5. DIMENSION

We are now in a position to define the *dimension* of a vector space with a basis, and to show that dimension is well-defined. Implicitly we have already seen this result in the tests for linear independent sets and for spanning sets. We showed in those tests that a linear independent subset of $\mathbb{R}^n$ cannot have more than n elements and that a spanning set of $\mathbb{R}^n$ cannot have fewer than n. The proof below has the merit of not relying on coordinates.

Theorem 127 (Steinitz Exchange Lemma) *Let V be a vector space over a field $\mathbb{F}$. Take $X = \{v_1, v_2, \ldots, v_n\} \subseteq V$. Suppose that $u \in \langle X \rangle$ but that $u \notin \langle X \backslash \{v_i\} \rangle$ for some i. Let*

$$Y = (X \backslash \{v_i\}) \cup \{u\}$$

(that is, we "exchange u for v_i"). Then $\langle Y \rangle = \langle X \rangle$.

Proof. Since $u \in \langle X \rangle$, there are $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

There is $v_i \in X$ such that $u \notin \langle X \backslash \{v_i\} \rangle$. Without loss of generality, we may assume that $i = n$. Since $u \notin \langle X \backslash \{v_n\} \rangle$, we see that $\alpha_n \neq 0$. So we can divide by α_n and rearrange, to obtain

$$v_n = \frac{1}{\alpha_n}(u - \alpha_1 v_1 - \cdots - \alpha_{n-1} v_{n-1}).$$

Now if $w \in \langle Y \rangle$ then we have an expression of w as a linear combination of elements of Y. We can replace u by $\alpha_1 v_1 + \cdots + \alpha_n v_n$ to express w as a linear combination of elements of X. So $\langle Y \rangle \subseteq \langle X \rangle$. And if $w \in \langle X \rangle$ then we have an expression of w as a linear combination of elements of X. We can replace v_n by

$$\frac{1}{\alpha_n}(u - \alpha_1 v_1 - \cdots - \alpha_{n-1} v_{n-1})$$

to express w as a linear combination of elements of Y. So $\langle Y \rangle \supseteq \langle X \rangle$. ∎

The Steinitz Exchange Lemma is called a lemma, which sounds unimportant, and it looks a bit like a niche technical result. But in fact it is fundamental to defining the dimension of a vector space.

Theorem 128 *Let V be a vector space. Let S, T be finite subsets of V. If S is linearly independent and T spans V, then $|S| \leqslant |T|$.*

Proof. Assume that S is linearly independent and that T spans V. List the elements of S as $u_1, \ldots, u_m$ and the elements of T as $v_1, \ldots, v_n$. We will use the Steinitz Exchange Lemma to swap out the elements of T with those of S, one at a time, ultimately exhausting S.

Let $T_0 = \{v_1, \ldots, v_n\}$. Since $\langle T_0 \rangle = V$, then $u_1 \in \langle v_1, \ldots, v_i \rangle$ for some $1 \leqslant i \leqslant n$ and choose i to be minimal in this regard. Note then that $u_1 \in \langle v_1, \ldots, v_i \rangle$ but that $u_1 \notin \langle v_1, \ldots, v_{i-1} \rangle$. The Steinitz Exchange Lemma then shows that

$$\langle v_1, \ldots, v_i \rangle = \langle u_1, v_1, \ldots, v_{i-1} \rangle$$

and hence

$$\begin{aligned} V &= \langle v_1, \ldots, v_n \rangle \\ &= \langle v_1, \ldots, v_i \rangle + \langle v_{i+1}, \ldots, v_n \rangle \\ &= \langle u_1, v_1, \ldots, v_{i-1} \rangle + \langle v_{i+1}, \ldots, v_n \rangle \\ &= \langle u_1, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle. \end{aligned}$$

Now, by relabelling the elements of T, we can assume without loss of generality assume that u_1 has been exchanged for v_1 and we set

$$T_1 = \{u_1, v_2, \ldots, v_n\} \qquad \text{noting that} \qquad \langle T_1 \rangle = V.$$

We proceed inductively in this manner creating sets

$$T_k = \{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\} \qquad \text{such that} \qquad \langle T_k \rangle = V..$$

Note that at each stage $u_{k+1} \in \langle T_k \rangle$ but that $u_{k+1} \notin \langle u_1, \ldots, u_k \rangle$ as the set S is independent. Hence we can keep continuing to replace elements of T with elements of S. The process can only terminate when S is exhausted which means that $m \leqslant n$. ■

Corollary 129 *Let V be a finite-dimensional vector space. All bases of V are finite and of the same size.*

Proof. Since V is finite-dimensional then V has a finite basis B. By Theorem 128 any finite linearly independent subset of V has size at most $|B|$. Given another basis S of V, it is linearly independent, so every finite subset of S is linearly independent. So in fact S must be finite, and $|S| \leqslant |B|$. But B is linearly independent and S is spanning and so by Theorem 128 $|B| \leqslant |S|$. ■

Definition 130 *Let V be a finite-dimensional vector space. The* ***dimension*** *of V, written $\dim V$, is the size of any basis of V.*

Definition 131 *We can now redefine row rank using this notion of dimension. The* ***row rank*** *of a matrix is the dimension of its* ***row space.*** *When in RRE form, the non-zero rows of the matrix are linearly independent. Further EROs do not affect the row space. So the non-zero rows of a matrix in RRE form are a basis of the row space.*

5.1 Subspaces and Dimension

We include the following result here as it fits in naturally with some of the subsequent results; in what follows we will show:

- A spanning set contains a basis.
- A linearly independent set can be extended to a basis. (This result requires the notion of dimension.)
- A basis is a maximal linearly independent set.
- A basis is a minimal spanning set.

Proposition 132 *Let V be a vector space over $\mathbb{F}$ and let S be a finite spanning set. Then S contains a basis.*

Remark 133 *That is, if V has a finite spanning set, then V has a basis. We say nothing here about what happens if V does not have a finite spanning set. This question is addressed in the Part B course on Set Theory (using the Axiom of Choice).*

Proof. Let S be a finite spanning set for V. Take $T \subseteq S$ such that T is linearly independent, and T is a largest such set (that is, no linearly independent subset of S strictly contains T). Suppose, for a contradiction, that $\langle T \rangle \neq V$. Then, since $\langle S \rangle = V$, there must exist $v \in S \backslash \langle T \rangle$.

Now by Lemma 102 we see that $T \cup \{v\}$ is linearly independent, and $T \cup \{v\} \subseteq S$, and $|T \cup \{v\}| > |T|$, which contradicts the maximality of T. So T spans V, is linearly independent, and thus a basis. ■

Proposition 134 *Let U be a subspace of a finite-dimensional vector space V. Then U is finite-dimensional, and $\dim U \leqslant \dim V$. Further if $\dim U = \dim V$ then $U = V$.*

Proof. Let $n = \dim V$. Then, by Theorem 128, every linearly independent subset of V has size at most n. Let S be a largest linearly independent set contained in U (and so in V), so $|S| \leqslant n$.

Suppose, for a contradiction, that $\langle S \rangle \neq U$. Then there exists $u \in U \backslash \langle S \rangle$. Now by Lemma 102 $S \cup \{u\}$ is linearly independent, and $|S \cup \{u\}| > |S|$, which contradicts our choice of S. So $U = \langle S \rangle$ and S is linearly independent, so S is a basis of U, and as we noted earlier $|S| \leqslant n$.

Say now that $\dim U = \dim V$ and $U \neq V$. Then there exists $v \in V \backslash U$. This v may then be added to a basis of U to create a linearly independent subset of V with

$$\dim U + 1 = \dim V + 1$$

vectors, which is a contradiction. Hence $\dim U = \dim V$ implies $U = V$. ■

Proposition 135 *Let V be a finite-dimensional vector space over $\mathbb{F}$ and let S be a linearly independent set. Then there exists a basis B such that $S \subseteq B$.*

Proof. If $\langle S\rangle = V$ then we are done as S is linearly independent and spanning, and so a basis. If $\langle S\rangle \neq V$ then by Lemma 102 we can extend S to $S_1 = S \cup \{u_1\}$ where $u_1 \in U \backslash \langle S\rangle$ to create a larger linearly independent set. If $\langle S_1\rangle = V$, then we are done as S_1 is a basis. This process can continue and only terminates at some S_k if S_k is a basis. However this process must terminate as we know every linearly independent subset of V must contain at most $\dim V$ elements. ■

Corollary 136 *A maximal linearly independent subset of a finite-dimensional vector space is a basis.*

Proof. Let S be a maximal linearly independent subset of a finite-dimensional vector space V. If $\langle S\rangle \neq V$ then by Lemma 102 we can extend $S_1 = S \cup \{u_1\}$ which is still linearly independent, but contradicts the maximality of S. So $\langle S\rangle = V$. ■

Corollary 137 *A minimal spanning subset of a finite-dimensional vector space is a basis.*

Proof. Let S be a minimal spanning subset of a finite-dimensional vector space V. If S is not linearly independent, then there exists $v \in S$ which can be written as a linear combination of elements of $S\backslash\{v\}$. Then $S\backslash\{v\}$ is linearly independent, and as shown in Lemma 102 $S\backslash\{v\}$ is still spanning, which contradicts the minimality of S. ■

Question Let S be a finite set of vectors in $\mathbb{R}^n$. How can we (efficiently) find a basis of $\langle S\rangle$?

Example 138 *Let* $S = \{(0,1,2,3),(1,2,3,4),(2,3,4,5)\} \subseteq \mathbb{R}^4$. *Define*

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{pmatrix}.$$

So $\langle S\rangle = \mathrm{Row}(A)$. *Applying EROs to* A *does not change the row space. Now*

$$\mathrm{RRE}(A) = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

As has been commented before, the non-zero rows are a basis for the row space, or equivalently for $\langle S\rangle$.

5.2 The dimension formula

We previously saw that the sum $U + W$ and intersection $U \cap W$ of two subspaces are subspaces. We now prove a useful theorem connecting their dimensions. Recall that we can extend bases of subspaces to bases of larger spaces, but in general a basis of a vector space won't contain a basis of a subspace (or possibly even any elements from the subspace). Thus it makes sense to begin with $U \cap W$, the smallest of the relevant spaces.

The next result is particularly useful.

Theorem 139 (Dimension Formula) *Let U, W be subspaces of a finite-dimensional vector space V over $\mathbb{F}$. Then*

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

Proof. Take a basis $v_1, \ldots, v_m$ of $U \cap W$. Now $U \cap W \leqslant U$ and $U \cap W \leqslant W$, so by Theorem 135 we can separately extend this set to a basis $v_1, \ldots, v_m, u_1, \ldots, u_p$ of U, and a basis $v_1, \ldots, v_m, w_1, \ldots, w_q$ of W. With this notation, we see that

$$\dim(U \cap W) = m, \qquad \dim U = m + p, \qquad \dim W = m + q.$$

We aim to show that $S = \{v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q\}$ is a basis of $U + W$. It contains

$$\begin{aligned} m + p + q &= (m + p) + (m + q) - m \\ &= \dim U + \dim W - \dim(U \cap W) \end{aligned}$$

elements. So if we can verify this aim then the result follows.

S is spanning: Take $x \in U + W$, so that $x = u + w$ for some $u \in U$, $w \in W$. Then

$$\begin{aligned} u &= \alpha_1 v_1 + \cdots + \alpha_m v_m + \alpha_1' u_1 + \cdots + \alpha_p' u_p, \\ w &= \beta_1 v_1 + \cdots + \beta_m v_m + \beta_1' w_1 + \cdots + \beta_q' w_q \end{aligned}$$

for some scalars α_i, α_i', β_i, $\beta_i' \in \mathbb{F}$. Then

$$x = u + w = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_m + \beta_m)v_m + \alpha_1' u_1 + \cdots + \alpha_p' u_p + \beta_1' w_1 + \cdots + \beta_q' w_q \in \langle S \rangle,$$

showing $\langle S \rangle = U + W$.

S is linearly independent: Take $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_q \in \mathbb{F}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 u_1 + \cdots + \beta_p u_p + \gamma_1 w_1 + \cdots + \gamma_q w_q = 0.$$

Then

$$\alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 u_1 + \cdots + \beta_p u_p = -(\gamma_1 w_1 + \cdots + \gamma_q w_q).$$

The vector on the LHS is in U, and the vector on the RHS is in W. So they are both in $U \cap W$. As $v_1, \ldots, v_m$ form a basis of $U \cap W$, there are $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ such that

$$-(\gamma_1 w_1 + \cdots + \gamma_q w_q) = \lambda_1 v_1 + \cdots + \lambda_m v_m,$$

which rearranges to

$$\gamma_1 w_1 + \cdots + \gamma_q w_q + \lambda_1 v_1 + \cdots + \lambda_m v_m = 0.$$

But $\{v_1, \ldots, v_m, w_1, \ldots, w_q\}$ is linearly independent (it's a basis for W), and so each γ_i is 0. This then implies that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 u_1 + \cdots + \beta_p u_p = 0.$$

But $\{v_1, \ldots, v_m, u_1, \ldots, u_p\}$ is linearly independent (it's a basis for U), so each α_i and β_i equals 0. So S is linearly independent and the result follows. ■

Example 140 *Let V be a vector space of dimension 10. Let X, Y be subspaces of dimension 6. Then $X + Y \leqslant V$ so $\dim(X + Y) \leqslant \dim V = 10$. So, by the dimension formula,*

$$\dim(X \cap Y) = \dim(X) + \dim(Y) - \dim(X + Y) \geqslant 6 + 6 - 10 = 2.$$

It is not hard to see that the possibilities

$$2 \leqslant \dim(X \cap Y) \leqslant 6,$$

are all possible. This is left as an exercise.

Definition 141 *Let U, W be subspaces of a vector space V. If $U \cap W = \{0_V\}$ and $U + W = V$, then we say that V is the **direct sum** of U and W, and we write $V = U \oplus W$.*

Proposition 142 *Let U, W be subspaces of a finite-dimensional vector space V. The following are equivalent:*
(a) $V = U \oplus W$;
(b) every $v \in V$ has a unique expression as $u + w$ where $u \in U$ and $w \in W$;
(c) $\dim V = \dim U + \dim W$ and $V = U + W$;
(d) $\dim V = \dim U + \dim W$ and $U \cap W = \{0_V\}$;
(e) if $u_1, \ldots, u_m$ is a basis for U and $w_1, \ldots, w_n$ is a basis for W, then $u_1, \ldots, u_m, w_1, \ldots, w_n$ is a basis for V.

Proof. Exercise. Hint: (a) $\Leftrightarrow$ (b) follows from the definition of direct sum.
Try using the dimension formula to prove that (a)/(b) are equivalent to (c)/(d)/(e). ■

Remark 143 *To conclude, two more general comments on direct sums.*

- *A vector space V is said to be the direct sum of subspaces $X_1, \ldots, X_k \leqslant V$ if every $v \in V$ can be uniquely written*

$$v = x_1 + x_2 + \cdots + x_k \qquad \text{where} \quad x_i \in X_i \quad \text{for all } i.$$

  Thus it is statement (b) in the proposition above which naturally generalizes.

- *Writing a vector space as a sum of subspaces is called an **internal direct sum**. Given vectors spaces $V_1, \ldots, V_k$ then the **external direct sum***

$$V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

  has the Cartesian product $V_1 \times V_2 \times \cdots \times V_k$ as the underlying set, with addition and scalar multiplication defined componentwise. That is

$$\begin{aligned} (v_1, \ldots, v_k) + (w_1, \ldots, w_k) &= (v_1 + w_1, \ldots, v_k + w_k); \\ \alpha.(v_1, \ldots, v_k) &= (\alpha.v_1, \ldots, \alpha.v_k) \end{aligned}$$