$$

so that $\text{span}((1, 1), (1, 2), (1, 0)) = \text{span}((1, 1), (1, 2))$.

The next results shows that linearly independent lists of vectors that span a finite-dimensional vector space are the smallest possible spanning sets.

Theorem 4. *Let V be a finite-dimensional vector space. Suppose that (v_1, \dots, v_m) is a linearly independent list of vectors that spans V , and let (w_1, \dots, w_n) be any list that spans V . Then $m \leq n$.*

Proof. The proof uses an iterative procedure. We start with an arbitrary list $\mathcal{S}_0 = (w_1, \dots, w_n)$ that spans V . At the k -th step of the procedure we construct a new list \mathcal{S}_k by replacing a w_{j_k} by v_k such that \mathcal{S}_k still spans V . Repeating this for all v_k finally produces a new list \mathcal{S}_m of length n that contains all v_1, \dots, v_m . This proves that indeed $m \leq n$. Let us now discuss each step in the procedure in detail:

Step 1. Since (w_1, \dots, w_n) spans V , adding a new vector to the list makes the new list linearly dependent. Hence (v_1, w_1, \dots, w_n) is linearly dependent. By Lemma 3 there exists an index j_1 such that

$$w_{j_1} \in \text{span}(v_1, w_1, \dots, w_{j_1-1}).$$

Hence $\mathcal{S}_1 = (v_1, w_1, \dots, \hat{w}_{j_1}, \dots, w_n)$ spans V . In this step we added the vector v_1 and removed the vector w_{j_1} from \mathcal{S}_0 .

Step k . Suppose that we already added v_1, \dots, v_{k-1} to our spanning list and removed the vectors $w_{j_1}, \dots, w_{j_{k-1}}$ in return. Call this list \mathcal{S}_{k-1} which spans V . Add the vector v_k to \mathcal{S}_{k-1} . By the same arguments as before, adjoining the extra vector v_k to the spanning list \mathcal{S}_{k-1} yields a list of linearly dependent vectors. Hence by Lemma 3 there exists an index j_k such that \mathcal{S}_{k-1} with v_k added and w_{j_k} removed still spans V . The fact that (v_1, \dots, v_k) is linearly independent ensures that the vector removed is indeed among the w_j . Call the new list \mathcal{S}_k which spans V .

The final list \mathcal{S}_m is \mathcal{S}_0 with all v_1, \dots, v_m added and w_{j_1}, \dots, w_{j_m} removed. It has length n and still spans V . Hence necessarily $m \leq n$. \square

3 Bases

A basis of a finite-dimensional vector space is a spanning list that is also linearly independent. We will see that all bases of finite-dimensional vector spaces have the same length. This length will be the dimension of our vector space.

Definition 5. A **basis** of a finite-dimensional vector space V is a list of vectors (v_1, \dots, v_m) in V that is linearly independent and spans V .

If (v_1, \dots, v_m) forms a basis of V , then by Lemma 2 every vector $v \in V$ can be uniquely written as a linear combination of (v_1, \dots, v_m) .

Example 7. (e_1, \dots, e_n) is a basis of \mathbb{F}^n . There are of course other bases. For example $((1, 2), (1, 1))$ is a basis of \mathbb{F}^2 . The list $((1, 1))$ is linearly independent, but does not span \mathbb{F}^2 and hence is not a basis.

Example 8. $(1, z, z^2, \dots, z^m)$ is a basis of $\mathcal{P}_m(\mathbb{F})$.

Theorem 5 (Basis Reduction Theorem). *If $V = \text{span}(v_1, \dots, v_m)$, then some v_i can be removed to obtain a basis of V .*

Proof. Suppose $V = \text{span}(v_1, \dots, v_m)$. We start with the list $\mathcal{S} = (v_1, \dots, v_m)$ and iteratively run through all vectors v_k for $k = 1, 2, \dots, m$ to determine whether to keep or remove them from \mathcal{S} :

Step 1. If $v_1 = 0$, remove v_1 from \mathcal{S} . Otherwise leave \mathcal{S} unchanged.

Step k . If $v_k \in \text{span}(v_1, \dots, v_{k-1})$, remove v_k from \mathcal{S} . Otherwise leave \mathcal{S} unchanged.

The final list \mathcal{S} still spans V since at each step a vector was only discarded if it was already in the span of the previous vectors. The process also ensures that no vector is in the span of the previous vectors. Hence by the Linear Dependence Lemma 3 the final list \mathcal{S} is linearly independent. Hence \mathcal{S} is a basis of V . \square

Example 9. To see how Basis Reduction Theorem 5 works, consider the list of vectors

$$\mathcal{S} = ((1, -1, 0), (2, -2, 0), (-1, 0, 1), (0, -1, 1), (0, 1, 0)).$$

This list does not form a basis for \mathbb{R}^3 as it is not linearly independent. However, it is clear that $\mathbb{R}^3 = \text{span}(\mathcal{S})$ since any arbitrary vector $v = (x, y, z) \in \mathbb{R}^3$ can be written as the following linear combination over \mathcal{S} :

$$v = (x + z)(1, -1, 0) + 0(2, -2, 0) + (z)(-1, 0, 1) + 0(0, -1, 1) + (x + y + z)(0, 1, 0).$$

In fact, since the coefficients of $(2, -2, 0)$ and $(0, -1, 1)$ in this linear combination are both zero, it suggests that they add nothing to the span of the subset

$$\mathcal{B} = ((1, -1, 0), (-1, 0, 1), (0, 1, 0))$$

of \mathcal{S} . Moreover, one can show that \mathcal{B} is a basis for \mathbb{R}^3 , and it is exactly the basis produced by applying the process from the proof of Theorem 5 (as you should be able to verify).

Corollary 6. *Every finite-dimensional vector space has a basis.*

Proof. By definition, a finite-dimensional vector space has a spanning list. By the Basis Reduction Theorem 5 any spanning list can be reduced to a basis. \square

Theorem 7 (Basis Extension Theorem). *Every linearly independent list of vectors in a finite-dimensional vector space V can be extended to a basis of V .*

Proof. Suppose V is finite-dimensional and (v_1, \dots, v_m) is linearly independent. Since V is finite-dimensional, there exists a list (w_1, \dots, w_n) of vectors that spans V . We wish to adjoin some of the w_k to (v_1, \dots, v_m) to create a basis of V .

Step 1. If $w_1 \in \text{span}(v_1, \dots, v_m)$, let $\mathcal{S} = (v_1, \dots, v_m)$. Otherwise set $\mathcal{S} = (v_1, \dots, v_m, w_1)$.

Step k . If $w_k \in \text{span}(\mathcal{S})$, leave \mathcal{S} unchanged. Otherwise adjoin w_k to \mathcal{S} .

After each step the list \mathcal{S} is still linearly independent since we only adjoined w_k if w_k was not in the span of the previous vectors. After n steps $w_k \in \text{span}(\mathcal{S})$ for all $k = 1, 2, \dots, n$. Since (w_1, \dots, w_n) was a spanning list, \mathcal{S} spans V , so that \mathcal{S} is indeed a basis of V . \square

4 Dimension

We now come to the important definition of the dimension of finite-dimensional vector spaces. Intuitively we know that the plane \mathbb{R}^2 has dimension 2, \mathbb{R}^3 has dimension 3, or more generally \mathbb{R}^n has dimension n . This is precisely the length of the bases of these vector spaces, which prompts the following definition.

Definition 6. We call the length of any basis of V (which is well-defined by Theorem 8), the **dimension** of V , also denoted $\dim V$.

Note that Definition 6 only makes sense, if in fact all bases of a given finite-dimensional vector space have the same length. This is true by the next Theorem.

Theorem 8. *Let V be a finite-dimensional vector space. Then any two bases of V have the same length.*

Proof. Let (v_1, \dots, v_m) and (w_1, \dots, w_n) be two bases of V . Both span V . By Theorem 4, we have $m \leq n$ since (v_1, \dots, v_m) is linearly independent. By the same theorem we also have $n \leq m$ since (w_1, \dots, w_n) is linearly independent. Hence $n = m$ as asserted. \square

Example 10. $\dim \mathbb{F}^n = n$ and $\dim \mathcal{P}_m(\mathbb{F}) = m + 1$. Note that $\dim \mathbb{C}^n = n$ as a \mathbb{C} vector space, but $\dim \mathbb{C}^n = 2n$ as an \mathbb{R} vector space. This comes from the fact that we can view \mathbb{C} itself as an \mathbb{R} vector space of dimension 2 with basis $(1, i)$.

Theorem 9. *Let V be a finite-dimensional vector space with $\dim V = n$. Then:*

1. If $U \subset V$ is a subspace of V , then $\dim U \leq \dim V$.
2. If $V = \text{span}(v_1, \dots, v_n)$, then (v_1, \dots, v_n) is a basis of V .
3. If (v_1, \dots, v_n) is linearly independent in V , then (v_1, \dots, v_n) is a basis of V .

Point 1 implies in particular, that every subspace of a finite-dimensional vector space is finite-dimensional. Points 2 and 3 show that if the dimension of a vector space is known to be n , then to check that a list of n vectors is a basis it is enough to check whether it spans V (resp. is linearly independent).

Proof. To prove point 1, let (u_1, \dots, u_m) be a basis of U . This list is linearly independent both in U and V . By the Basis Extension Theorem 7 it can be extended to a basis of V , which is of length n since $\dim V = n$. This implies that $m \leq n$ as desired.

To prove point 2 suppose that (v_1, \dots, v_n) spans V . Then by the Basis Reduction Theorem 5 this list can be reduced to a basis. However, every basis of V has length n , hence no vector needs to be removed from (v_1, \dots, v_n) . Hence (v_1, \dots, v_n) is already a basis of V .

Point 3 is proved in a very similar fashion. Suppose (v_1, \dots, v_n) is linearly independent. By the Basis Extension Theorem 7 this list can be extended to a basis. However, every basis has length n , hence no vector needs to be added to (v_1, \dots, v_n) . Hence (v_1, \dots, v_n) is already a basis of V . \square

We conclude this chapter with some additional interesting results on bases and dimensions. The first one combines concepts of bases and direct sums of vector spaces.

Theorem 10. *Let $U \subset V$ be a subspace of a finite-dimensional vector space V . Then there exists a subspace $W \subset V$ such that $V = U \oplus W$.*

Proof. Let (u_1, \dots, u_m) be a basis of U . By point 1 of Theorem 9 we know that $m \leq \dim V$. Hence by the Basis Extension Theorem 7 (u_1, \dots, u_m) can be extended to a basis $(u_1, \dots, u_m, w_1, \dots, w_n)$ of V . Let $W = \text{span}(w_1, \dots, w_n)$.

To show that $V = U \oplus W$, we need to show that $V = U + W$ and $U \cap W = \{0\}$. Since $V = \text{span}(u_1, \dots, u_m, w_1, \dots, w_n)$ and (u_1, \dots, u_m) spans U and (w_1, \dots, w_n) spans W , it is clear that $V = U + W$.

To show that $U \cap W = \{0\}$ let $v \in U \cap W$. Then there exist scalars $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{F}$ such that

$$v = a_1u_1 + \dots + a_mu_m = b_1w_1 + \dots + b_nw_n$$

or equivalently

$$a_1u_1 + \dots + a_mu_m - b_1w_1 - \dots - b_nw_n = 0.$$

Since $(u_1, \dots, u_m, w_1, \dots, w_n)$ forms a basis of V and hence is linearly independent, the only solution to this equation is $a_1 = \dots = a_m = b_1 = \dots = b_n = 0$. Hence $v = 0$, proving that indeed $U \cap W = \{0\}$. \square

Theorem 11. *If $U, W \subset V$ are subspaces of a finite-dimensional vector space, then*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Proof. Let (v_1, \dots, v_n) be a basis of $U \cap W$. By the Basis Extension Theorem 7, there exist (u_1, \dots, u_k) and (w_1, \dots, w_ℓ) such that $(v_1, \dots, v_n, u_1, \dots, u_k)$ is a basis of U and $(v_1, \dots, v_n, w_1, \dots, w_\ell)$ is a basis of W . It suffices to show that

$$\mathcal{B} = (v_1, \dots, v_n, u_1, \dots, u_k, w_1, \dots, w_\ell)$$

is a basis of $U + W$, since then

$$\dim(U + W) = n + k + \ell = (n + k) + (n + \ell) - n = \dim U + \dim W - \dim(U \cap W).$$

Clearly $\text{span}(v_1, \dots, v_n, u_1, \dots, u_k, w_1, \dots, w_\ell)$ contains U and W and hence $U + W$. To show that \mathcal{B} is a basis it hence remains to show that \mathcal{B} is linearly independent. Suppose

$$a_1v_1 + \dots + a_nv_n + b_1u_1 + \dots + b_ku_k + c_1w_1 + \dots + c_\ell w_\ell = 0. \quad (3)$$

Let $u = a_1v_1 + \dots + a_nv_n + b_1u_1 + \dots + b_ku_k \in U$. Then by (3) also $u = -c_1w_1 - \dots - c_\ell w_\ell \in W$, which implies that $u \in U \cap W$. Hence there exist scalars $a'_1, \dots, a'_n \in \mathbb{F}$ such that $u = a'_1v_1 + \dots + a'_nv_n$. Since there is only a unique linear combination of the linearly independent vectors $(v_1, \dots, v_n, u_1, \dots, u_k)$ that describes u , we must have $b_1 = \dots = b_k = 0$ and $a_1 = a'_1, \dots, a_n = a'_n$. Since $(v_1, \dots, v_n, w_1, \dots, w_\ell)$ is also linearly independent, it further follows that $a_1 = \dots = a_n = c_1 = \dots = c_\ell = 0$. Hence (3) only has the trivial solution which implies that \mathcal{B} is a basis. \square