Linear Algebra 1: Lecture 7

Irena Penev

Winter 2022/2023

1 Bases of vector spaces

Given a vector space V over a field $\mathbb{F}$, and given vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$, we say that $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a *linearly independent set*, or that vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are *linearly independent*, if for all $\alpha_1, \dots, \alpha_k$ such that

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k \quad = \quad \mathbf{0},$$

we have that $\alpha_1 = \cdots = \alpha_k = \mathbf{0}$. Otherwise, we say that $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a *linearly dependent set*, or that vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly dependent.

We note that $\emptyset$ is linearly independent in any vector space.

A *finite basis* (or simply *basis*) of a vector space V over a field $\mathbb{F}$ is a set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of vectors in V that satisfies the following two conditions:

1. $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is linearly independent in V;
2. $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a spanning set of V, i.e. $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = V$.

A vector space is *finite-dimensional* if it has a finite basis. A vector space that does not have a finite basis is *infinite-dimensional*.

Not all vector spaces have a finite basis (we shall see some examples of this later). It is, indeed, possible to define a basis more generally, so that it may possibly be an infinite set. This is briefly discussed in subsection 1.1 below. However, in this course, we will only study finite bases. (We will occasionally deal with infinite-dimensional vector spaces, but we will not deal with their bases.)

Remarks: Suppose that V is a vector space over a field $\mathbb{F}$.

- Obviously, any subset of a linearly independent set of vectors in V is linearly independent. Similarly, any superset of a spanning set of V is a spanning set of V.[1]

[1] A set A is a *superset* of a set B provided that $B \subseteq A$.

- $\{\mathbf{0}\}$ is **not** a linearly independent set in V (because $1 \cdot \mathbf{0} = \mathbf{0}$ and $1 \neq 0$), and so by the previous bullet point, no linearly independent set of vectors in V, and in particular, no basis of V, contains the zero vector.

- $\emptyset$ is a basis of the trivial vector space $\{\mathbf{0}\}$ (over any field $\mathbb{F}$), and in particular, $\{\mathbf{0}\}$ is finite dimensional. In fact, $\emptyset$ is the unique basis of $\{\mathbf{0}\}$ (because, by the previous bullet point, no linearly independent set contains $\mathbf{0}$).

- Suppose we are given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, and we are trying to check if $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a spanning set of V, i.e. whether $V = \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ (this is one of the two conditions from the definition of a basis). Obviously, $\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \subseteq V$, and so the only question is whether $V \subseteq \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. But "$V \subseteq \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$" simply means "every vector in $\mathbf{v}$ is a linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$." So, the second condition from the definition of a basis holds if and only if every vector in V is a linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

- In the literature, there is a bit ambiguity about whether (finite) bases are sets or **ordered** sets. An "ordered set" is a set in which order and repetitions matter. For instance, $\{1, 2, 3\}$, $\{1, 2, 2, 3\}$, and $\{3, 1, 2\}$ are the same as sets, but they are pairwise distinct as ordered sets. In what follows, we will implicitly treat finite sets (when discussed in the context of linearly independent sets, spanning sets, and bases) as ordered, and in particular, we will care about repetitions. It is important to note that no linearly independent set (and in particular, no basis), may contain more than one copy of the same vector. Indeed, if $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a list of vectors that contains more than one copy of some vector (say, $\mathbf{v}_i = \mathbf{v}_j$ for some $i \neq j$), then we can set $\alpha_i = 1$, $\alpha_j = -1$, and $\alpha_k = 0$ for all $k \in \{1, \ldots, n\} \setminus \{i, j\}$, and we get $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$; so $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are not linearly independent.

  - In what follows, if A and B are ordered sets (possibly with repeating elements), then $A \subseteq B$ means that every element of A appears at least as many times in B as in A. Moreover, for $x \in A$, $A \setminus \{x\}$ is the set obtained from A by deleting one copy of x.

Example 1.1. *Let $\mathbb{F}$ be a field, and let n be a positive integer. For each $i \in \{1, \ldots, n\}$, let $\mathbf{e}_i^n$ to be the vector in $\mathbb{F}^n$ whose i-th entry is 1, and all of whose other entries are zero. Then $\{\mathbf{e}_1^n, \mathbf{e}_2^n, \ldots, \mathbf{e}_n^n\}$ is a basis of $\mathbb{F}^n$,[2] and it is called the* standard basis *of $\mathbb{F}^n$.*

[2]Let us check this! We first show that $\{\mathbf{e}_1^n, \ldots, \mathbf{e}_n^n\}$ is linearly independent. Fix scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\alpha_1\mathbf{e}_1^n + \cdots + \alpha_n\mathbf{e}_n^n = \mathbf{0}$. Clearly, $\alpha_1\mathbf{e}_1^n + \cdots + \alpha_n\mathbf{e}_n^n = \begin{bmatrix} \alpha_1 & \ldots & \alpha_n \end{bmatrix}^T$. So, $\begin{bmatrix} \alpha_1 & \ldots & \alpha_n \end{bmatrix}^T = \mathbf{0}$, and it follows that $\alpha_1 = \cdots = \alpha_n = 0$. So, $\{\mathbf{e}_1^n, \ldots, \mathbf{e}_n^n\}$ is linearly independent. Let us now show that $Span(\mathbf{e}_1^n, \ldots, \mathbf{e}_n^n) = \mathbb{F}^n$, i.e.

Example 1.2. *Let $\mathbb{F}$ be a field, and let n be a positive integer. Consider the set $\mathbb{P}^n_{\mathbb{F}}$ of all polynomials (in variable x) of degree at most n, with coefficients in $\mathbb{F}$. (Then $\mathbb{P}^n_{\mathbb{F}}$ is a vector space over the field $\mathbb{F}$.) Then $\{1, x, \ldots, x^n\}$ is a basis of $\mathbb{P}^n_{\mathbb{F}}$.*[3]

Theorem 1.3. *Let V be a vector space over a field $\mathbb{F}$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$. Then the following are equivalent:*

(i) $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V;

(ii) for all $\mathbf{v} \in V$, there exist unique scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$.

Proof. Suppose first that (i) holds; we must show that (ii) holds. Fix $\mathbf{v} \in V$. We must show that there exist unique scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V, we know that every vector in V is a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. This proves existence. It remains to prove uniqueness. Fix scalars $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$ and $\mathbf{v} = \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n$. Then

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n \quad = \quad \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n,$$

and consequently,

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n \quad = \quad \mathbf{0}.$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent (because it is a basis of V), we deduce that $\alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0$. So, $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$. This proves uniqueness, and (ii) follows.

Suppose now that (ii) holds; we must show that (i) holds. By (ii), every vector in V is a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and so $V = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. It remains to show that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent. Clearly, $0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n = \mathbf{0}$. By the uniqueness part of (ii) applied to $\mathbf{v} := \mathbf{0}$, it follows that for all scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$, if $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$, then $\alpha_1 = \cdots = \alpha_n = 0$. Thus, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent, and (i) follows. □

---

that every vector in V is a linear combination of the vectors $\mathbf{e}^n_1, \ldots, \mathbf{e}^n_n$. Fix any $\mathbf{v} \in V$, and set $\mathbf{v} = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}^T$. But now $\mathbf{v} = v_1\mathbf{e}^n_1 + \cdots + v_n\mathbf{e}^n_n$, i.e. every vector in $\mathbb{F}^n$ is a linear combination of the vectors $\mathbf{e}^n_1, \ldots, \mathbf{e}^n_n$. So, $\{\mathbf{e}^n_1, \ldots, \mathbf{e}^n_n\}$ is indeed a basis of $\mathbb{F}^n$.

[3]This is immediate from the relevant definitions. Indeed, to show that $\{1, x, \ldots, x^n\}$ is linearly independent, we fix $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\alpha_0 \cdot 1 + \alpha_1 x + \cdots + \alpha_n x^n = 0$. Then clearly, $\alpha_0 = \alpha_1 = \cdots = \alpha_0 = 0$. So, $\{1, x, \ldots, x^n\}$ is linearly independent. On the other hand, by definition, for every $p(x) \in \mathbb{P}^n_{\mathbb{F}}$, there exist $a_0, a_1, \ldots, a_n \in \mathbb{F}$ such that $p(x) = a_0 + a_1 x + \cdots + a_n x^n$. But then $p(x) = a_0 \cdot 1 + a_1 x + \cdots + a_n x^n$, and so $p(x)$ is a linear combination of $1, x, \ldots, x^n$. Thus, $\{1, x, \ldots, x^n\}$ is indeed a basis of $\mathbb{P}^n_{\mathbb{F}}$.

Remark: Theorem 1.3 is one of the main reasons why we care about bases. Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of a vector space V over a field $\mathbb{F}$. Then by Theorem 1.3, every vector $\mathbf{v}$ can be associated to a unique vector $\begin{bmatrix} \alpha_1 & \ldots & \alpha_n \end{bmatrix}^T$ in $\mathbb{F}^n$ such that $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$. So, V is in some sense "equivalent" to $\mathbb{F}^n$. The technical word here is "isomorphic": V is "isomorphic" to $\mathbb{F}^n$. We will discuss this more formally in a subsequent lecture.

Proposition 1.4. *Let V be a vector space over a field $\mathbb{F}$, and let $\mathbf{a}_1, \ldots, \mathbf{a}_k \in V$. Set $A := \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$. Then the following hold:*

(a) A is linearly independent if and only if no vector in A is a linear combination of the other vectors in A;[4]

(b) if A is a spanning set of V, and some vector $\mathbf{a}_i \in A$ is a linear combination of the other vectors in A, then $A \setminus \{\mathbf{a}_i\}$ is a spanning set of V.[5]

Proof. We first prove (a). We prove the following equivalent statement: A is linearly dependent if and only if some vector of A is a linear combination of the other vectors in A.

Suppose first that A is linearly dependent. Then there exist scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$, not all zero, such that $\alpha_1\mathbf{a}_1 + \cdots + \alpha_k\mathbf{a}_k = \mathbf{0}$. Fix an index $i \in \{1, \ldots, k\}$ such that $\alpha_i \neq 0$. Then

$$\mathbf{a}_i \;=\; -\alpha_i^{-1}\alpha_1\mathbf{a}_1 - \cdots - \alpha_i^{-1}\alpha_{i-1}\mathbf{a}_{i-1} - \alpha_i^{-1}\alpha_{i+1}\mathbf{a}_{i+1} - \cdots - -\alpha_k^{-1}\alpha_1\mathbf{a}_k,$$

and so $\mathbf{a}_i$ is a linear combination of the other vectors in A.

Suppose now that some vector in A is a linear combination of the other vectors in A. Say, $\mathbf{a}_i$ is a linear combination of the vectors $\mathbf{a}_1, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_k$. Then there exist scalars $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_k \in \mathbb{F}$ such that

$$\mathbf{a}_i \;=\; \alpha_1\mathbf{a}_1 + \cdots + \alpha_{i-1}\mathbf{a}_{i-1} + \alpha_{i+1}\mathbf{a}_{i+1} + \cdots + \alpha_k\mathbf{a}_k.$$

We now set $\alpha_i = -1$, and we observe that

$$\alpha_1\mathbf{a}_1 + \cdots + \alpha_{i-1}\mathbf{a}_{i-1} + \alpha_i\mathbf{a}_i + \alpha_{i+1}\mathbf{a}_{i+1} + \cdots + \alpha_k\mathbf{a}_k \;=\; \mathbf{0}.$$

Since not all of $\alpha_1, \ldots, \alpha_k$ are zero (indeed, $\alpha_i \neq 0$), we see that $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is linearly dependent. This proves (a).

[4] If A contains more than one copy of the same vector, then we treat each copy as distinct. So, when expressing a vector $\mathbf{v}$ in A as a linear combination of the "other" vectors in A, we are allowed to use any additional copies of $\mathbf{v}$ (if there are any) in that linear combination.

[5] If $\mathbf{a}_i$ appears more than once in A, then $A \setminus \{\mathbf{a}_i\}$ is understood to be the set obtained from A by removing only one copy of $\mathbf{a}_i$.

We now prove (b). Assume that some $\mathbf{a}_i \in A$ is a linear combination of the other vectors in A. Then there exist scalars $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_k \in \mathbb{F}$ such that

$$\mathbf{a}_i = \alpha_1\mathbf{a}_1 + \cdots + \alpha_{i-1}\mathbf{a}_{i-1} + \alpha_{i+1}\mathbf{a}_{i+1} + \cdots + \alpha_k\mathbf{a}_k.$$

Now, fix any vector $\mathbf{v} \in V$. We must show that $\mathbf{v}$ is a linear combination of vectors in $A \setminus \{\mathbf{a}_i\} = \{\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_k\}$. Since $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is a spanning set of V, we know that there exist scalars $\beta_1, \ldots, \beta_k \in \mathbb{F}$ such that $\mathbf{v} = \beta_1\mathbf{a}_1 + \cdots + \beta_k\mathbf{a}_k$. We now compute:

$$\begin{aligned}
\mathbf{v} &= \beta_1\mathbf{a}_1 + \cdots + \beta_{i-1}\mathbf{a}_{i-1} + \beta_i\mathbf{a}_i + \beta_{i+1}\mathbf{a}_{i+1} + \beta_k\mathbf{a}_k \\
&= \beta_1\mathbf{a}_1 + \cdots + \beta_{i-1}\mathbf{a}_{i-1} + \\
&\quad + \beta_i(\alpha_1\mathbf{a}_1 + \cdots + \alpha_{i-1}\mathbf{a}_{i-1} + \alpha_{i+1}\mathbf{a}_{i+1} + \cdots + \alpha_k\mathbf{a}_k) \\
&\quad + \beta_{i+1}\mathbf{a}_{i+1} + \beta_k\mathbf{a}_k \\
&= (\beta_1 + \beta_i\alpha_1)\mathbf{a}_1 + \cdots + (\beta_{i-1} + \beta_i\alpha_{i-1})\mathbf{a}_{i-1} + \\
&\quad + (\beta_{i+1} + \beta_i\alpha_{i+1})\mathbf{a}_{i+1} + \cdots + (\beta_k + \beta_i\alpha_k)\mathbf{a}_k.
\end{aligned}$$

So, $\mathbf{v}$ is a linear combination of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_k$, and (b) follows. □

Proposition 1.5. *Let V be a vector space over a field $\mathbb{F}$, and let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ be a spanning set of V. Let $B' \subseteq B$ be such that every vector in B is a linear combination of vectors in B'. Then B' is a spanning set of V.*

Proof. Choose a set $\widetilde{B}$ such that

- $B' \subseteq \widetilde{B} \subseteq B$,
- B' is a spanning set of V;
- subject to the above, $\widetilde{B}$ is as small as possible.

(The fact that $\widetilde{B}$ exists follows from the fact that $B' \subseteq B \subseteq B$, and B is a spanning set of V.) If $\widetilde{B} = B'$, then we are done. So, assume that $B' \subsetneq \widetilde{B}$, and fix some $\mathbf{v} \in \widetilde{B} \setminus B'$. Then $\mathbf{v}$ is a linear combination of the other vectors in $\widetilde{B}$ (because $\mathbf{v}$ is a linear combination of the vectors in B'), and so by Proposition 1.4(b), $\widetilde{B} \setminus \{\mathbf{v}\}$ is a spanning set of V. But now $\widetilde{B} \setminus \mathbf{v}$ contradicts the minimality of $\widetilde{B}$. □

Our next proposition (Proposition 1.6) states that, given any spanning set B of a vector space V over a field $\mathbb{F}$, we can obtain a basis of V by possibly removing some vectors from B. As we shall see later (see Proposition 1.10), any linearly independent set in a finite-dimensional vector space can be extended to a basis; however, we cannot prove this yet.

Proposition 1.6. *Let V be a vector space over a field $\mathbb{F}$, and let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ be a spanning set of V. Then some subset of B is a basis of V.*

Proof. Let $B' \subseteq B$ be a spanning set of V that has as few elements as possible.[6] We claim that B' is a basis of V. It suffices to show that B is linearly independent. Suppose otherwise. Then Proposition 1.4(a) guarantees that some $\mathbf{b} \in B'$ is a linear combination of the other vectors in B'; but then by Proposition 1.4(b), $B' \setminus \{\mathbf{b}\}$ is a spanning set of V, contrary to the minimality of B'. □

Lemma 1.7. *Let V be a vector space over a field $\mathbb{F}$. Let $\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b}_1, \ldots, \mathbf{b}_\ell \in V$, and assume that $\mathbf{a}_1, \ldots, \mathbf{a}_k$ are pairwise distinct and $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$ are pairwise distinct. Assume that $A := \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is a linearly independent set in V, and that $B := \{\mathbf{b}_1, \ldots, \mathbf{b}_\ell\}$ is a spanning set of V. Then for all $\mathbf{a} \in A \setminus B$, there exists some $\mathbf{b} \in B \setminus A$ such that $(B \setminus \{\mathbf{b}\}) \cup \{\mathbf{a}\}$ is a spanning set of V.*[7]

Proof. We may assume that $A \not\subseteq B$, for otherwise, the lemma is vacuously true. Fix any $\mathbf{a} \in A \setminus B$. Then there exists an index $i \in \{1, \ldots, k\}$ such that $\mathbf{a} = \mathbf{a}_i$. Since $\mathbf{a}_i \in V = \mathrm{Span}(B)$, we know that there exist scalars $\alpha_1, \ldots, \alpha_\ell$ such that

$$\mathbf{a}_i \;=\; \alpha_1\mathbf{b}_1 + \cdots + \alpha_\ell\mathbf{b}_\ell.$$

Since A is linearly independent, it does not contain the zero vector; since $\mathbf{a}_i \in A$, it follows that $\mathbf{a}_i \neq \mathbf{0}$. So, at least one of the scalars $\alpha_1, \ldots, \alpha_\ell$ is nonzero. If for all $j \in \{1, \ldots, \ell\}$ such that $\alpha_j \neq 0$, we have that $\mathbf{b}_j \in A \setminus \{\mathbf{a}_i\}$, then $\mathbf{a}_i$ is a linear combination of vectors in $A \setminus \{\mathbf{a}_i\}$, contrary to Proposition 1.4(a). So, there exists some $j \in \{1, \ldots, \ell\}$ such that $\alpha_j \neq 0$ and $\mathbf{b}_j \notin A \setminus \{\mathbf{a}_i\}$. Since $\mathbf{a}_i \notin B$, it follows that $\mathbf{b}_j \neq \mathbf{a}_i$ and $\mathbf{b}_j \notin A$.

It remains to show that $(B \setminus \{\mathbf{b}_j\}) \cup \{\mathbf{a}_i\}$ is a spanning set of V. Since $\mathbf{b}_j \neq \mathbf{a}_i$, we see that $(B \setminus \{\mathbf{b}_j\}) \cup \{\mathbf{a}_i\} = (B \cup \{\mathbf{a}_i\}) \setminus \{\mathbf{b}_j\}$, and we need to show that $(B \cup \{\mathbf{a}_i\}) \setminus \{\mathbf{b}_j\}$ is a spanning set of V. Since B is a spanning set of V, so is $B \cup \{\mathbf{a}_i\}$. In view of Proposition 1.4(b), it now suffices to show that $\mathbf{b}_j$ is a linear combination of the other vectors in $B \cup \{a_i\}$. Since $\mathbf{a}_i = \alpha_1\mathbf{b}_1 + \cdots + \alpha_\ell\mathbf{b}_\ell$, we see that

$$\alpha_j\mathbf{b}_j \;=\; \mathbf{a}_i - \alpha_1\mathbf{b}_1 - \cdots - \alpha_{j-1}\mathbf{b}_{j-1} - \alpha_{j+1}\mathbf{b}_{j+1} - \cdots - \alpha_\ell\mathbf{b}_\ell.$$

Since $\alpha_j \neq 0$, we deduce that

$$\begin{aligned}\mathbf{b}_j \;=\;& \alpha_j^{-1}\mathbf{a}_i - \alpha_j^{-1}\alpha_1\mathbf{b}_1 - \cdots - \alpha_j^{-1}\alpha_{j-1}\mathbf{b}_{j-1} - \\ & -\alpha_j^{-1}\alpha_{j+1}\mathbf{b}_{j+1} - \cdots - \alpha_j^{-1}\alpha_\ell\mathbf{b}_\ell.\end{aligned}$$

[6] Let us explain why B' exists. Clearly, B has a subset (namely itself) that is a spanning set of V. Of all subsets of B that span V, we choose B' to be one of minimum size.

[7] Note that this lemma is vacuously true if $A \subseteq B$, since in that case, there are no vectors $\mathbf{a} \in A \setminus B$.

So, $\mathbf{b}_j$ is indeed a linear combination of the other vectors in $B \cup \{\mathbf{a}_i\}$, and we are done. $\square$

Steinitz exchange lemma. *Let V be a vector space over a field $\mathbb{F}$, let $\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b}_1, \ldots, \mathbf{b}_\ell \in V$, and assume that $\mathbf{a}_1, \ldots, \mathbf{a}_k$ are pairwise distinct and $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$ are pairwise distinct. Assume that $A := \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is a linearly independent set in V, and assume that $B := \{\mathbf{b}_1, \ldots, \mathbf{b}_\ell\}$ is a spanning set of V. Then $k \leq \ell$ (i.e. $|A| \leq |B|$). Moreover, there exists a set $B' \subseteq B \setminus A$ such that $|B'| = |B| - |A| = \ell - k$ and $A \cup B'$ is a spanning set of V.*

Proof. We may assume that $A \not\subseteq B$, for otherwise, the result is immediate.[8] Set $p := |A \cap B|$.[9] After possibly permuting the elements of A and B, we may assume that the following hold:

- $\mathbf{a}_1 = \mathbf{a}_1, \ldots, \mathbf{a}_p = \mathbf{b}_p$;
- $\{\mathbf{a}_{p+1}, \ldots, \mathbf{a}_k\} \cap \{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \neq \emptyset$.

We now prove a technical claim.

> **Claim.** For all $t \in \{0, \ldots, k-p\}$, there exist pairwise distinct indices $i_1, \ldots, i_t \in \{p+1, \ldots, \ell\}$ such that
>
> $$\{\mathbf{a}_1, \ldots, \mathbf{a}_p\} \cup \{\mathbf{a}_{p+1}, \ldots, \mathbf{a}_{p+t}\} \cup \Big(\{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \setminus \{\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_t}\}\Big)$$
>
> is a spanning set of V.

Proof of the Claim. We proceed by induction on t, using Lemma 1.7.

For $t = 0$, we need only show that $\{\mathbf{a}_1, \ldots, \mathbf{a}_p\} \cup \{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\}$ is a spanning set of V. But note that $\{\mathbf{a}_1, \ldots, \mathbf{a}_p\} \cup \{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} = B$, and by hypothesis, B is a spanning set of V.

Now, fix some $t \in \{0, \ldots, k-p-1\}$, and assume inductively that the statement is true for t, i.e. that there exist pairwise distinct indices $i_1, \ldots, i_t \in \{p+1, \ldots, \ell\}$ such that

$$B_t \;:=\; \{\mathbf{a}_1, \ldots, \mathbf{a}_p\} \cup \{\mathbf{a}_{p+1}, \ldots, \mathbf{a}_{p+t}\} \cup \Big(\{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \setminus \{\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_t}\}\Big)$$

is a spanning set of V. Now, $\mathbf{a}_{p+t+1} \in A \setminus B_t$, and so by Lemma 1.7, there exists some $\mathbf{b} \in B_t \setminus A$ such that $(B_t \setminus \{\mathbf{b}\}) \cup \{\mathbf{a}_{p+t+1}\}$ is a spanning set of V. Since $\mathbf{b} \in B_t \setminus A$, we see that $\mathbf{b} \in \{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \setminus \{\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_t}\}$; consequently, there exists some $i_{t+1} \in \{p+1, \ldots, \ell\} \setminus \{i_1, \ldots, i_t\}$ such that

[8] Indeed, if $A \subseteq B$, then $|A| \leq |B|$, and we may set $B' := B \setminus A$.

[9] Since $A \not\subseteq B$, we see that $p < k$.

$\mathbf{b} = \mathbf{b}_{i_{t+1}}$. Now $i_1, \ldots, i_t, i_{t+1}$ are pairwise distinct indices in $\{p+1, \ldots, \ell\}$, and

$$\begin{aligned}(B_t \setminus \{\mathbf{b}\}) \cup \{\mathbf{a}_{p+t+1}\} &= \{\mathbf{a}_1, \ldots, \mathbf{a}_p\} \cup \{\mathbf{a}_{p+1}, \ldots, \mathbf{a}_{p+t}, \mathbf{a}_{p+t+1}\} \cup \\ &\quad \cup \Big(\{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \setminus \{\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_t}, \mathbf{b}_{i_{t+1}}\} \Big)\end{aligned}$$

is a spanning set of V. This completes the induction. ♦

We now apply the claim for $t = k-p$, and we get that there exist pairwise distinct indices $i_1, \ldots, i_{k-p} \in \{p+1, \ldots, \ell\}$ such that

$$C := \{\mathbf{a}_1, \ldots, \mathbf{a}_p\} \cup \{\mathbf{a}_{p+1}, \ldots, \mathbf{a}_k\} \cup \Big(\{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \setminus \{\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_{k-p}}\} \Big)$$

is a spanning set of V. But note that $|C| = \ell = |B|$ and $A \subseteq C$. Thus, $|A| \leq |C| = |B|$, and so $k \leq \ell$. Next, set $B' := \{\mathbf{b}_{p+1}, \ldots, \mathbf{b}_\ell\} \setminus \{\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_{k-p}}\}$. Then $B' \subseteq B \setminus A$, $|B'| = (\ell - p) - (k - p) = \ell - k = |B| - |A|$, and $C = A \cup B'$ is a spanning set of V. This completes the argument. □

Remark: For technical reasons (in order to get the set B'), the Steinitz exchange lemma assumes that the sets A and B contain no repetitions. However, if we only care about the $|A| \leq |B|$ part of the Steinitz exchange lemma (which is what we usually care about), then this assumption is not necessary. Indeed, suppose V is a vector space over a field $\mathbb{F}$, and suppose that A is a linearly independent set of vectors in V and that B is a spanning set of V (with repetitions allowed). Since A is linearly independent, it contains no repetitions; however, B may possibly contain repetitions. But then we let $\widetilde{B}$ be the set obtained from B by eliminating repetitions. Then $\widetilde{B}$ is still a spanning set of V, and by the Steinitz exchange lemma, we get that $|A| \leq |\widetilde{B}| \leq |B|$.

Corollary 1.8. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$. Then all bases of V are of the same size.*

Proof. We apply the Steinitz exchange lemma. Fix bases $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V. Since $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is linearly independent and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a spanning set of V, the Steinitz exchange lemma guarantees that $m \leq n$. On the other hand, since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly independent set and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is a spanning set of V, the Steinitz exchange lemma guarantees that $n \leq m$. So, $m = n$. □

The *dimension* of a finite-dimensional vector space V over a field $\mathbb{F}$, denoted by $\dim(V)$, is the number of elements in any basis of V (by Corollary 1.8, this is well-defined).

Remarks:

- Note that $\dim(\{\mathbf{0}\}) = 0$ (where $\{\mathbf{0}\}$ is understood to be a vector space over an arbitrary field $\mathbb{F}$), because $\emptyset$ is a basis of $\{\mathbf{0}\}$.
- For any field $\mathbb{F}$, we have that $\dim(\mathbb{F}^n) = n$, because the standard basis of $\mathbb{F}^n$ has n elements. The standard basis is not the only basis of $\mathbb{F}^n$ (except in some very special cases). See Theorem 4.1.

Proposition 1.9. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. Then both the following hold:*

(a) every linearly independent set of vectors in V has at most n vectors;

(b) every spanning set of V has at least n vectors.

Proof. Fix a basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of V. Then B is both a linearly independent set and a spanning set of V. Now, by the Steinitz exchange lemma, the number of vectors in any linearly independent set of V is at most the number of vectors in the spanning set B of V, which is n; so, (a) holds. On the other hand, by the Steinitz exchange lemma, any spanning set of V has at least as many vectors as the linearly independent set B; so, (b) holds. □

Our next proposition states that any linearly independent set of vectors in a finite-dimensional vector space can be extended to a basis of that vector space. (Compare Proposition 1.10 below to Proposition 1.6 above.)

Proposition 1.10. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$, and let $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ be a linearly independent set of vectors in V. Then there exists some basis of V that contains all of $\mathbf{a}_1, \ldots, \mathbf{a}_k$.*

Proof. Set $n := \dim(V)$. By Proposition 1.9, any linearly independent set of vectors in V has at most n vectors; in particular, $k \leq n$ (because $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is linearly independent). Now, let A be a linearly independent set that contains vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$, and subject to that, is of maximum possible size.[10] Set $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_{k+\ell}\}$. We claim that A is a basis of V. Since A is linearly independent, it suffices to show that A is a spanning set of V. Fix $\mathbf{v} \in V$; we must show that $\mathbf{v}$ is a linear combination of vectors in A. If $\mathbf{v} \in A$, then this is immediate.[11] So, assume that $\mathbf{v} \notin A$. Then by the maximality of A, the set $\{\mathbf{v}\} \cup A$ is not linearly independent. So, there exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_{k+\ell} \in \mathbb{F}$, not all zero, such that

$$\alpha_0\mathbf{v} + \alpha_1\mathbf{a}_1 + \cdots + \alpha_{k+\ell}\mathbf{a}_{k+\ell} \quad = \quad \mathbf{0}.$$

[10] Let us explain why A exists. There exists at least one linearly independent set that contains vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$, namely, the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$. On the other hand, all linearly independent sets are of size at most n, and in particular, there is an upper bound on the size of linearly independent sets containing $\mathbf{a}_1, \ldots, \mathbf{a}_k$. So, A exists.

[11] Indeed, suppose $\mathbf{v} \in A$. Then there exists an index $i \in \{1, \ldots, k+\ell\}$ such that $\mathbf{v} = \mathbf{a}_i$. Now set $\alpha_i = 1$, and for all $j \in \{1, \ldots, k+\ell\}$, set $\alpha_j = 0$. Then $\mathbf{v} = \alpha_1\mathbf{a}_1 + \cdots + \alpha_{k+\ell}\mathbf{a}_{k+\ell}$, and so $\mathbf{v}$ is linear combination of vectors in A.

If $\alpha_0 = 0$, then at least one of $\alpha_1, \ldots, \alpha_{k+\ell}$ is non-zero and $\alpha_1\mathbf{a}_1 + \cdots + \alpha_{k+\ell}\mathbf{a}_{k+\ell} = \mathbf{0}$, contrary to the fact that A is linearly independent. So, $\alpha_0 \neq 0$, and it follows that

$$\mathbf{v} \;=\; (-\alpha_0^{-1}\alpha_1)\mathbf{a}_1 + \cdots + (-\alpha_0^{-1}\alpha_{k+\ell})\mathbf{a}_{k+\ell},$$

and it follows that $\mathbf{v}$ is a linear combination of vectors in A. This proves that A is a basis of V, and we are done. □

Proposition 1.11. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. Then both the following hold:*

(a) any linearly independent set of n vectors of V is a basis of V;

(b) any set of n vectors of V that spans V is a basis of V.

Proof. We first prove (a). Let A be any linearly independent set of vectors in V such that $|A| = n$. By Proposition 1.10, V has a basis A' such that $A \subseteq A'$. Since $\dim(V) = n$, we see that $|A'| = n$. Since $|A| = n$ and $A \subseteq A'$, it follows that $A = A'$. Since A' is a basis of V, we see that A is a basis of V. This proves (a).

It remains to prove (b). Let B be any set of n vectors of V such that $V = \mathrm{Span}(B)$. Then by Proposition 1.6, V has a basis B' such that $B' \subseteq B$. Since $\dim(V) = n$, we see that $|B'| = n$. Since $|B| = n$ and $B' \subseteq B$, it follows that $B' = B$. Since B' is a basis of V, we see that B is a basis of V. This proves (b). □

The following theorem summarizes some of the main results that we have proven so far.

Theorem 1.12. *Let V be a finite-dimensional vector space over a vector space $\mathbb{F}$, and let U be a subspace of V. Then:*

- *U is finite-dimensional;*
- *$\dim(U) \leq \dim(V)$;*
- *if $\dim(U) = \dim(V)$, then $U = V$.*

Proof. Set $n := \dim(V)$. By Proposition 1.9, any linearly independent set of vectors in V contains at most n vectors. Since U is a subspace in V, we see that any linearly independent set of vectors in U is a linearly independent set of vectors in V, and consequently, it contains at most n vectors. Now, let $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a linearly independent set of vectors in U of maximum possible size.[12] (Then $k \leq n$.) Let us show that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ spans U. Fix $\mathbf{u} \in U$; we must show that $\mathbf{u}$ is a linear combination of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

[12] Possibly, $k = 0$, in which case, our linearly independent set is empty.

If $\mathbf{u} \in \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, then this is immediate. So, assume that $\mathbf{u} \notin \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. By the maximality of $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, we see that $\{\mathbf{u}, \mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is linearly dependent. So, there exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_k$, not all zero, such that $\alpha_0\mathbf{u} + \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k = \mathbf{0}$. If $\alpha_0 = 0$, then $\alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k = \mathbf{0}$ and at least one of the scalars $\alpha_1, \ldots, \alpha_k$ is non-zero, contrary to the fact that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is linearly independent. So, $\alpha_0 \neq 0$, and we deduce that $\mathbf{u} = (-\alpha_0^{-1}\alpha_1)\mathbf{u}_1 + \cdots + (-\alpha_0^{-1}\alpha_k)\mathbf{u}_k$. So, $\mathbf{u} \in \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, and we deduce that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a spanning set of U. So, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a basis of U, and it follows that U is finite-dimensional, with $\dim(U) = k$. So, $\dim(U) = k \leq n = \dim(V)$. Now, suppose that $\dim(U) = \dim(V)$, i.e. $k = n$. But now $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a linearly independent set of n vectors in V, and so Proposition 1.11 guarantees that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a basis of V. So, $U = \mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) = V$, and we are done. □

1.1 Infinite bases (optional)

We can define a basis of a vector space in more generality, as follows. Let V be a vector space over a field $\mathbb{F}$, and let $B \subseteq V$. (B may possibly be infinite. However, we do not allow repetitions in B.)

- B is *linearly independent* provided that for all pairwise distinct vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, and all scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$, if $\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0}$, then $\alpha_1 = \cdots = \alpha_k = 0$.[13]
- $\mathrm{Span}(B) = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k \mid \mathbf{v}_1, \ldots, \mathbf{v}_k \in V,\ \alpha_1, \ldots, \alpha_k \in \mathbb{F}\}$.[14]
- B is a basis of V if it satisfies the following two conditions:
  1. B is linearly independent;
  2. $V = \mathrm{Span}(B)$.

With a basis defined in this way, it is possible to show that every vector space has a (possibly infinite) basis. However, the proof uses "Zorn's lemma" (an equivalent of the "Axiom of Choice," which is studied in set theory) and is non-constructive. So, it is possible to show that every vector space has a basis, but for some vector spaces, we have no idea what a basis might look like. For instance, consider the set of all functions from $\mathbb{R}$ to $\mathbb{R}$; this set is a vector space (over $\mathbb{R}$) and therefore has a basis, but it is not known what a basis for this vector space might look like.

In some cases, though, we can get a "nice" infinite basis. For instance, $\mathbb{P}_{\mathbb{R}}$ has a basis $\{1, x, x^2, x^3, x^4, \ldots\}$.

[13] So, B is linearly independent if and only if all finite subsets of B are linearly independent.

[14] So, $\mathrm{Span}(B)$ is the set of vectors that can be expressed as a linear combination of finitely many vectors in B.

2 The rank of a matrix

Given a field $\mathbb{F}$, the *rank* of a matrix $A \in \mathbb{F}^{n\times m}$, denoted by $\text{rank}(A)$, is the number of non-zero rows of $\text{RREF}(A)$.

Example 2.1. *Find the rank of the matrix*

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_2$.

Solution. By row reducing, we get that

$$\text{RREF}(A) = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\text{RREF}(A)$ has three non-zero rows, and it follows that $\text{rank}(A) = 3$. □

Remark: Note that, for a matrix A, the number of non-zero rows of $\text{RREF}(A)$ is equal to the number of non-zero rows of any row echelon form (not necessarily reduced) of A. So, to find $\text{rank}(A)$, we need only perform the "forward" part of the row reduction algorithm in order to transform A into a matrix in row echelon form, and then we count the number of non-zero rows of the matrix that we obtain. The "backward" part of the row reduction algorithm is optional for computing rank.

Proposition 2.2. *Let* $\mathbb{F}$ *be a field, and let* $A \in \mathbb{F}^{n\times m}$. *Then rank*$(A)$ *is equal to the number of pivot columns of* A. *Moreover, rank*$(A) \leq \min\{n, m\}$.[15]

Proof. By definition, $\text{rank}(A)$ is equal to the number of non-zero rows of $\text{RREF}(A)$. The number of non-zero rows of $\text{RREF}(A)$ is precisely equal to the number of pivot positions of $\text{RREF}(A)$, which is equal to the number of pivot columns of $\text{RREF}(A)$. The pivot columns of A correspond to the ones in $\text{RREF}(A)$, and it follows that $\text{rank}(A)$ is equal to the number of pivot columns of A.

Note that n is the number of rows and m is the number of columns of A. The fact that $\text{rank}(A) \leq n$ follows immediately from the definition of rank, and the fact that $\text{rank}(A) \leq m$ follows from the fact (proven above) that $\text{rank}(A)$ is equal to the number of pivot columns of A. □

[15] So, $\text{rank}(A)$ is at most the number of rows of A, and similarly, $\text{rank}(A)$ is at most the number of columns of A.

3 The row space and the column space of a matrix

A *row vector* is a matrix that has only one row.

Given a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times m}$,

- the *column space* of A, denoted by $\mathrm{Col}(A)$, is the subspace of $\mathbb{F}^n$ spanned by the columns of A;[16]
- the *row space* of A, denoted by $\mathrm{Row}(A)$, is the subspace of $\mathbb{F}^{1\times m}$ spanned by the rows of A.[17]

Our goal in this section is to give a recipe for finding a basis of the column space and the row space of a matrix. As we shall see, both of those spaces have dimension precisely $\mathrm{rank}(A)$.

We begin with a technical proposition.

Proposition 3.1. *Let $\mathbb{F}$ be a field, let $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{F}^n$, and let $B \in \mathbb{F}^{n\times n}$ be an invertible matrix. Then all the following hold:*

(a) $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is linearly independent if and only if $\{B\mathbf{a}_1, \ldots, B\mathbf{a}_k\}$ is linearly independent;

(b) for all $\mathbf{v} \in \mathbb{F}^n$, $\mathbf{v} \in \mathit{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ if and only if $B\mathbf{v} \in \mathit{Span}(B\mathbf{a}_1, \ldots, B\mathbf{a}_k)$;

Proof. We first prove (a). Suppose first that $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is linearly independent. We must show that $\{B\mathbf{a}_1, \ldots, B\mathbf{a}_k\}$. Fix scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that

$$\alpha_1 B\mathbf{a}_1 + \cdots + \alpha_k B\mathbf{a}_k \quad = \quad \mathbf{0}.$$

Since B is invertible, it has an inverse B^{-1}. By multiplying both sides of the equation above by B^{-1}, we obtain

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_k \mathbf{a}_k \quad = \quad \mathbf{0}.$$

Since $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is linearly independent, we have that $\alpha_1 = \cdots = \alpha_k = 0$. So, $\{B\mathbf{a}_1, \ldots, B\mathbf{a}_k\}$ is linearly independent.

Suppose, conversely, that $\{B\mathbf{a}_1, \ldots, B\mathbf{a}_k\}$ is linearly independent. Fix scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_k \mathbf{a}_k \quad = \quad \mathbf{0}.$$

[16] More precisely, if $A = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$ (i.e. $\mathbf{a}_1, \ldots, \mathbf{a}_m$ are the columns of A, appearing in A that order, from left to right), then $\mathrm{Col}(A) := \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_m)$.

[17] More precisely, if $A = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix}$ (i.e. $\mathbf{r}_1, \ldots, \mathbf{r}_n$ are the rows of A, appearing in A in that order, from top to bottom), then $\mathrm{Row}(\mathbf{r}_1, \ldots, \mathbf{r}_n)$.

We now multiply both sides by B, and we obtain

$$\alpha_1(B\mathbf{a}_1) + \cdots + \alpha_k(B\mathbf{a}_k) \quad = \quad \mathbf{0}.$$

Since $\{B\mathbf{a}_1, \ldots, B\mathbf{a}_k\}$ is linearly independent, it follows that $\alpha_1 = \cdots = \alpha_k = 0$. So, $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ is linearly independent. This completes the proof of (a).

We now prove (b). Fix $\mathbf{v} \in \mathbb{F}^n$. Suppose first that $\mathbf{v} \in \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$. Then there exist scalars $\alpha_1, \ldots, \alpha_k$ such that $\mathbf{v} = \alpha_1\mathbf{a}_1 + \cdots + \alpha_k\mathbf{a}_k$. By multiplying both sides by B, we get $B\mathbf{v} = \alpha_1(B\mathbf{a}_1) + \cdots + \alpha_k(B\mathbf{a}_k)$, and so $B\mathbf{v} \in \mathrm{Span}(B\mathbf{a}_1, \ldots, B\mathbf{a}_k)$.

Suppose, conversely, that $B\mathbf{v} \in \mathrm{Span}(B\mathbf{a}_1, \ldots, B\mathbf{a}_k)$. Then there exist scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that

$$B\mathbf{v} \quad = \quad \alpha_1(B\mathbf{a}_1) + \cdots + \alpha_k(B\mathbf{a}_k).$$

Since B is invertible, it has an inverse B^{-1}. We now multiply both sides of the equation by B^{-1}, and we obtain $\mathbf{v} = \alpha_1\mathbf{a}_1 + \cdots + \alpha_k\mathbf{a}_k$. So, $\mathbf{v} \in \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$. This proves (b). □

Theorem 3.2. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times m}$. Then the pivot columns of A form a basis of Col(A).*[18] *Moreover, dim$\Big($Col(A)$\Big)$ = rank(A).*

Proof. Set $r := \mathrm{rank}(A)$. By Proposition 2.2, r is equal to the number of pivot columns of A. So, the first statement implies the second.

It remains to prove the first statement. Set $A = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{bmatrix}$. Let $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ (with $1 \leq i_1 < \cdots < i_r \leq m$) be the pivot columns of A. We must show that $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$ is a basis of $\mathrm{Col}(A)$.

Set $U = \mathrm{RREF}(A)$. Then there exist elementary matrices $E_1, \ldots, E_k$ such that $E_k \ldots E_1 A = U$. Set $B := E_k \ldots E_1$. Then B is invertible and $U = BA = \begin{bmatrix} B\mathbf{a}_1 & \ldots & B\mathbf{a}_m \end{bmatrix}$. Moreover, since U is in reduced row echelon form, all the following hold:

(i) $B\mathbf{a}_{i_1}, \ldots, B\mathbf{a}_{i_r}$ are the pivot columns of U;

(ii) for all $j \in \{1, \ldots, r\}$, we have that $B\mathbf{a}_{i_j} = \mathbf{e}^n_j$;

(iii) in any column of U, only the first r entries may possibly be non-zero (the other entries are all zero).

Clearly, $\{\mathbf{e}^n_1, \ldots, \mathbf{e}^n_r\}$ is a linearly independent set; so, by (ii), we have that $\{B\mathbf{a}_{i_1}, \ldots, B\mathbf{a}_{i_r}\}$ is a linearly independent set. Consequently, by Proposition 3.1(a), $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$ is a linearly independent set. Moreover, it is

[18] **Warning:** We need to take the pivot columns of the original matrix A, not of RREF(A).

clear that any vector in $\mathbb{F}^n$ in which only the top r entries may possibly be non-zero (and the other entries are all zero), is a linear combination of vectors $\mathbf{e}_1^n, \ldots, \mathbf{e}_r^n$. So, (i), (ii), and (iii) together imply that every column of $U = \left[\begin{array}{ccc} B\mathbf{a}_1 & \ldots & B\mathbf{a}_m \end{array} \right]$ is a linear combination of vectors $B\mathbf{a}_{i_1}, \ldots, B\mathbf{a}_{i_r}$. But now by Proposition 3.1(b), we see that every column of $A = \left[\begin{array}{ccc} \mathbf{a}_1 & \ldots & \mathbf{a}_m \end{array} \right]$ is a linear combination of vectors $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$. So, by Propositon 1.5, $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$ is a spanning set of $\mathrm{Col}(A)$.[19] It now follows that $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$ is a basis of $\mathrm{Col}(A)$, and we are done. □

Proposition 3.3. *Let $\mathbb{F}$ be a field. Then any two row equivalent matrices in $\mathbb{F}^{n\times m}$ have the same row space.*

Proof.

> **Claim.** Let $A, B \in \mathbb{F}^{n\times m}$ be matrices such that B is obtained from A by performing one elementary row operation. Then $\mathrm{Row}(A) = \mathrm{Row}(B)$.

Proof of the Claim. Set $A = \left[\begin{array}{c} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{array} \right]$ and $B = \left[\begin{array}{c} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{array} \right]$ (so, $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the rows of A appearing in that order in A, from top to bottom, and similar for B). By definition, $\mathrm{Row}(A) = \mathrm{Span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ and $\mathrm{Row}(B) = \mathrm{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_n)$.

Since B is obtained from A by performing one elementary row operation R, we know that A can be obtained from B by performing one elementary row operation (the one that "undoes" R). So, it is enough to show that $\mathrm{Row}(A) \subseteq \mathrm{Row}(B)$, for then analogous argument will establish that $\mathrm{Row}(B) \subseteq \mathrm{Row}(A)$, and then the result will follow.

If B is obtained by swapping two rows of A, then obviously, $\mathrm{Row}(A) = \mathrm{Row}(B)$. Next, suppose B is obtained by multiplying one row of A (say, row $\mathbf{a}_i$) by a non-zero scalar $\alpha \in \mathbb{F}$. Now, fix $\mathbf{v} \in \mathrm{Row}(A)$; we must show that $\mathbf{v} \in \mathrm{Row}(B)$. Since $\mathbf{v} \in \mathrm{Row}(A)$, there exist scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1\mathbf{a}_1 + \cdots + \alpha_n\mathbf{a}_n$. But now

$$\begin{aligned} \mathbf{v} &= \alpha_1\mathbf{a}_1 + \cdots + \alpha_{i-1}\mathbf{a}_{i-1} + \alpha_i\mathbf{a}_i + \alpha_{i+1}\mathbf{a}_{i+1} + \cdots + \alpha_n\mathbf{a}_n \\ &= \alpha_1\mathbf{a}_1 + \cdots + \alpha_{i-1}\mathbf{a}_{i-1} + (\alpha_i\alpha^{-1})(\alpha\mathbf{a}_i) + \alpha_{i+1}\mathbf{a}_{i+1} + \cdots + \alpha_n\mathbf{a}_n \\ &= \alpha_1\mathbf{b}_1 + \cdots + \alpha_{i-1}\mathbf{b}_{i-1} + (\alpha_i\alpha^{-1})\mathbf{b}_i + \alpha_{i+1}\mathbf{b}_{i+1} + \cdots + \alpha_n\mathbf{b}_n, \end{aligned}$$

and so $\mathbf{v} \in \mathrm{Row}(B)$. Thus, $\mathrm{Row}(A) \subseteq \mathrm{Row}(B)$.

[19] By definition, $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ is a spanning set of $\mathrm{Col}(A)$. By what we just showed, every vector in $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ is a linear combination of vectors in $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$. So, by Propositon 1.5 $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}\}$ is a spanning set of $\mathrm{Col}(A)$.

Finally, suppose that B is obtained from A by adding a scalar multiple of one row to another row. Then there exist distinct indices $i, j \in \{1, \ldots, n\}$ and a scalar $\alpha \in \mathbb{F}$ such that $\mathbf{b}_i = \mathbf{a}_i + \alpha \mathbf{a}_j$, and $\mathbf{b}_k = \mathbf{a}_k$ for all $k \in \{1, \ldots, n\} \setminus \{i\}$. Now, fix $\mathbf{v} \in \text{Row}(A)$. Then there exist scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n$. We now set $\beta_j := \alpha_j - \alpha_i \alpha$, and we set $\beta_k := \alpha_k$ for all $k \in \{1, \ldots, k\}$. Then $\mathbf{v} = \beta_1 \mathbf{b}_1 + \cdots + \beta_n \mathbf{b}_n$,[20] and so $\mathbf{v} \in \text{Row}(B)$. Thus, $\text{Row}(A) \subseteq \text{Row}(B)$. ♦

Now, fix $A, B \in \mathbb{F}^{n \times m}$ such that $A \sim B$.[21] Then there exists a sequence $R_1, \ldots, R_k$ of elementary row operations such that, by starting with A and then successively applying $R_1, \ldots, R_k$ to it, we obtain B. By the Claim, each time we apply an elementary row operation, the row space remains unchanged. So, $\text{Row}(A) = \text{Row}(B)$.[22] □

Theorem 3.4. *Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times m}$, and let U be any matrix in row echelon form that is row equivalent to A.*[23] *Then the non-zero rows of U form a basis of Row(A). Moreover, dim$\Big($Row(A)$\Big)$ = rank(A).*

Proof. Set $r := \text{rank}(A)$. By definition, r is equal to the number of non-zero rows of U. So, the first statement implies the second. Moreover, by Proposition 3.3, $\text{Row}(A) = \text{Row}(U)$. So, it suffices to show that the non-zero rows of U form a basis of $\text{Row}(U)$. Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be the non-zero rows of U, appearing in that order (from top to bottom) in U. We must show that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a basis of $\text{Row}(U)$. Clearly, $\text{Row}(U) = \text{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. It remains to show that $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is a linearly independent set. Fix scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ such that $\alpha_1 \mathbf{u}_1 + \ldots \alpha_k \mathbf{u} = \mathbf{0}$. We must show that $\alpha_1 = \cdots = \alpha_k = 0$. Suppose otherwise, and let $i \in \{1, \ldots, k\}$ be the smallest index such that $\alpha_i \neq 0$. We may assume that the leading entry (i.e. the leftmost non-zero entry) of the row $\mathbf{u}_i$ is in position j. But since U is in row echelon form, the leading entries of $\mathbf{u}_{i+1}, \ldots, \mathbf{u}_k$ are all strictly to the right of the leading entry of $\mathbf{u}_i$, and so their j-th entry is 0. Since $\alpha_1 = \cdots = \alpha_{i-1} = 0$

[20] Indeed, if $i < j$, then

$$\begin{aligned} \mathbf{v} &= \alpha_1 \mathbf{a}_1 + \cdots + \alpha_{i-1} \mathbf{a}_{i-1} + \alpha_i \mathbf{a}_i + \cdots + \alpha_j \mathbf{a}_j + \alpha_{j+1} \mathbf{a}_{j+1} + \cdots + \alpha_n \mathbf{a}_n \\ &= \alpha_1 \mathbf{a}_1 + \cdots + \alpha_{i-1} \mathbf{a}_{i-1} + \alpha_i (\mathbf{a}_i + \alpha \mathbf{a}_j) + \cdots + \\ &\quad + (\alpha_j - \alpha_i \alpha) \mathbf{a}_j + \alpha_{j+1} \mathbf{a}_{j+1} + \cdots + \alpha_n \mathbf{a}_n \\ &= \beta_1 \mathbf{b}_1 + \cdots + \beta_n \mathbf{b}_n. \end{aligned}$$

The calculation is almost identical when $j < i$.

[21] As usual, $A \sim B$ means that A and B are row equivalent.

[22] Technically, we are doing an induction on the number of elementary row operations. (Details?)

[23] It may be that $U = \text{RREF}(A)$, but this assumption is not necessary. U may be any matrix in row echelon form obtained from A via a sequence of elementary row operations. For instance, U may be the matrix obtained from A by only performing the "forward" part of the row reduction algorithm in order to transform A into a matrix in row echelon form.

(by the minimality of i), it follows that the j-th entry of $\alpha_1\mathbf{u}_1 + \ldots \alpha_k\mathbf{u}$ is $\alpha_i \neq 0$, contrary to the fact that $\alpha_1\mathbf{u}_1 + \ldots \alpha_k\mathbf{u} = \mathbf{0}$. □

Example 3.5. *Consider the matrix*

$$A = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 2 & 1 \end{bmatrix}$$

with entries understood to be in $\mathbb{R}$.

(a) Compute rank(A).

(b) Find a basis of Col(A).

(c) Find a basis of Row(A).

Solution. By performing the forward part of the row reduction algorithm, we see that the following matrix is a row echelon form A:

$$U = \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(a) The matrix U has three non-zero rows, and so $\text{rank}(A) = 3$.

(b) The pivot columns of U are its first, second, and fifth column. So, the pivot columns of A are its first, second, and fifth column, and so those columns of A form a basis for $\text{Col}(A)$. More precisely, the following is a basis of $\text{Col}(A)$:

$$\left\{ \begin{bmatrix} 0 \\ 3 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -7 \\ -9 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \\ 6 \\ 2 \end{bmatrix} \right\}.$$

(c) The non-zero rows of U form a basis of $\text{Row}(A)$. So, the following is a basis of $\text{Row}(A)$:

$$\left\{ \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}, \begin{bmatrix} 0 & 2 & -4 & 4 & 2 & -6 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \right\}.$$

□

Corollary 3.6. *Let* $\mathbb{F}$ *be a field, and let* $A \in \mathbb{F}^{n\times m}$. *Then*

$$dim\Big(Col(A)\Big) = dim\Big(Row(A)\Big) = rank(A).$$

Proof. This follows immediately from Theorems 3.2 and 3.4. □

Theorem 3.7. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times m}$. Then both the following hold:*

(a) the columns of A are linearly independent if and only if rank(A) is equal to the number of columns of A (i.e. rank$(A) = m$);

(b) the columns of A span $\mathbb{F}^n$ if and only if rank(A) is equal to the number of rows of A (i.e. rank$(A) = n$).

Proof. We first prove (a). Suppose first that the columns of A are linearly independent. Then the columns of A form a basis of $\mathrm{Col}(A)$, and consequently, $\dim\Big(\mathrm{Col}(A)\Big) = m$. But then by Theorem 3.2, we have that $\mathrm{rank}(A) = \dim\Big(\mathrm{Col}(A)\Big) = m$.

Suppose, conversely, that $\mathrm{rank}(A) = m$. Then by Proposition 2.2, A has m pivot columns, i.e. all columns of A are pivot columns. But by Theorem 3.2, the pivot columns of A form a basis of $\mathrm{Col}(A)$, and in particular, they are linearly independent. It follows that the columns of A are linearly independent. This proves (a).

It remains to prove (b). Set $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$. We now have the following sequence of equivalent statements:

$$\begin{aligned}
&\text{the columns of } A \text{ span } \mathbb{F}^n && \iff && \underbrace{\mathrm{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)}_{=\mathrm{Col}(A)} = \mathbb{F}^n \\
& && \iff && \mathrm{Col}(A) = \mathbb{F}^n \\
& && \overset{\text{by Theorem 1.12}}{\iff} && \dim\Big(\mathrm{Col}(A)\Big) = \underbrace{\dim(\mathbb{F}^n)}_{=n} \\
& && \overset{\text{by Theorem 3.2}}{\iff} && \mathrm{rank}(A) = n
\end{aligned}$$

So, (b) holds. □

4 More on invertible matrices and transposes

Theorem 4.1. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$.*[24] *Then the following are equivalent:*

(a) A is invertible;

[24] Note that we are assuming that A is a square matrix. Do not apply this theorem to non-square matrices!

(b) RREF$(A) = I_n$*;*

(c) rank$(A) = n$*;*

(d) the columns of A are linearly independent;[25]

(e) the columns of A span $\mathbb{F}^n$*;*

(f) the columns of A form a basis of $\mathbb{F}^n$*.*[26]

Proof. The fact that (a) and (b) are equivalent follows from Corollary 5.1 from Lecture Notes 4. Further, since A is an $n \times n$ matrix, it is clear that $\text{rank}(A) = n$ if and only if $\text{RREF}(A) = I_n$; so, (b) and (c) are equivalent. We have now shown that (a), (b), and (c) are equivalent.

By definition, (f) implies (d) and (e). On the other hand, by Proposition 1.11(a), (d) implies (f), and by Proposition 1.11(b), (e) implies (f). We now have that (d), (e) and (f) are equivalent.

Finally, by Theorem 3.7(a), we have that (c) and (d) are equivalent. This completes the argument. □

Corollary 4.2. *Let* $\mathbb{F}$ *be a field. Then both the following hold:*

(a) for all $A \in \mathbb{F}^{n\times m}$*, rank*$(A) =$ *rank*(A^T)*;*

(b) for all $A \in \mathbb{F}^{n\times n}$*, A is invertible if and only if* A^T *is invertible.*

Proof. We first prove (a). Fix $A \in \mathbb{F}^{n\times m}$. Then

$$\text{rank}(A^T) \overset{(*)}{=} \dim(\text{Col}(A^T)) = \dim(\text{Row}(A)) \overset{(**)}{=} \text{rank}(A),$$

where both (*) and (**) follow from Corollary 3.6. This proves (a).

It remains to prove (b). Fix $A \in \mathbb{F}^{n\times n}$. Then we have the following sequence of equivalences:

$$\begin{aligned} A \text{ is invertible} &\overset{\text{by Theorem 4.1}}{\iff} \text{rank}(A) = n \\ &\overset{\text{by part (a)}}{\iff} \text{rank}(A^T) = n \\ &\overset{\text{by Theorem 4.1}}{\iff} A^T \text{ is invertible} \end{aligned}$$

This proves (b). □

[25] Here, repetitions count. In particular, if A has two identical columns, then the columns of A are **not** linearly independent.

[26] Again, repetitions count. In particular, if A has two identical columns, then the columns of A are **not** linearly independent, and therefore, they do not form a basis of $\mathbb{F}^n$.

5 The null space of a matrix. The rank-nullity theorem for matrices

For field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times m}$, we define the *null space* of A, denoted by $\mathrm{Nul}(A)$, to be the set of all solutions of the equation $A\mathbf{x} = \mathbf{0}$, i.e.

$$\mathrm{Nul}(A) \;:=\; \{\mathbf{x} \in \mathbb{F}^m \mid A\mathbf{x} = \mathbf{0}\}.$$

Notation: In some texts, notation $\mathrm{Ker}(A)$ is used instead of $\mathrm{Nul}(A)$. "Ker" stands for "kernel."

Proposition 5.1. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times m}$. Then Nul$(A)$ is a subspace of $\mathbb{F}^m$.*

Proof. We apply Theorem 2.7 from Lecture Notes 6. First, $A\mathbf{0} = \mathbf{0}$, and so $\mathbf{0} \in \mathrm{Nul}(A)$. Next, if $\mathbf{u}, \mathbf{v} \in \mathrm{Nul}(A)$, then

$$\begin{aligned} A(\mathbf{u}+\mathbf{v}) &= A\mathbf{u} + A\mathbf{v} \\ &= \mathbf{0} + \mathbf{0} \qquad \text{because } \mathbf{u}, \mathbf{v} \in \mathrm{Nul}(A) \\ &= \mathbf{0}, \end{aligned}$$

and so $\mathbf{u} + \mathbf{v} \in A$. Finally, if $\mathbf{u} \in \mathrm{Nul}(A)$ and $\alpha \in \mathbb{F}$, then

$$\begin{aligned} A(\alpha\mathbf{u}) &= \alpha(A\mathbf{u}) \\ &= \alpha\mathbf{0} \qquad \text{because } \mathbf{u} \in \mathrm{Nul}(A) \\ &= \mathbf{0}, \end{aligned}$$

and so $\alpha\mathbf{u} \in \mathrm{Nul}(A)$. It now follows from Theorem 2.7 from Lecture Notes 6 that $\mathrm{Nul}(A)$ is a subspace of $\mathbb{F}^n$. □

Example 5.2. *Let*

$$A \;=\; \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

with entries understood to be in $\mathbb{Z}_2$. Find a basis for Nul(A). What is $\dim\Big(\mathrm{Nul}(A)\Big)$?

Proof. By row reducing, we see that

$$\mathrm{RREF}(\left[\begin{array}{c|c} A & \mathbf{0} \end{array}\right]) \;=\; \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The general solution of $\mathbf{Ax} = \mathbf{0}$ is

$$\mathbf{x} \;=\; \begin{bmatrix} r+t \\ s \\ r \\ s \\ t \end{bmatrix}, \qquad r, s, t \in \mathbb{Z}_2,$$

that is,

$$\mathbf{x} \;=\; r\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \qquad r, s, t \in \mathbb{Z}_2.$$

So,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis of $\mathrm{Nul}(A)$, and it follows that $\dim\Big(\mathrm{Nul}(A)\Big) = 3$. □

Note that for the matrix A from Example 5.2 satisfies $\mathrm{rank}(A) = 2$ and $\dim\Big(\mathrm{Nul}(A)\Big) = 3$. The sum of these two numbers is 5, which is the number of columns of A. As the the rank-nullity theorem for matrices (see below) shows, this is not an accident. We give a slightly informal proof of the rank-nullity theorem for matrices (however, this proof hopefully provides the right intuition). We will give a fully formal proof of the (more general) rank-nullity theorem for linear transformations in a subsequent lecture.

Rank–nullity theorem (matrix version). *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times m}$. Then*

$$\mathit{rank}(A) + \mathit{dim}\Big(\mathit{Nul}(A)\Big) \;=\; \underbrace{m}_{\substack{= \text{ number of} \\ \text{columns of } A}}.$$

Proof (outline/informal). By Theorem 3.2, we know that $\mathrm{rank}(A)$ is equal to the number of pivot columns of A. On the other hand, when computing the general solution of $\mathbf{Ax} = \mathbf{0}$, the number of free variables is equal to the number of non-pivot columns of A, and the number of vectors in a basis of $\mathrm{Nul}(A)$ is equal to the number of free variables.[27] So, $\dim\Big(\mathrm{Nul}(A)\Big)$

[27] This last part ("the number of vectors in a basis of $\mathrm{Nul}(A)$ is equal to the number of free variables") is not fully justified, and we omit the full details. Can you convince yourself this is true?

is equal to the number of non-pivot columns of A. It now follows that $\text{rank}(A) + \dim\Big(\text{Nul}(A)\Big)$ is equal to the number of columns of A, and we are done. $\square$