

Chapter 4

Vector spaces

Learning goals

After this chapter you should be able to:

- test whether a subset of $\mathbb{F}^n$ is a subspace, and find the column space and null space of a matrix;
- find a basis of a subspace of $\mathbb{F}^n$ and its dimension;
- work with vector spaces over a field: polynomials, matrices and functions as vectors;
- form sums and intersections of subspaces and recognise direct sums;
- extend independent lists to bases and use the dimension of a vector space;
- use the dimension formula $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

4.1 Subspaces of $\mathbb{R}^n$

Lecture 19

Two kinds of sets have kept turning up. In Section 1.3, the span of one vector in $\mathbb{R}^3$ was a line through the origin, and the span of two vectors in different directions was a plane through the origin. In Section 2.5, the solutions of a homogeneous system formed a line or a plane through the origin as well. These sets share three properties: each contains the zero vector, and adding or scaling vectors in it never leads out of it. A set with these properties is called a *subspace*. This section defines subspaces, shows that every span is one, and introduces the two subspaces that come with every matrix: the *column space*, which tells us for which $\mathbf{b}$ the system $A\mathbf{x} = \mathbf{b}$ can be solved, and the *null space*, the solution set of $A\mathbf{x} = \mathbf{0}$.

Throughout, $\mathbb{F}$ is a field (Section 2.1) and $n \geq 1$. The rules of vector algebra (Theorem 1.2.7) and their consequences (Proposition 1.2.8) hold in $\mathbb{F}^n$ with the same proofs, and we use them freely. The pictures are drawn in $\mathbb{R}^2$ and $\mathbb{R}^3$.

4.1.1 Subspaces

Definition 4.1.1 (Subspace). A subset H of $\mathbb{F}^n$ is a *subspace* of $\mathbb{F}^n$ if

- (a) $\mathbf{0} \in H$;
- (b) H is closed under addition: if $\mathbf{u}, \mathbf{v} \in H$, then $\mathbf{u} + \mathbf{v} \in H$;
- (c) H is closed under scalar multiplication: if $\mathbf{u} \in H$ and $c \in \mathbb{F}$, then $c\mathbf{u} \in H$.

Closure says that once we are inside H , adding and scaling never take us out of it. Condition (a) makes sure that H is not empty; given (c), it is the same as asking that H contain at least one vector $\mathbf{u}$, since then $\mathbf{0} = 0\mathbf{u} \in H$ (Proposition 1.2.8(d)).

Example 4.1.2 (First subspaces). (a) The set $\{\mathbf{0}\}$ is a subspace of $\mathbb{F}^n$, called the *zero subspace*: $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $c\mathbf{0} = \mathbf{0}$ (rule (3) and Proposition 1.2.8(e)). The whole space $\mathbb{F}^n$ is a subspace of itself, since sums and multiples of vectors of $\mathbb{F}^n$ lie in $\mathbb{F}^n$. These two are the *trivial* subspaces.

(b) The line $L = \{(x, 2x) : x \in \mathbb{R}\}$ through the origin of $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$. Indeed, $\mathbf{0} = (0, 2 \cdot 0) \in L$. If $\mathbf{u} = (s, 2s)$ and $\mathbf{v} = (t, 2t)$ lie in L , then

$$\mathbf{u} + \mathbf{v} = (s + t, 2s + 2t) = (s + t, 2(s + t)) \in L, \quad c\mathbf{u} = (cs, c \cdot 2s) = (cs, 2(cs)) \in L.$$

In the plane: the sum of two arrows along the line, drawn tip to tail, stays on the line, and so does every stretch or flip of an arrow along it (Figure 4.1, left).

Example 4.1.3 (Three sets that are not subspaces). Each of these subsets of $\mathbb{R}^2$ fails at least one of the three conditions (Figure 4.1).

(a) The line $y = x + 1$ does not pass through the origin: $0 \neq 0 + 1$, so $\mathbf{0}$ is not on it, and (a) fails. (Conditions (b) and (c) fail too: $(0, 1)$ lies on it, but $(0, 1) + (0, 1) = 2(0, 1) = (0, 2)$ does not.)

(b) The first quadrant $Q = \{(x, y) : x \geq 0, y \geq 0\}$ contains $\mathbf{0}$ and is closed under addition, because sums of non-negative numbers are non-negative. But $(1, 1) \in Q$ and $(-1)(1, 1) = (-1, -1) \notin Q$, so (c) fails.

(c) The union of the two axes, $X = \{(x, y) : x = 0 \text{ or } y = 0\}$, contains $\mathbf{0}$ and is closed under scalar multiplication: if $x = 0$ then $cx = 0$, and if $y = 0$ then $cy = 0$. But $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ lie in X while $\mathbf{e}_1 + \mathbf{e}_2 = (1, 1)$ does not, so (b) fails.

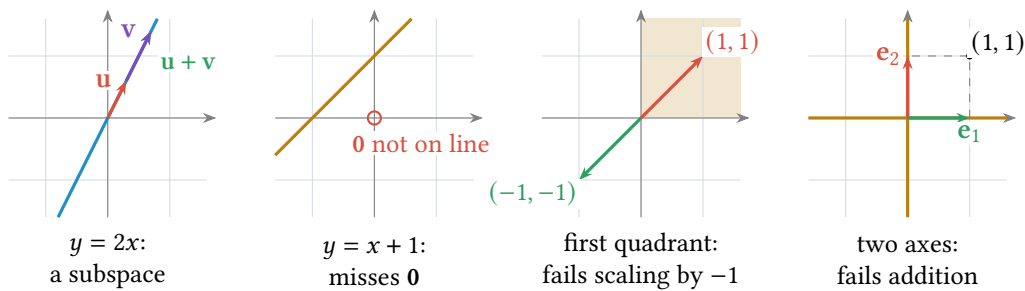


Figure 4.1: Left: the line $y = 2x$ is a subspace (Example 4.1.2(b)); a sum of two vectors on it stays on it. Then the three sets of Example 4.1.3: the line $y = x + 1$ misses $\mathbf{0}$; the first quadrant loses $(1, 1)$ when it is scaled by -1 ; the two axes lose $\mathbf{e}_1 + \mathbf{e}_2$.

Proposition 4.1.4 (Closed under linear combinations). Let H be a subspace of $\mathbb{F}^n$. If $\mathbf{v}_1, \dots, \mathbf{v}_k \in H$ and $c_1, \dots, c_k \in \mathbb{F}$, then $c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k \in H$.

Proof. Each $c_i \mathbf{v}_i$ lies in H by closure under scalar multiplication. We show by induction on j that the partial sum $\mathbf{s}_j = c_1 \mathbf{v}_1 + \cdots + c_j \mathbf{v}_j$ lies in H . For $j = 1$ this is $c_1 \mathbf{v}_1 \in H$. If $\mathbf{s}_j \in H$ and $j < k$, then $\mathbf{s}_{j+1} = \mathbf{s}_j + c_{j+1} \mathbf{v}_{j+1}$ is a sum of two vectors of H , so it lies in H by closure under addition. For $j = k$ we get the claim. $\square$

Remark 4.1.5 (A one-line test). A subset $H \subseteq \mathbb{F}^n$ is a subspace if and only if $\mathbf{0} \in H$ and $c \mathbf{u} + \mathbf{v} \in H$ for all $\mathbf{u}, \mathbf{v} \in H$ and $c \in \mathbb{F}$. If H is a subspace, this holds by Proposition 4.1.4 with weights c and 1 (and $1 \mathbf{v} = \mathbf{v}$). Conversely, suppose the test holds. With $c = 1$ it gives $\mathbf{u} + \mathbf{v} \in H$; with $\mathbf{v} = \mathbf{0}$ it gives $c \mathbf{u} + \mathbf{0} = c \mathbf{u} \in H$. So (a), (b) and (c) hold.

4.1.2 Spans are subspaces

Theorem 4.1.6 (Spans are subspaces). Let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{F}^n$ and $S = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

- (a) S is a subspace of $\mathbb{F}^n$ that contains $\mathbf{v}_1, \dots, \mathbf{v}_k$.
- (b) S is the smallest such subspace: every subspace H of $\mathbb{F}^n$ that contains $\mathbf{v}_1, \dots, \mathbf{v}_k$ contains S .

Proof. This is Theorem 1.3.10, now read over any field; we repeat the argument in the new language.

(a) First, $\mathbf{0} = 0 \mathbf{v}_1 + \cdots + 0 \mathbf{v}_k \in S$, by Proposition 1.2.8(d) and rule (3). Next, let $\mathbf{x} = c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k$ and $\mathbf{y} = d_1 \mathbf{v}_1 + \cdots + d_k \mathbf{v}_k$ lie in S . Regrouping by rules (1) and (2) and combining with rule (8),

$$\mathbf{x} + \mathbf{y} = (c_1 + d_1) \mathbf{v}_1 + \cdots + (c_k + d_k) \mathbf{v}_k \in S.$$

For $c \in \mathbb{F}$, rules (7) and (6) give $c \mathbf{x} = (cc_1) \mathbf{v}_1 + \cdots + (cc_k) \mathbf{v}_k \in S$. Finally, $\mathbf{v}_i \in S$: take weight 1 on $\mathbf{v}_i$ and 0 on the others.

(b) Every element of S is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$, which lie in H , so it lies in H by Proposition 4.1.4. $\square$

So every line through the origin, $\text{Span}\{\mathbf{v}\}$ with $\mathbf{v} \neq \mathbf{0}$, is a subspace, and so is every plane through the origin in $\mathbb{R}^3$. The zero subspace is a span too: $\{\mathbf{0}\} = \text{Span}\{\mathbf{0}\}$. In $\mathbb{F}^2$ these are all the subspaces there are.

Proposition 4.1.7 (The subspaces of $\mathbb{F}^2$). Every subspace H of $\mathbb{F}^2$ is one of: the zero subspace $\{\mathbf{0}\}$; a line $\text{Span}\{\mathbf{v}\}$ with $\mathbf{v} \neq \mathbf{0}$; or $\mathbb{F}^2$ itself.

Proof. Suppose $H \neq \{\mathbf{0}\}$, and pick $\mathbf{v} \in H$ with $\mathbf{v} \neq \mathbf{0}$. By Theorem 4.1.6(b), $\text{Span}\{\mathbf{v}\} \subseteq H$. If equality holds, H is a line and we are done.

Otherwise pick $\mathbf{w} \in H$ that is not in $\text{Span}\{\mathbf{v}\}$. Then $\mathbf{v}, \mathbf{w}$ are linearly independent. Indeed, by Proposition 2.6.5(c) they would otherwise be multiples of one another. Now $\mathbf{w} = c \mathbf{v}$ is impossible, since $\mathbf{w} \notin \text{Span}\{\mathbf{v}\}$. And $\mathbf{v} = c \mathbf{w}$ is impossible too: $c = 0$ would give $\mathbf{v} = \mathbf{0}$, and $c \neq 0$ would give $\mathbf{w} = c^{-1} \mathbf{v} \in \text{Span}\{\mathbf{v}\}$.

Let B be the 2×2 matrix with columns $\mathbf{v}, \mathbf{w}$. By the elimination test (Theorem 2.6.8) an echelon form of B has a pivot in each of its 2 columns. These 2 pivots lie in different rows, and B has only 2 rows, so every row has a pivot. By Theorem 3.1.9 the columns span $\mathbb{F}^2$. Hence, by Theorem 4.1.6(b) again, $\mathbb{F}^2 = \text{Span}\{\mathbf{v}, \mathbf{w}\} \subseteq H$, so $H = \mathbb{F}^2$. $\square$

Preview (Section 4.2, Lecture 20). Let H be a subspace of $\mathbb{F}^n$. Then H is the span of at most n vectors of H ; the least number of vectors needed is called the *dimension* of H . In particular, the subspaces of $\mathbb{R}^3$ are exactly $\{0\}$, the lines through the origin, the planes through the origin, and $\mathbb{R}^3$ itself. Both statements are proved in Section 4.2.

4.1.3 The column space

Definition 4.1.8 (Column space). Let A be an $m \times n$ matrix over $\mathbb{F}$ with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$. The *column space* of A is

$$\text{Col } A = \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subseteq \mathbb{F}^m.$$

Proposition 4.1.9. Let A be an $m \times n$ matrix over $\mathbb{F}$.

- (a) $\text{Col } A = \{A\mathbf{x} : \mathbf{x} \in \mathbb{F}^n\}$, and $\text{Col } A$ is a subspace of $\mathbb{F}^m$.
- (b) For $\mathbf{b} \in \mathbb{F}^m$, the system $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b} \in \text{Col } A$.

Proof. (a) By Definition 3.1.1, $A\mathbf{x} = x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n$. As $\mathbf{x}$ runs through $\mathbb{F}^n$, the weights $x_1, \dots, x_n$ run through all choices in $\mathbb{F}$, so the vectors $A\mathbf{x}$ are exactly the linear combinations of the columns. It is a subspace by Theorem 4.1.6(a).

(b) By Theorem 3.1.7, $A\mathbf{s} = \mathbf{b}$ exactly when $s_1\mathbf{a}_1 + \dots + s_n\mathbf{a}_n = \mathbf{b}$. So a solution exists exactly when $\mathbf{b}$ is a combination of the columns. $\square$

In the language of Section 3.2: $\text{Col } A$ is the set of all outputs of the transformation $\mathbf{x} \mapsto A\mathbf{x}$, its *range*. Deciding whether $\mathbf{b} \in \text{Col } A$ takes a computation, the row reduction of $[A \mid \mathbf{b}]$.

Example 4.1.10. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{a}_1 = (1, 0, 1), \quad \mathbf{a}_2 = (0, 1, 1).$$

The columns point in different directions, so $\text{Col } A$ is a plane through the origin of $\mathbb{R}^3$ (Figure 4.2). Which $\mathbf{b} = (b_1, b_2, b_3)$ lie in it? We row reduce $[A \mid \mathbf{b}]$:

$$\left[\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 1 & b_2 \\ 1 & 1 & b_3 \end{array} \right] \xrightarrow{R_3 \rightarrow R_3 - R_1} \left[\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 1 & b_2 \\ 0 & 1 & b_3 - b_1 \end{array} \right] \xrightarrow{R_3 \rightarrow R_3 - R_2} \left[\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 1 & b_2 \\ 0 & 0 & b_3 - b_1 - b_2 \end{array} \right].$$

By Theorem 2.5.3(a) the system is consistent exactly when $b_3 - b_1 - b_2 = 0$. So

$$\text{Col } A = \{\mathbf{b} \in \mathbb{R}^3 : b_3 = b_1 + b_2\}.$$

For instance $\mathbf{b} = (1, 2, 3)$ lies in $\text{Col } A$, since $3 = 1 + 2$; indeed $1\mathbf{a}_1 + 2\mathbf{a}_2 = (1, 2, 1+2) = (1, 2, 3)$. But $(1, 2, 0)$ does not, since $0 \neq 3$, and the system $A\mathbf{x} = (1, 2, 0)$ has no solution.

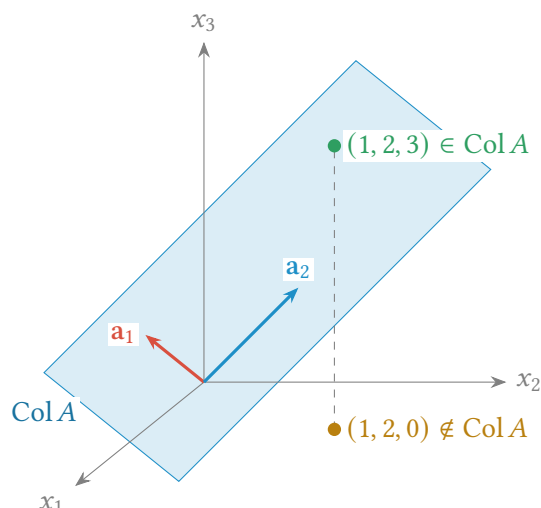


Figure 4.2: Example 4.1.10: $\text{Col } A$ is the plane $b_3 = b_1 + b_2$ (blue) spanned by the columns $\mathbf{a}_1$ (coral) and $\mathbf{a}_2$ (blue). The point $(1, 2, 3)$ lies on it; $(1, 2, 0)$ does not.

4.1.4 The null space

Definition 4.1.11 (Null space). Let A be an $m \times n$ matrix over $\mathbb{F}$. The *null space* of A is the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$:

$$\text{Nul } A = \{ \mathbf{x} \in \mathbb{F}^n : A\mathbf{x} = \mathbf{0} \} \subseteq \mathbb{F}^n.$$

Note where the two spaces live: for an $m \times n$ matrix, $\text{Col } A$ lies in $\mathbb{F}^m$, the space of outputs, while $\text{Nul } A$ lies in $\mathbb{F}^n$, the space of inputs.

Theorem 4.1.12. For every $m \times n$ matrix A over $\mathbb{F}$, $\text{Nul } A$ is a subspace of $\mathbb{F}^n$.

Proof. We check the three conditions, using the linearity of the product (Theorem 3.1.6). (a) $A\mathbf{0} = 0\mathbf{a}_1 + \cdots + 0\mathbf{a}_n = \mathbf{0}$, so $\mathbf{0} \in \text{Nul } A$. (b) If $A\mathbf{u} = \mathbf{0}$ and $A\mathbf{v} = \mathbf{0}$, then $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$. (c) If $A\mathbf{u} = \mathbf{0}$ and $c \in \mathbb{F}$, then $A(c\mathbf{u}) = c(A\mathbf{u}) = c\mathbf{0} = \mathbf{0}$ (Proposition 1.2.8(e)). $\square$

Remark 4.1.13 (Solutions of $A\mathbf{x} = \mathbf{b}$ with $\mathbf{b} \neq \mathbf{0}$). If $\mathbf{b} \neq \mathbf{0}$, the solution set of $A\mathbf{x} = \mathbf{b}$ is *not* a subspace, since $A\mathbf{0} = \mathbf{0} \neq \mathbf{b}$. When it is not empty, it is a shifted copy of the null space: by Theorem 3.1.11, for any one solution $\mathbf{p}$ the solutions are exactly the vectors $\mathbf{p} + \mathbf{h}$ with $\mathbf{h} \in \text{Nul } A$.

Example 4.1.14. Let

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}.$$

To describe $\text{Nul } A$ we solve $A\mathbf{x} = \mathbf{0}$; the right-hand side stays zero (Proposition 2.5.8), so we reduce A alone. One step, $R_2 \rightarrow R_2 - 2R_1$, gives the reduced row echelon form

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The pivot variable is x_1 , and $x_2 = s$, $x_3 = t$ are free. The only equation left is $x_1 + 2x_2 + x_3 = 0$, so

$x_1 = -2s - t$ and

$$\mathbf{x} = \begin{bmatrix} -2s - t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = s \mathbf{v} + t \mathbf{w}.$$

Hence $\text{Nul } A = \text{Span}\{\mathbf{v}, \mathbf{w}\}$, a plane through the origin of $\mathbb{R}^3$ (Figure 4.3). As a check, $A\mathbf{v} = -2(1, 2) + 1(2, 4) + 0(1, 2) = (0, 0)$ and $A\mathbf{w} = -1(1, 2) + 0(2, 4) + 1(1, 2) = (0, 0)$.

The column space of the same matrix lives in $\mathbb{R}^2$. The columns are $\mathbf{a}_1 = (1, 2)$, $\mathbf{a}_2 = (2, 4) = 2\mathbf{a}_1$ and $\mathbf{a}_3 = (1, 2) = \mathbf{a}_1$, so by Theorem 1.3.11 (applied twice, over any field) $\text{Col } A = \text{Span}\{\mathbf{a}_1\}$, the line through $(1, 2)$.

For $\mathbf{b} = (2, 4)$, one solution of $A\mathbf{x} = \mathbf{b}$ is $\mathbf{p} = (2, 0, 0)$, so by Remark 4.1.13 its solutions form the plane $\mathbf{p} + \text{Nul } A$, parallel to $\text{Nul } A$ and not through the origin.

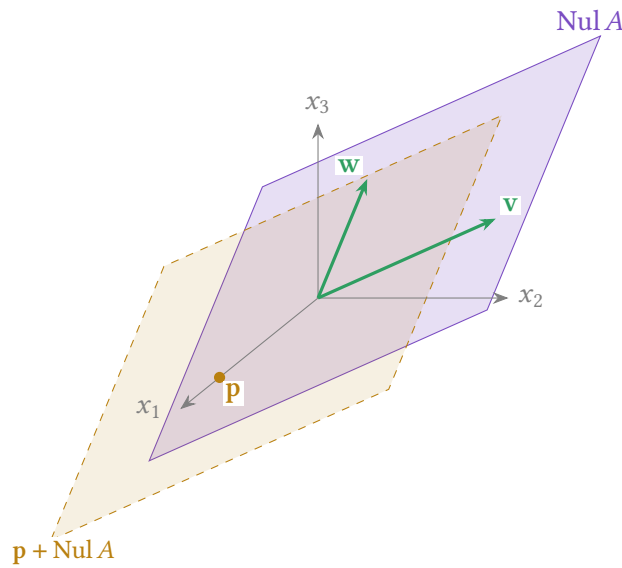


Figure 4.3: Example 4.1.14: $\text{Nul } A$ (violet) is the plane spanned by $\mathbf{v} = (-2, 1, 0)$ and $\mathbf{w} = (-1, 0, 1)$ (green). The solutions of $A\mathbf{x} = (2, 4)$ form the parallel plane through $\mathbf{p} = (2, 0, 0)$ (gold).

The method of Example 4.1.14 works for every matrix.

Proposition 4.1.15 (A spanning list for $\text{Nul } A$). Let A be an $m \times n$ matrix over $\mathbb{F}$. If $A\mathbf{x} = \mathbf{0}$ has no free variables, then $\text{Nul } A = \{\mathbf{0}\}$. Otherwise, write its general solution in parametric vector form, $\mathbf{x} = t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k$, one vector for each free variable. Then $\text{Nul } A = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

Proof. If there are no free variables, the solution is unique by Theorem 2.5.3(b), and $\mathbf{0}$ is a solution, so $\text{Nul } A = \{\mathbf{0}\}$.

Otherwise, row reduce $[A \mid \mathbf{0}]$ to reduced row echelon form; its last column stays zero (Proposition 2.5.8), so the constant part of the parametric vector form, computed from that column, is $\mathbf{p} = \mathbf{0}$ (Definition 2.5.2 and the procedure after it). By that definition every solution has the form $t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k$, so $\text{Nul } A \subseteq \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. Conversely, each $\mathbf{v}_i$ is a solution: it is the one with $t_i = 1$ and every other parameter 0. As $\text{Nul } A$ is a subspace (Theorem 4.1.12), Theorem 4.1.6(b) gives $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq \text{Nul } A$. $\square$

The two descriptions are opposite in a useful way. $\text{Col } A$ is given *explicitly*, by spanning vectors: listing members is easy (take any $A\mathbf{x}$), but testing whether a given $\mathbf{b}$ is a member needs a row reduction. $\text{Nul } A$

is given *implicitly*, by a condition: testing a given $\mathbf{x}$ is easy (compute $A\mathbf{x}$), but listing members needs a row reduction, which turns the condition into spanning vectors.

4.1.5 What the two spaces say about $A\mathbf{x} = \mathbf{b}$

Theorem 4.1.16. Let A be an $m \times n$ matrix over $\mathbb{F}$.

- (a) $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{F}^m$ if and only if $\text{Col } A = \mathbb{F}^m$, if and only if an echelon form of A has a pivot in every row.
- (b) $A\mathbf{x} = \mathbf{b}$ has at most one solution for every $\mathbf{b} \in \mathbb{F}^m$ if and only if $\text{Nul } A = \{\mathbf{0}\}$, if and only if an echelon form of A has a pivot in every column.
- (c) If $m = n$, then A is invertible if and only if $\text{Nul } A = \{\mathbf{0}\}$, if and only if $\text{Col } A = \mathbb{F}^n$.

Proof. (a) By Proposition 4.1.9(b), the first statement says that every $\mathbf{b} \in \mathbb{F}^m$ lies in $\text{Col } A$, that is, $\text{Col } A = \mathbb{F}^m$ (it is always a subset of $\mathbb{F}^m$). The equivalence with a pivot in every row is Theorem 3.1.9.

(b) $\text{Nul } A = \{\mathbf{0}\}$ says that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. By Proposition 3.1.13 this is equivalent to the uniqueness statement and to a pivot in every column.

(c) Let R be the reduced row echelon form of A . By Theorem 3.5.10(c), A is invertible exactly when $R = I_n$. If $R = I_n$, then R is an echelon form of A with a pivot in every row and every column, so $\text{Col } A = \mathbb{F}^n$ by (a) and $\text{Nul } A = \{\mathbf{0}\}$ by (b).

Suppose $\text{Nul } A = \{\mathbf{0}\}$. By (b), applied to the echelon form R , every one of the n columns of R has a pivot. In reduced form each pivot is 1, the only nonzero entry of its column, and the pivots step down and to the right. With n pivots in n columns, the pivot of column j must be in row j , so $R = I_n$ and A is invertible.

Suppose $\text{Col } A = \mathbb{F}^n$. By (a), applied to R , every one of the n rows of R has a pivot. A column holds at most one pivot, so n pivots fill all n columns, and $R = I_n$ as before. $\square$

Together with Remark 4.1.13, the theorem says how to read a linear system through its two subspaces: $\text{Col } A$ tells us *which* right-hand sides can be reached, and $\text{Nul } A$ tells us *how many* ways there are to reach each of them, since the solutions of a solvable system form one shifted copy of $\text{Nul } A$.

Preview. Section 4.2 (Lecture 20) measures the size of a subspace by its dimension, and proves that $\dim \text{Col } A + \dim \text{Nul } A = n$ for every $m \times n$ matrix A over $\mathbb{F}$ (Theorem 4.2.14(c)), the first form of the rank–nullity theorem.

Check yourself

- (a) Which of these subsets of $\mathbb{R}^3$ are subspaces? $H_1 = \{(x, y, z) : x - y + 2z = 0\}$, $H_2 = \{(x, y, z) : x - y + 2z = 1\}$, $H_3 = \{(x, y, z) : z = x^2\}$. Give a reason for each.
- (b) Let $A = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{bmatrix}$. Find vectors that span $\text{Nul } A$, and decide whether $\text{Col } A = \mathbb{R}^2$.
- (c) True or false: if A is a 3×5 matrix, then $\text{Nul } A \neq \{\mathbf{0}\}$.

4.2 Basis and dimension in $\mathbb{R}^n$

Lecture 20

Section 4.1 met the subspaces of $\mathbb{F}^n$: the lines and planes through the origin, the column space and the null space of a matrix. A subspace usually contains infinitely many vectors, so we want a short list that describes it. Take the floor of $\mathbb{R}^3$, the plane H of vectors with third entry 0. It is spanned by

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

but $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$ adds nothing (Section 2.6). Leave it out and $\mathbf{v}_1, \mathbf{v}_2$ still span the floor, with nothing redundant left. Such a list, spanning and independent, is called a *basis*. The floor has many bases, but every one of them has exactly two vectors; that number is its *dimension*, and proving that it does not depend on the basis is the main work of this section.

Throughout, $\mathbb{F}$ is a field and H is a subspace of $\mathbb{F}^n$: a subset that contains $\mathbf{0}$ and is closed under addition and scalar multiplication (Section 4.1). In the pictures $\mathbb{F} = \mathbb{R}$. We use two facts about spans. A subspace that contains $\mathbf{v}_1, \dots, \mathbf{v}_k$ contains every linear combination of them, that is, $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq H$ (Theorem 4.1.6(b)). And adding a vector to a list leaves the span unchanged exactly when that vector is already in the span (Theorem 1.3.11, proved for $\mathbb{R}^n$ with a proof that works word for word over any field).

4.2.1 Bases

Definition 4.2.1 (Basis). Let H be a subspace of $\mathbb{F}^n$. A list $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_k)$ of vectors in H is a *basis* of H if

- (i) $\mathbf{b}_1, \dots, \mathbf{b}_k$ are linearly independent, and
- (ii) $\mathbf{b}_1, \dots, \mathbf{b}_k$ span H , that is, $\text{Span}\{\mathbf{b}_1, \dots, \mathbf{b}_k\} = H$.

A basis is a list, so its vectors come in an order; the same vectors in another order form another basis of the same subspace. The zero subspace $\{\mathbf{0}\}$ needs a convention. Every list in it contains $\mathbf{0}$ and is dependent (Proposition 2.6.5(b)), but the span of the empty list is $\{\mathbf{0}\}$ (Definition 1.3.5), and the empty list has no nontrivial relation, because it has no weights at all. So the empty list counts as the basis of $\{\mathbf{0}\}$.

Example 4.2.2 (The standard basis). The standard vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

form a basis of $\mathbb{F}^n$, the *standard basis*. They are independent (Example 2.6.2), and they span $\mathbb{F}^n$, since every $\mathbf{x} = (x_1, \dots, x_n)$ is

$$\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n,$$

as the i -th entry of the right-hand side is $x_1 \cdot 0 + \dots + x_i \cdot 1 + \dots + x_n \cdot 0 = x_i$.

Example 4.2.3 (A second basis of $\mathbb{R}^2$). Let $\mathbf{b}_1 = (1, 2)$ and $\mathbf{b}_2 = (3, 1)$. They are independent, because neither is a multiple of the other (Example 2.6.6). They span $\mathbb{R}^2$: the replacement $R_2 \rightarrow R_2 - 2R_1$ takes the matrix with these columns to an echelon form with a pivot in every row,

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \begin{bmatrix} 1 & 3 \\ 0 & -5 \end{bmatrix} \quad (2 - 2 \cdot 1 = 0, \quad 1 - 2 \cdot 3 = -5),$$

so Theorem 3.1.9 applies. Hence $(\mathbf{b}_1, \mathbf{b}_2)$ is a basis of $\mathbb{R}^2$. Its slanted grid is drawn in Figure 4.4.

Two lists that are *not* bases of $\mathbb{R}^2$: the list $(1, 0), (2, 0), (0, 1)$ spans $\mathbb{R}^2$ (it contains $\mathbf{e}_1$ and $\mathbf{e}_2$), but it is dependent, since $2(1, 0) - (2, 0) = \mathbf{0}$. And the single vector $(1, 2)$ is independent but does not span $\mathbb{R}^2$: its multiples $(c, 2c)$ never equal $(1, 0)$, because $c = 1$ would give second entry 2.

So a basis is a balance between two demands. A spanning list must not be too short, and an independent list must not be too long; a basis is both at once.

4.2.2 Coordinates

Theorem 4.2.4 (Unique representation). Let $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_k)$ be a basis of a subspace H of $\mathbb{F}^n$. Then every $\mathbf{x} \in H$ can be written as

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_k\mathbf{b}_k$$

with scalars $c_1, \dots, c_k \in \mathbb{F}$ in exactly one way.

Proof. *At least one way.* The vectors $\mathbf{b}_1, \dots, \mathbf{b}_k$ span H , so every $\mathbf{x} \in H$ is a linear combination of them.

At most one way. Suppose $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k$ and also $\mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_k\mathbf{b}_k$. Subtracting the second equation from the first, and collecting terms with the rules of vector algebra (Theorem 1.2.7),

$$\mathbf{0} = (c_1 - d_1)\mathbf{b}_1 + (c_2 - d_2)\mathbf{b}_2 + \cdots + (c_k - d_k)\mathbf{b}_k.$$

The vectors $\mathbf{b}_1, \dots, \mathbf{b}_k$ are independent, so every weight is zero: $c_i - d_i = 0$, that is, $c_i = d_i$ for every i . The two ways are the same. $\square$

The converse holds too: if every $\mathbf{x} \in H$ is a combination of $\mathbf{b}_1, \dots, \mathbf{b}_k \in H$ in exactly one way, then $(\mathbf{b}_1, \dots, \mathbf{b}_k)$ is a basis of H . Indeed the vectors span H , and $\mathbf{0}$ has only one representation, the trivial one $0\mathbf{b}_1 + \cdots + 0\mathbf{b}_k$, which says that the vectors are independent.

The numbers $c_1, \dots, c_k$ are the *coordinates* of $\mathbf{x}$ in the basis $\mathcal{B}$. In the standard basis of $\mathbb{F}^n$, the coordinates of $\mathbf{x}$ are its entries (Example 4.2.2). Coordinate vectors are the subject of Section 5.1.

Example 4.2.5. In the basis $\mathbf{b}_1 = (1, 2)$, $\mathbf{b}_2 = (3, 1)$ of Example 4.2.3, find the coordinates of $\mathbf{x} = (5, 5)$. We need $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 = \mathbf{x}$, that is,

$$c_1 + 3c_2 = 5, \quad 2c_1 + c_2 = 5.$$

The second equation gives $c_2 = 5 - 2c_1$, and then the first gives $c_1 + 15 - 6c_1 = 5$, so $-5c_1 = -10$ and $c_1 = 2$, $c_2 = 5 - 4 = 1$. *Check:* $2(1, 2) + 1(3, 1) = (2 + 3, 4 + 1) = (5, 5)$. By Theorem 4.2.4, $(2, 1)$ is the only pair of coordinates: on the slanted grid, $\mathbf{x}$ sits at exactly one crossing (Figure 4.4).

4.2.3 Finding a basis: shrink or grow

There are two natural ways to reach a basis. Start with a list that spans, and throw out redundant vectors; or start with an independent list, and add vectors until it spans.

Theorem 4.2.6 (Shrinking a spanning list). Let $H = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be a subspace of $\mathbb{F}^n$ with $H \neq \{\mathbf{0}\}$. Then some of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ form a basis of H .

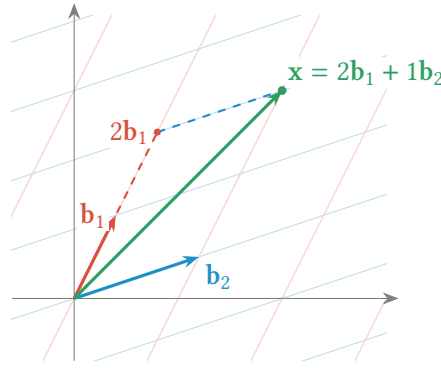


Figure 4.4: Examples 4.2.3 and 4.2.5: the basis $\mathbf{b}_1 = (1, 2)$ (coral), $\mathbf{b}_2 = (3, 1)$ (blue) and its slanted grid. The point $\mathbf{x} = (5, 5)$ (green) is reached in exactly one way, $\mathbf{x} = 2\mathbf{b}_1 + 1\mathbf{b}_2$.

Proof. If the list is independent, it is a basis of H . Otherwise it is dependent. It has at least two vectors: a list of one vector $\mathbf{v}_1$ that spans $H \neq \{\mathbf{0}\}$ has $\mathbf{v}_1 \neq \mathbf{0}$, and is then independent (Proposition 2.6.5(a)). So by Theorem 2.6.3 one of the vectors is a linear combination of the others, and by Theorem 1.3.11 removing it leaves the span equal to H . Repeat with the shorter list. Each step removes a vector and keeps the span equal to H , so the process stops after fewer than k steps, and it stops only at an independent list that spans H : a basis made of some of the $\mathbf{v}_i$. $\square$

Example 4.2.7. The floor H of the introduction is spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, which are dependent since $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$. In this relation every weight is nonzero, so any one of the three vectors is a combination of the other two and may be removed. Removing $\mathbf{v}_3$ gives the basis $(\mathbf{v}_1, \mathbf{v}_2)$; removing $\mathbf{v}_1$ gives the basis $(\mathbf{v}_2, \mathbf{v}_3)$, which spans H because $\mathbf{v}_1 = \mathbf{v}_3 - \mathbf{v}_2$, and is independent because neither vector is a multiple of the other. Different choices give different bases, each with two vectors (Figure 4.5).

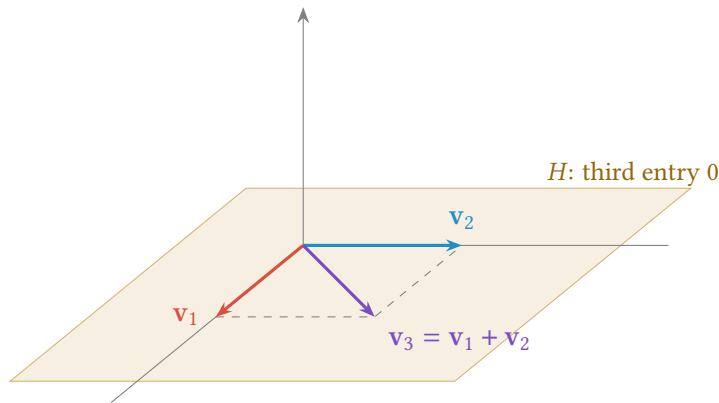


Figure 4.5: Example 4.2.7: the floor of $\mathbb{R}^3$ and three vectors in it, $\mathbf{v}_1$ (coral), $\mathbf{v}_2$ (blue) and $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$ (violet). Any two of them form a basis of the floor.

Theorem 4.2.8 (Growing an independent list). Let H be a subspace of $\mathbb{F}^n$.

- (a) Every linearly independent list $\mathbf{u}_1, \dots, \mathbf{u}_m$ of vectors in H can be extended to a basis $(\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{w}_1, \dots, \mathbf{w}_l)$ of H , with $l \geq 0$. Here $m \geq 0$: the empty list ($m = 0$) is independent, and the theorem extends it too.
- (b) H has a basis, and every basis of H has at most n vectors.

Proof. (a) Let L be the current list, starting with $L = (\mathbf{u}_1, \dots, \mathbf{u}_m)$. Its span lies in H , because H is a

subspace containing its vectors (Theorem 4.1.6(b)). If the span is all of H , then L is a basis of H and we stop. Otherwise choose $\mathbf{w} \in H$ that is not in the span of L , and add it at the end of L . The longer list is still independent, by Theorem 2.6.7: the condition of that theorem holds for the vectors of L , since L is independent, and it holds for the new last vector $\mathbf{w}$ by its choice. Every list we produce is an independent list in $\mathbb{F}^n$, so it has at most n vectors (Theorem 2.6.10). Hence the process stops after at most $n - m$ additions, and it stops only at a basis of H .

(b) If $H = \{\mathbf{0}\}$, the empty list is its basis. Otherwise pick $\mathbf{u} \in H$ with $\mathbf{u} \neq \mathbf{0}$; the list $(\mathbf{u})$ is independent (Proposition 2.6.5(a)), and (a) extends it to a basis. Every basis of H is an independent list in $\mathbb{F}^n$, so it has at most n vectors by Theorem 2.6.10. $\square$

4.2.4 Independent lists are never longer than spanning lists

The floor has bases $(\mathbf{v}_1, \mathbf{v}_2)$ and $(\mathbf{v}_2, \mathbf{v}_3)$, both with two vectors. The next theorem is the reason no basis of the floor can have three vectors, or one.

Theorem 4.2.9. Let H be a subspace of $\mathbb{F}^n$ spanned by k vectors $\mathbf{w}_1, \dots, \mathbf{w}_k$. Then every linearly independent list of vectors in H has at most k vectors. Equivalently, any $m > k$ vectors in H are linearly dependent.

Proof. Let $\mathbf{u}_1, \dots, \mathbf{u}_m \in H$ with $m > k$; we show that they are dependent. Each $\mathbf{u}_j$ lies in $H = \text{Span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$, so

$$\mathbf{u}_j = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \cdots + a_{kj}\mathbf{w}_k \quad (j = 1, \dots, m)$$

for some scalars a_{ij} . Let W be the $n \times k$ matrix with columns $\mathbf{w}_1, \dots, \mathbf{w}_k$, let M be the $n \times m$ matrix with columns $\mathbf{u}_1, \dots, \mathbf{u}_m$, and let $C = [a_{ij}]$ be the $k \times m$ matrix of the weights, whose column j is $\mathbf{c}_j = (a_{1j}, \dots, a_{kj})$. By the column rule (Definition 3.1.1), $W\mathbf{c}_j = \mathbf{u}_j$, and column j of WC is $W\mathbf{c}_j$ (Section 3.3). So

$$M = WC.$$

The homogeneous system $C\mathbf{x} = \mathbf{0}$ has m unknowns and only $k < m$ equations, so it has a nontrivial solution $\mathbf{x} = (x_1, \dots, x_m) \neq \mathbf{0}$ (Theorem 2.5.9). Then

$$M\mathbf{x} = (WC)\mathbf{x} = W(C\mathbf{x}) = W\mathbf{0} = \mathbf{0},$$

and by the column rule $M\mathbf{x} = x_1\mathbf{u}_1 + \cdots + x_m\mathbf{u}_m$. So $x_1\mathbf{u}_1 + \cdots + x_m\mathbf{u}_m = \mathbf{0}$ with weights not all zero: the vectors $\mathbf{u}_1, \dots, \mathbf{u}_m$ are dependent. $\square$

The proof hides an idea that is worth seeing on its own. Each independent vector can take the place of one spanning vector, and the list still spans; after m such exchanges, m spanning vectors have been used up, so $m \leq k$. This is the *exchange idea* of Steinitz. The example below shows it in $\mathbb{R}^2$, and the supplement at the end of the section proves it in general. Section 4.5 (Lecture 23) uses it again for vector spaces that are not inside any $\mathbb{F}^n$.

Example 4.2.10 (The exchange idea). In $\mathbb{R}^2$, start with the spanning list $(\mathbf{e}_1, \mathbf{e}_2)$ and the independent vectors $\mathbf{u}_1 = (1, 2)$, $\mathbf{u}_2 = (3, 1)$. First, $\mathbf{u}_1 = 1\mathbf{e}_1 + 2\mathbf{e}_2$. The weight of $\mathbf{e}_1$ is not zero, so we can solve for it: $\mathbf{e}_1 = \mathbf{u}_1 - 2\mathbf{e}_2$. Hence $\text{Span}\{\mathbf{u}_1, \mathbf{e}_2\}$ contains $\mathbf{e}_1$ and $\mathbf{e}_2$, so it contains every combination of them and equals $\mathbb{R}^2$: the list $(\mathbf{u}_1, \mathbf{e}_2)$ still spans. Next, write $\mathbf{u}_2$ in this new list:

$$\mathbf{u}_2 = 3\mathbf{e}_1 + \mathbf{e}_2 = 3(\mathbf{u}_1 - 2\mathbf{e}_2) + \mathbf{e}_2 = 3\mathbf{u}_1 - 5\mathbf{e}_2 \quad (\text{check: } (3, 6) - (0, 5) = (3, 1)).$$

The weight of $\mathbf{e}_2$ is $-5 \neq 0$, so $\mathbf{e}_2 = \frac{3}{5}\mathbf{u}_1 - \frac{1}{5}\mathbf{u}_2$, and $(\mathbf{u}_1, \mathbf{u}_2)$ spans $\mathbb{R}^2$. Each new independent vector pushed out one old spanning vector. A weight on an old vector must be nonzero at each step: if all of them were zero, the new $\mathbf{u}$ would be a combination of the $\mathbf{u}$'s already in the list, contradicting independence.

4.2.5 Dimension

Theorem 4.2.11 (Invariance of dimension). Let H be a subspace of $\mathbb{F}^n$. Any two bases of H have the same number of vectors.

Proof. Let $(\mathbf{b}_1, \dots, \mathbf{b}_k)$ and $(\mathbf{c}_1, \dots, \mathbf{c}_m)$ be bases of H . The $\mathbf{b}$'s span H and the $\mathbf{c}$'s are independent vectors in H , so $m \leq k$ by Theorem 4.2.9. Exchanging the roles of the two bases gives $k \leq m$. Hence $k = m$. (If $H = \{\mathbf{0}\}$, its only basis is the empty list, since every nonempty list in $\{\mathbf{0}\}$ contains $\mathbf{0}$ and is dependent.) $\square$

Definition 4.2.12 (Dimension). Let H be a subspace of $\mathbb{F}^n$. The *dimension* of H , written $\dim H$, is the number of vectors in any basis of H . In particular $\dim\{\mathbf{0}\} = 0$.

Every subspace of $\mathbb{F}^n$ has a basis (Theorem 4.2.8(b)), and all its bases have the same length (Theorem 4.2.11), so $\dim H$ is a well-defined number, with $0 \leq \dim H \leq n$. Dimension counts the independent directions in H : the number of parameters needed to describe its vectors.

The dimension is also the least number of vectors that span H . A basis spans H with $\dim H$ vectors. Conversely, if k vectors span H , then the vectors of a basis form an independent list of $\dim H$ vectors in H , so $\dim H \leq k$ by Theorem 4.2.9. In particular every subspace of $\mathbb{F}^n$ is the span of at most n of its vectors.

Example 4.2.13 (Some dimensions). We read off a basis in each case and count its vectors.

- (a) $\dim \mathbb{F}^n = n$, since the standard basis has n vectors (Example 4.2.2). So $\dim \mathbb{R}^2 = 2$ and $\dim \mathbb{R}^3 = 3$.
- (b) A line through the origin, $L = \text{Span}\{\mathbf{v}\}$ with $\mathbf{v} \neq \mathbf{0}$, has the basis $(\mathbf{v})$, so $\dim L = 1$.
- (c) The floor of $\mathbb{R}^3$ has the basis $(\mathbf{v}_1, \mathbf{v}_2)$, so its dimension is 2. More generally, a plane $\text{Span}\{\mathbf{v}, \mathbf{w}\}$ with $\mathbf{v}, \mathbf{w}$ independent has dimension 2.

So in $\mathbb{R}^3$ the subspaces are, by dimension: $\{\mathbf{0}\}$ (dimension 0), lines through the origin (1), planes through the origin (2), and $\mathbb{R}^3$ itself (3). That the list is complete follows from Theorem 4.2.16(c) below: a subspace of dimension 1 or 2 is spanned by a basis of one or two vectors, and one of dimension 3 is all of $\mathbb{R}^3$.

4.2.6 Bases of the column space and the null space

For the subspaces that come from a matrix, a basis can be read off the reduced row echelon form.

Theorem 4.2.14. Let A be an $m \times n$ matrix over $\mathbb{F}$ with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and let R be its reduced row echelon form, with r pivots.

- (a) The pivot columns of A form a basis of $\text{Col } A$. Hence $\dim \text{Col } A = r$.
- (b) Write the general solution of $A\mathbf{x} = \mathbf{0}$ in parametric vector form, $\mathbf{x} = t_1\mathbf{s}_1 + \dots + t_f\mathbf{s}_f$, with

one parameter for each free variable (Definition 2.5.2). Then $(\mathbf{s}_1, \dots, \mathbf{s}_f)$ is a basis of $\text{Nul } A$. Hence $\dim \text{Nul } A = f$, the number of free variables.

(c) $\dim \text{Col } A + \dim \text{Nul } A = n$.

Proof. (a) Let the pivot columns be $j_1 < \dots < j_r$, and write $\mathbf{r}_1, \dots, \mathbf{r}_n$ for the columns of R . Since A and R are row equivalent, they have the same column relations (Proposition 2.4.11).

Independent. Suppose $c_1 \mathbf{a}_{j_1} + \dots + c_r \mathbf{a}_{j_r} = \mathbf{0}$. Giving every other column the weight 0, this is a column relation of A , hence of R : $c_1 \mathbf{r}_{j_1} + \dots + c_r \mathbf{r}_{j_r} = \mathbf{0}$. By Proposition 2.4.12(a), $\mathbf{r}_{j_i} = \mathbf{e}_i$, so the left side is $(c_1, \dots, c_r, 0, \dots, 0)$, and every $c_i = 0$.

Spanning. Let column k be a non-pivot column. By Proposition 2.4.12(c), $\mathbf{r}_k - \sum_{i:j_i < k} r_{ik} \mathbf{r}_{j_i} = \mathbf{0}$, a column relation of R , hence of A : $\mathbf{a}_k = \sum_{i:j_i < k} r_{ik} \mathbf{a}_{j_i}$. So every column of A lies in the span of the pivot columns, which is a subspace (Theorem 4.1.6(a)). By Theorem 4.1.6(b) it then contains the span of all the columns: $\text{Col } A$ is contained in the span of the pivot columns. The reverse inclusion holds because the pivot columns are columns of A . So the pivot columns span $\text{Col } A$, and they form a basis.

(b) Let the free variables be $x_{k_1}, \dots, x_{k_f}$, with parameters $t_1, \dots, t_f$. By the construction after Definition 2.5.2, the vector $\mathbf{s}_q$ has entry 1 in position k_q and entry 0 in every other free position. *Spanning:* the $\mathbf{s}_q$ span $\text{Nul } A$ by Proposition 4.1.15 (every solution of $A\mathbf{x} = \mathbf{0}$ is $t_1 \mathbf{s}_1 + \dots + t_f \mathbf{s}_f$, and each $\mathbf{s}_q$ is the solution with $t_q = 1$ and the other parameters 0). *Independent:* if $t_1 \mathbf{s}_1 + \dots + t_f \mathbf{s}_f = \mathbf{0}$, then the entry in position k_q of the left side is $t_q \cdot 1 + \sum_{p \neq q} t_p \cdot 0 = t_q$, so every $t_q = 0$. If there are no free variables, then $\text{Nul } A = \{\mathbf{0}\}$ and the empty list is its basis, in agreement with $f = 0$.

(c) Every column of A is either a pivot column or belongs to a free variable, so $r + f = n$. $\square$

A warning: the basis of $\text{Col } A$ consists of the pivot columns of A , not of R . Row operations keep the column relations but change the columns themselves, and usually $\text{Col } R \neq \text{Col } A$. In Example 4.2.15 every column of R has third entry 0, while the columns of A do not.

Example 4.2.15. Take the coefficient matrix of Example 2.5.1 and its reduced row echelon form:

$$A = \begin{bmatrix} 1 & 2 & -1 & -1 \\ 2 & 4 & -1 & 0 \\ -1 & -2 & 2 & 3 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The pivots are in columns 1 and 3. By Theorem 4.2.14(a), a basis of $\text{Col } A$ is

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix},$$

and $\dim \text{Col } A = 2$: the column space is a plane in $\mathbb{R}^3$. The relations of R , $\mathbf{r}_2 = 2\mathbf{r}_1$ and $\mathbf{r}_4 = \mathbf{r}_1 + 2\mathbf{r}_3$, hold in A too: $\mathbf{a}_2 = (2, 4, -2) = 2\mathbf{a}_1$, and $\mathbf{a}_1 + 2\mathbf{a}_3 = (1 - 2, 2 - 2, -1 + 4) = (-1, 0, 3) = \mathbf{a}_4$.

For the null space, the free variables are $x_2 = s$ and $x_4 = t$, and R gives $x_1 = -2s - t$ and $x_3 = -2t$. So

$$\mathbf{x} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ -2 \\ 1 \end{bmatrix} = s \mathbf{s}_1 + t \mathbf{s}_2,$$

the vectors $\mathbf{v}$ and $\mathbf{w}$ of Example 2.5.1. By Theorem 4.2.14(b), $(\mathbf{s}_1, \mathbf{s}_2)$ is a basis of $\text{Nul } A$, and $\dim \text{Nul } A = 2$: a plane in $\mathbb{R}^4$. *Check:* $A\mathbf{s}_1 = -2\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{0}$, and $A\mathbf{s}_2 = -\mathbf{a}_1 - 2\mathbf{a}_3 + \mathbf{a}_4 = (-1 + 2 - 1, -2 + 2 + 0, 1 - 4 + 3) = \mathbf{0}$. Finally $2 + 2 = 4$, the number of columns, as part (c) says.

Part (c) is the first form of the *rank–nullity theorem*. *Preview.* For an $m \times n$ matrix A over a field $\mathbb{F}$, the number $\dim \operatorname{Col} A$ is called the *rank* of A (Section 5.2). In Section 6.2 (Lecture 30), part (c) becomes a statement about a linear map $T: V \rightarrow W$ between vector spaces over $\mathbb{F}$, with V finite-dimensional: $\dim V$ equals the dimension of the image of T plus the dimension of its kernel.

4.2.7 When the count is right

If we already know the dimension, half of the work of checking a basis comes free.

Theorem 4.2.16. Let H be a subspace of $\mathbb{F}^n$ with $\dim H = k$.

- (a) Any k linearly independent vectors in H form a basis of H .
- (b) If $k \geq 1$, any k vectors that span H form a basis of H .
- (c) If U is a subspace of $\mathbb{F}^n$ with $U \subseteq H$, then $\dim U \leq \dim H$, and $\dim U = \dim H$ only if $U = H$.

Proof. (a) By Theorem 4.2.8(a) the k independent vectors can be extended to a basis of H . That basis has $\dim H = k$ vectors (Theorem 4.2.11), so nothing was added, and the k vectors are already a basis.

(b) Since $k \geq 1$, $H \neq \{0\}$, so by Theorem 4.2.6 some of the k vectors form a basis of H . That basis has k vectors, so none was removed, and the k vectors are a basis.

(c) Let $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ be a basis of U , so $p = \dim U$. It is an independent list in H , so by Theorem 4.2.8(a) it extends to a basis of H , which has $\dim H$ vectors; hence $p \leq \dim H$. If $p = \dim H$, nothing was added, so $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ is a basis of H and $H = \operatorname{Span}\{\mathbf{u}_1, \dots, \mathbf{u}_p\} = U$. (If $U = \{0\}$, its basis is the empty list, $p = 0$, and (a) applies with $m = 0$: the process in its proof starts from the empty list.) $\square$

Example 4.2.17. The vectors $(1, 2, 3)$, $(4, 5, 6)$, $(7, 8, 10)$ are independent (Example 2.6.9). They are three independent vectors in $\mathbb{R}^3$, and $\dim \mathbb{R}^3 = 3$, so by Theorem 4.2.16(a) they form a basis of $\mathbb{R}^3$, with no need to check that they span. By contrast, $(1, 2, 3)$, $(4, 5, 6)$, $(7, 8, 9)$ are dependent, so they are not a basis; their span is the column space of the matrix in Example 2.6.9, which has pivots in columns 1 and 2, so it is the plane with basis $(1, 2, 3)$, $(4, 5, 6)$ (Theorem 4.2.14(a)).

For n vectors in $\mathbb{F}^n$, Theorem 4.2.16 says: they form a basis of $\mathbb{F}^n$ exactly when they are independent, and exactly when they span. With Proposition 3.1.13 and Theorem 3.1.9, both conditions say that the square matrix with these columns has a pivot in every row and every column. For a square matrix, a pivot in every column means that its reduced row echelon form is the identity, so by Theorem 3.5.10 the n vectors form a basis of $\mathbb{F}^n$ exactly when the matrix with these columns is invertible.

Check yourself

- (a) Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \end{bmatrix}$. Find a basis of $\operatorname{Col} A$ and a basis of $\operatorname{Nul} A$, and check that the dimensions add up to 3.
- (b) Is $((1, 1, 0), (0, 1, 1), (1, 0, 1))$ a basis of $\mathbb{R}^3$? Is it a basis of $\mathbb{Z}_2^3$? (Use Example 2.6.11.)
- (c) True or false: if H is a subspace of $\mathbb{R}^4$ and $\dim H = 4$, then $H = \mathbb{R}^4$. Justify your answer.

Supplement: beyond the lecture

This material is not needed later in this chapter, but it rounds out the picture.

The exchange lemma. Example 4.2.10 works in general.

Steinitz exchange lemma. Let $\mathbf{u}_1, \dots, \mathbf{u}_m$ be linearly independent vectors in $\text{Span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\} \subseteq \mathbb{F}^n$. Then $m \leq k$, and after renumbering the $\mathbf{w}$'s, the list $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{w}_{m+1}, \dots, \mathbf{w}_k$ spans the same subspace as $\mathbf{w}_1, \dots, \mathbf{w}_k$.

Proof. Write $S = \text{Span}\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$. We show by induction on $j = 0, 1, \dots, m$ that $j \leq k$ and that, after renumbering, $\mathbf{u}_1, \dots, \mathbf{u}_j, \mathbf{w}_{j+1}, \dots, \mathbf{w}_k$ spans S . For $j = 0$ this is the assumption. Suppose it holds for some $j < m$. Since $\mathbf{u}_{j+1} \in S$,

$$\mathbf{u}_{j+1} = c_1\mathbf{u}_1 + \dots + c_j\mathbf{u}_j + d_{j+1}\mathbf{w}_{j+1} + \dots + d_k\mathbf{w}_k$$

for some scalars. If $j = k$ (no $\mathbf{w}$'s left) or every $d_i = 0$, this writes $\mathbf{u}_{j+1}$ as a combination of $\mathbf{u}_1, \dots, \mathbf{u}_j$, which is impossible for an independent list (Theorem 2.6.7). So $j + 1 \leq k$ and some $d_i \neq 0$; renumber so that $d_{j+1} \neq 0$. Solving for $\mathbf{w}_{j+1}$,

$$\mathbf{w}_{j+1} = d_{j+1}^{-1}(\mathbf{u}_{j+1} - c_1\mathbf{u}_1 - \dots - c_j\mathbf{u}_j - d_{j+2}\mathbf{w}_{j+2} - \dots - d_k\mathbf{w}_k),$$

so $\mathbf{w}_{j+1}$ lies in the span S' of $\mathbf{u}_1, \dots, \mathbf{u}_{j+1}, \mathbf{w}_{j+2}, \dots, \mathbf{w}_k$. Hence S' contains every vector of the spanning list $\mathbf{u}_1, \dots, \mathbf{u}_j, \mathbf{w}_{j+1}, \dots, \mathbf{w}_k$ of S , and so $S \subseteq S'$ (Theorem 4.1.6(b)); and $S' \subseteq S$ because all its vectors lie in the subspace S . So $S' = S$, which is the statement for $j + 1$. For $j = m$ we obtain $m \leq k$ and the spanning list. $\square$

This gives a second proof of Theorem 4.2.9, without matrices, and it is the form of the argument that Section 4.5 uses for general vector spaces.

Counting over a finite field. Over $\mathbb{Z}_p$, a subspace H of $\mathbb{Z}_p^n$ with $\dim H = k$ has exactly p^k vectors. Indeed, fix a basis $(\mathbf{b}_1, \dots, \mathbf{b}_k)$ of H . By Theorem 4.2.4, the map $(c_1, \dots, c_k) \mapsto c_1\mathbf{b}_1 + \dots + c_k\mathbf{b}_k$ from lists of k scalars to H hits every vector of H exactly once, and there are p^k lists of k elements of $\mathbb{Z}_p$. For example, in $\mathbb{Z}_2^3$ the vectors $(1, 1, 0)$ and $(0, 1, 1)$ are independent (neither is a multiple of the other: the only multiples of a vector over $\mathbb{Z}_2$ are $\mathbf{0}$ and the vector itself), so they span a subspace of dimension 2 with $2^2 = 4$ vectors:

$$\mathbf{0}, \quad (1, 1, 0), \quad (0, 1, 1), \quad (1, 1, 0) + (0, 1, 1) = (1, 0, 1).$$

In particular the dimension of a subspace of $\mathbb{Z}_p^n$ can be found by counting its vectors.