Spanning and Linear Independence

References are to Anton–Rorres, 7th edition

Coordinates Let V be a given vector space. We wish to equip V with a coordinate system, much as we did geometrically for the plane and space. We have the origin $\mathbf{0}$. However, because V is only a vector space, the concepts of length and orthogonality *do not apply.*

Take *any* set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ of vectors in V. There is an associated linear transformation $L: \mathbf{R}^r \to V$ (well hidden in Anton–Rorres), given by

$$L(k_1, k_2, \ldots k_r) = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots k_r\mathbf{v}_r \tag{1}$$

So $L(\mathbf{e}_i) = \mathbf{v}_i$ for each i. It is easy to check that L is linear. The idea is to choose S to make L an isomorphism of vector spaces, which will allow us to transfer everything from the general vector space V to the familiar vector space $\mathbf{R}^r$.

Definition 2 The set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ of vectors in V is a *basis* [plural: *bases*] *of* V if the above linear transformation (1) satisfies the two conditions:

(i) The *range* $R(L)$ of L is the whole of V;

(ii) The *kernel* $\mathrm{Ker}(L)$ of L is $\{\mathbf{0}\}$.

Then by Theorem 8.3.1, L is 1–1 and we can restate the definition explicitly.

Theorem 3 (=Thm. 5.4.1) *If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a basis of V, then any vector $\mathbf{v} \in V$ can be uniquely expressed as a linear combination*

$$\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots k_r\mathbf{v}_r = L(k_1, k_2, \ldots, k_r) \quad \square \tag{4}$$

We then have the *inverse* linear transformation $L^{-1}: V \to \mathbf{R}^r$ and $L^{-1}(\mathbf{v}_i) = \mathbf{e}_i$.

Definition 5 If S *is* a basis of V, we define the *coordinate vector relative to* S of any vector $\mathbf{v} \in V$ to be $L^{-1}(\mathbf{v})$, and write it $(\mathbf{v})_S$. This is a vector in $\mathbf{R}^r$.

Explicitly, if $\mathbf{v}$ is the linear combination (4), then $(\mathbf{v})_S = (k_1, k_2, \ldots, k_r)$.

Example We have the standard basis $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r\}$ of $\mathbf{R}^r$. In this case, L is the identity linear transformation and $(\mathbf{v})_S = \mathbf{v}$.

We break up Definition 2 and discuss the two conditions separately.

Spanning In any case, the range $R(L)$ of L is always a subspace of V.

Definition 6 For any set S in V, we define the *span* of S to be the range $R(L)$ of the linear transformation L in equation (1), and write $\mathrm{span}(S) = R(L)$.

Explicitly, $\mathrm{span}(S)$ is the set of all linear combinations (4). Many different sets of vectors S can span the same subspace. Clearly, we can omit the zero vector $\mathbf{0}$ if it is present in S. More generally, as a direct application of Theorem 5.2.4, we have the following reduction, known as the Minus Theorem.

Lemma 7 (=Thm. 5.4.4(b)) *Suppose $\mathbf{v}_i \in S$ is a linear combination of the other vectors in S. Let S' denote the set S with $\mathbf{v}_i$ removed. Then* $\mathrm{span}(S') = \mathrm{span}(S)$. $\square$

Linear independence

DEFINITION 8 The set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is *linearly independent* if the kernel $\mathrm{Ker}(L)$ of the linear transformation L in equation (1) is $\{\mathbf{0}\}$, i. e. L is 1–1 (see Thm. 8.3.1). Otherwise, S is *linearly dependent*. [As linear independence is clearly the desirable condition, we shall eschew the term "linearly dependent".]

Explicitly, S is linearly independent if there is no *linear relation*

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_r\mathbf{v}_r = \mathbf{0} \tag{9}$$

between the $\mathbf{v}$'s, other than the obvious *trivial* relation

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_r = \mathbf{0}$$

The following property is clear enough, but note the direction of the implication.

LEMMA 10 *Let $T: V \to W$ be a linear transformation and $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ a set of vectors in V. If the image set $T(S) = \{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_r)\}$ is linearly independent in W, then S is linearly independent in V.*

Proof Suppose the vectors in S satisfy the linear relation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_r\mathbf{v}_r = \mathbf{0}$$

We apply T to this to see that $T(S)$ satisfies the corresponding linear relation

$$k_1T(\mathbf{v}_1) + k_2T(\mathbf{v}_2) + \ldots + k_rT(\mathbf{v}_r) = \mathbf{0} \quad \square$$

We need to recast the definition of linear independence in a more useful form. Roughly stated, S is linearly independent if each vector in S is new in the sense that it cannot be expressed in terms of the previous members of S.

LEMMA 11 (=Thm. 5.3.1(b), but sharper) *The set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ of vectors is linearly independent if and only if $\mathbf{v}_1 \neq \mathbf{0}$ and no vector $\mathbf{v}_i \in S$ is a linear combination of the preceding vectors in S, i. e. for $2 \le i \le r$, $\mathbf{v}_i \notin \mathrm{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}\}$.*

Proof Assume there is a nontrivial linear relation (9), with k_m as the last nonzero coefficient. Then we can divide by k_m and rearrange equation (9) as

$$\mathbf{v}_m = l_1\mathbf{v}_1 + l_2\mathbf{v}_2 + \ldots + l_{m-1}\mathbf{v}_{m-1} \tag{12}$$

where each $l_i = -k_i/k_m$. This expresses $\mathbf{v}_m$ as a linear combination of the preceding vectors. (If $m = 1$, equation (12) degenerates to $\mathbf{v}_1 = \mathbf{0}$.) Conversely, if equation (12) holds for some m, we can rearrange it as the linear relation

$$l_1\mathbf{v}_1 + l_2\mathbf{v}_2 + \ldots + l_{m-1}\mathbf{v}_{m-1} - \mathbf{v}_m = \mathbf{0} \quad \square$$

COROLLARY 13 (=Thm. 5.4.4(a), the Plus Theorem) *Suppose the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ of vectors in V is linearly independent but does not span V. Take any vector $\mathbf{v}_{r+1} \notin \mathrm{span}(S)$. Then the enlarged set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}\}$ remains linearly independent.*

Proof The condition of Lemma 11 holds for $\mathbf{v}_i$ if $i \le r$ because S is linearly independent. It holds for $\mathbf{v}_{r+1}$ by hypothesis. $\square$

Occasionally, a variant of Lemma 11 is useful.

Corollary 14 *The set* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ *of vectors in* V *is linearly independent if and only if* $\mathbf{v}_r \neq \mathbf{0}$ *and for* $1 \le i < r$, $\mathbf{v}_i$ *is not a linear combination of the later vectors in* S.

Proof We simply write the set S in reverse order and apply Lemma 11. □

Vectors in $\mathbf{R}^n$ All the main results depend ultimately on the following fact, which is intuitively obvious but *not* trivial to prove. However, the real work has already been done. Roughly, the consequence is that in a given vector space, a spanning set of vectors cannot be too small, and a linearly independent set cannot be too large.

Lemma 15 (=Thm. 5.3.3) *A linearly independent set* S *of vectors in* $\mathbf{R}^n$ *has at most* n *members.*

Proof Suppose S has r members, and consider the linear transformation $L: \mathbf{R}^r \to \mathbf{R}^n$ in equation (1). We are given $\mathrm{Ker}(L) = \{\mathbf{0}\}$. Let A be the matrix of L, so that $L(\mathbf{x}) = A\mathbf{x}$. We know $A\mathbf{x} = \mathbf{0}$ has no nontrivial solutions. Since A is an $n \times r$ matrix, Theorem 1.2.1 shows that we must have $r \le n$. □

From this we deduce the result we really want.

Theorem 16 *Suppose the vector space* V *is spanned by a set containing* n *vectors. Then any linearly independent set of vectors in* V *contains at most* n *members.*

Proof From the given spanning set, we construct as in equation (1) a linear transformation $L: \mathbf{R}^n \to V$ such that $R(L) = V$. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be any linearly independent set of vectors in V. Since $R(L) = V$, we can choose for each i a vector $\mathbf{u}_i \in \mathbf{R}^n$ such that $L(\mathbf{u}_i) = \mathbf{v}_i$. Then by Lemma 10, the set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ is linearly independent in $\mathbf{R}^n$. Lemma 15 now shows that $r \le n$. □

Corollary 17 (=Thm. 5.4.3) *Any two bases of* V *contain the same number of vectors.* □

Definition 18 A vector space V is *finite-dimensional* if it has a basis that contains n vectors for some finite n. The number n is the *dimension* of V and is written $\dim(V)$. (By Corollary 17, it is well defined. For completeness, the zero vector space is considered to have dimension 0, and the *empty* set (not $\{\mathbf{0}\}$) as a basis; this works.)

We say that V is *infinite-dimensional* if it does not have a finite basis.

We are primarily interested in finite-dimensional vector spaces.

Theorem 19 *Every finite-dimensional vector space is isomorphic to the standard vector space* $\mathbf{R}^n$ *for a unique integer* n. □

We collect in one place all the information about subsets of V.

Theorem 20 (=Thms. 5.4.2 and 5.4.5) *Let* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ *be any set of* r *vectors in the* n*-dimensional vector space* V. *Then:*

(a) *If* $r < n$, S *does not span* V. (It may or may not be linearly independent.)

(b) *If* $r = n$, S *spans* V *if and only if it is linearly independent. Thus* S *is a basis of* V *if* either *of these conditions holds.*

(c) *If* $r > n$, S *is not linearly independent.* (S may or may not span V.)

Proof Parts (a) and (c) both follow immediately from Theorem 16.

For (b), suppose first that S spans V. If S is not linearly independent, Lemma 11 shows that some $\mathbf{v}_i \in S$ is a linear combination of the other members. We remove $\mathbf{v}_i$ from S to get a set S' of $n-1$ vectors. By Lemma 7, S' still spans V; but this contradicts (a).

Conversely, suppose that S is linearly independent. If S does not span V, we could use Corollary 13 to add another vector to S to form a linearly independent set of $n+1$ vectors in V, which would contradict (c). □

It is not immediately obvious that any subspace of a finite-dimensional vector space is finite-dimensional.

THEOREM 21 (=Thm. 5.4.7) *If W is a subspace of the finite-dimensional vector space V, then W is again finite-dimensional and* $\dim(W) \leq \dim(V)$*, with equality only if $W = V$.*

Proof Choose a linearly independent set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ of vectors in W with r as large as possible; such a set exists because these vectors also lie in V, so that by Theorem 20(c), $r \leq n$ where $n = \dim(V)$. Then $\mathrm{span}(S) = W$, otherwise Corollary 13 (applied in W) would allow us to extend S by one more vector and increase r by 1. So S must be a basis of W and $\dim(W) = r$.

If $W \neq V$, choose any vector $\mathbf{v}_{r+1} \in V$ that is not in W; by Corollary 13, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}\}$ is still linearly independent. By Theorem 20(c), $r+1 \leq n$. □

Constructing bases Obviously, if a subset S of V spans V, so does any subset S' of V that contains S. Any subset of a linearly independent set S remains linearly independent. Beyond these restrictions, we can construct bases of V as follows.

THEOREM 22 (=Thm. 5.4.6(a)) *Suppose the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ spans the vector space V. Then we can* thin out *S to find a subset $S' \subset S$ that is a basis of V. In particular, V is finite-dimensional.* (Explicitly, one possibility is to take S' as the set of all the nonzero $\mathbf{v}_i \in S$ that are not linear combinations of the preceding members of S, but there are other choices.)

Proof If $\mathbf{v}_i \in S$ is a linear combination of the other members of S, we can delete $\mathbf{v}_i$ from S without affecting $\mathrm{span}(S)$, by Lemma 7. We repeat this, deleting elements of S *one at a time*, until we can go no further. By Lemma 11, the end result S' is linearly independent and therefore a basis of V. (We leave the suggested candidate for S' as an exercise.) □

THEOREM 23 (=Thm. 5.4.6(b)) *Suppose given any linearly independent set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ of vectors in a finite-dimensional vector space V that does not span V. Then we can* extend *S to a basis $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \ldots, \mathbf{v}_n\}$ of V.*

Proof By Corollary 13, we extend S by one more vector to get a larger subset that is still linearly independent. We repeat as long as possible, until we find a linearly independent set S' that does span V and is therefore a basis. The process must terminate, because by Theorem 20(c), the set S' can never have more than n members, where $n = \dim(V)$. (In fact, it has exactly n, by Corollary 17.) □