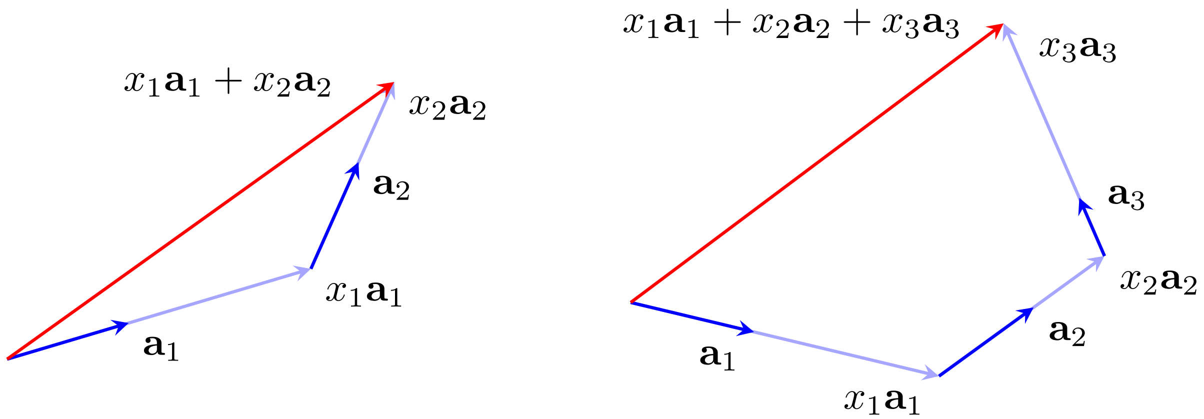

1. The equation $A\mathbf{x} = \mathbf{b}$

Let $A \in \mathbb{R}^{m\times n}$. The nature of the solution set of the equation $A\mathbf{x} = \mathbf{b}$, for $\mathbf{b} \in \mathbb{R}^m$, is governed by two sets:

Definition 1.1.

(1) The set
$$N(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$
is called the *nullspace of* A. It's the solution set of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

(2) The set
$$C(A) := \{\mathbf{b} \in \mathbb{R}^m : \mathbf{b} = A\mathbf{x},\ \text{for some } \mathbf{x} \in \mathbb{R}^n\}$$
is called the *column space of* A. It's the set of vectors $\mathbf{b} \in \mathbb{R}^m$ for which the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

To determine $N(A)$ is to solve the equation $A\mathbf{x} = \mathbf{0}$.

Example 1.2. Let
$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}.$$
The nullspace of A is the solution set of
$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}:$$
$$N(A) = \left\{ t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

Example 1.3. Let
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Setting

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \in \mathbb{R}^3,$$

and going through the motions of solving the system

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{✠}$$

we identify a condition under which $\mathbf{b}$ belongs to $C(A)$.

Application of two elementary row operations (which ones?) transforms

$$\begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & b_1 \\ 4 & 5 & 6 & b_2 \\ 7 & 8 & 9 & b_3 \end{bmatrix},$$

the augmented matrix of (✠), into

$$\begin{bmatrix} 1 & 2 & 3 & b_1 \\ 0 & -3 & -6 & b_2 - 4b_1 \\ 0 & 0 & 0 & b_1 - 2b_2 + b_3 \end{bmatrix},$$

It follows that $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $b_1 - 2b_2 + b_3 = 0$. Thus,

$$C(A) = \left\{ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} : b_1 - 2b_2 + b_3 = 0 \right\}.$$

Note that this description makes it easy to test a vector $\mathbf{b}$ for membership in $C(A)$ — simply check whether $b_1 - 2b_2 + b_3$ is 0 or not.

Finally, notice that

$$\begin{aligned} C(A) &= \left\{ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} : \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = 0 \right\} \\ &= N\left(\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}\right). \end{aligned}$$

The column space of one matrix can be the nullspace of another!

Exercise 1.4. Let $A \in \mathbb{R}^{m\times m}$ be an invertible matrix. Prove that $N(A) = \{\mathbf{0}\}$ and that $C(A) = \mathbb{R}^m$.

The set $C(A)$ governs the solubility of $A\mathbf{x} = \mathbf{b}$, while $N(A)$ determines the size of the corresponding set. More precisely:

Theorem 1.5. *Let $\mathbf{b} \in C(A)$, so that $A\mathbf{x} = \mathbf{b}$ has at least one solution, $\mathbf{x}_0$, say. Then*

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} = \{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(A)\}.$$

Proof. Let $\mathbf{b} \in C(A)$. Let $\mathbf{x} \in \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$. Then

$$A(\mathbf{x} - \mathbf{x}_0) = A\mathbf{x} - A\mathbf{x}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Thus, $\mathbf{y} := \mathbf{x} - \mathbf{x}_0 \in N(A)$ and $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$. Therefore, $\mathbf{x} \in \{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(A)\}$. Since $\mathbf{x} \in \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ was arbitrary,

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} \subseteq \{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(A)\}.$$

Conversely, suppose $\mathbf{x} \in \{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(A)\}$, i.e., $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$, for some $\mathbf{y} \in N(A)$. Now $A\mathbf{x}_0 = \mathbf{0}$ by hypothesis and $A\mathbf{y} = \mathbf{0}$ by definition of $N(A)$. Therefore,

$$A\mathbf{x} = A(\mathbf{x}_0 + y) = A\mathbf{x}_0 + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Thus, $\mathbf{x} \in N(A)$. Since $\mathbf{x} \in \{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(A)\}$ was arbitrary,

$$\{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(A)\} \subseteq \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}.$$

Having proved the reverse containment above, equality holds. $\square$

Example 1.6. Let

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}.$$

Since

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} =: \mathbf{b}$$

and, by Example 1.3,

$$N(A) = \left\{ t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\},$$

Theorem 1.5 implies that

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

2. Subspaces

Definition 2.1. A subset U of $\mathbb{R}^m$ is a *subspace of* $\mathbb{R}^m$ if:

(1) The zero vector of $\mathbb{R}^m$ belongs to U:

$$\mathbf{0} \in U.$$

(2) U is *closed under addition*:

$$\text{If } \mathbf{u}_1 \in U \text{ and } \mathbf{u}_2 \in U \text{ then } \mathbf{u}_1 + \mathbf{u}_2 \in U.$$

(3) U is *closed under scalar multiplication*:

$$\text{If } \mathbf{u} \in U \text{ and } s \in \mathbb{R} \text{ then } s\mathbf{u} \in U.$$

Remark 2.2.

- When attempting to prove that a subset U of $\mathbb{R}^m$ satisfies property (2), you need to show that $\mathbf{u}_1 + \mathbf{u}_2 \in U$ for *any* pair of elements $\mathbf{u}_1, \mathbf{u}_2$ of U. Similarly, to prove that a subset U of $\mathbb{R}^m$ satisfies property (3), you need to show that $s\mathbf{u} \in U$ for *any* element $\mathbf{u}_1$ of U and *any* real number s.
- To show that a given subset U of $\mathbb{R}^m$ is *not* a subspace of $\mathbb{R}^m$, you need only show that *one of* (1), (2), and (3) fail. To show that (2) fails, you need only produce *one* pair $\mathbf{u}_1, bu_2$ of elements of U such that $\mathbf{u}_1 + \mathbf{u}_2 \notin U$. Similarly, to prove that (3) fails, you need only produce *one* element $\mathbf{u}$ of U and *one* real number s such that $s\mathbf{u} \notin U$.

Example 2.3. Let

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x + y = 0 \right\}.$$

Then U is a subspace of $\mathbb{R}^2$. We prove this by verifying that U satisfies properties (1), (2), and (3) of Definition 2.1:

(1) Since $0 + 0 = 0$,

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in U,$$

by defintion of U.

(2) Let

$$\begin{bmatrix} x_1 \\ y_2 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \in U.$$

Then, by definition of U,

$$x_1 + y_1 = 0 \quad \text{and} \quad x_2 + y_2 = 0.$$

Therefore,

$$(x_1 + x_2) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) = 0.$$

Thus,

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} \in U,$$

by definition of U.

(3) Let

$$\begin{bmatrix} x \\ y \end{bmatrix} \in U, \quad s \in \mathbb{R}.$$

Then, by defintion of U, we have

$$x + y = 0.$$

Therefore,

$$sx + xy = s(x + y) = s(0) = 0.$$

Thus,

$$s \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} sx \\ sy \end{bmatrix} \in U.$$

Example 2.4. Let

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 + y^2 = 1 \right\}.$$

Then U is not a subspace of $\mathbb{R}^2$ as $\mathbf{0} \notin U$:

$$0^2 + 0^2 = 0 \neq 1.$$

Example 2.5. Let

$$U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x^2 = y^2 \right\}.$$

Then U is not a subspace of $\mathbb{R}^2$ since it fails property (2) of Definition 2.1:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in U, \text{ but } \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \notin U.$$

Example 2.6. Let $\mathbb{Z}$, the set of integers, viewed as a subset of in $\mathbb{R} = \mathbb{R}^1$. Then $\mathbb{Z}$ is not a subspace of $\mathbb{R}^1$ as $1 \in \mathbb{Z}$ but

$$\tfrac{1}{2} \cdot 1 = \tfrac{1}{2} \notin \mathbb{Z}.$$

Theorem 2.7. *Let $A \in \mathbb{R}^{m \times n}$. Then $N(A)$ is a subspace of $\mathbb{R}^n$.*

Proof. We prove properties (1), (2), and (3) of Definition 2.1.

(1) Since $A\mathbf{0} = \mathbf{0}$ we have $\mathbf{0} \in N(A)$ by defintion of $N(A)$.

(2) Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be arbitrary elements of $N(A)$. Then, by definition of $N(A)$,

$$A\mathbf{x}_1 = \mathbf{0} \quad \text{and} \quad A\mathbf{x}_2 = \mathbf{0}.$$

Therefore,

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = 0.$$

Thus, $\mathbf{x}_1 + \mathbf{x}_2 \in N(A)$, by definition of $N(A)$.

(3) Let $\mathbf{x}$ be an arbitrary element of $N(A)$ and let s be an y real number. As $\mathbf{x} \in N(A)$, we have $A\mathbf{x} = \mathbf{0}$, by defintion of $N(A)$. Therefore,

$$A(s\mathbf{x}) = s(A\mathbf{x}) = s\mathbf{0} = \mathbf{0}.$$

Thus, $s\mathbf{x} \in N(A)$, by definition of $N(A)$. □

Example 2.8. We give a shorter proof that the set U of Example 2.3 is a subspace of $\mathbb{R}^2$. By the identity

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \end{bmatrix},$$

and the defintion of $N\left(\begin{bmatrix} 1 & 1 \end{bmatrix}\right)$, we have

$$\begin{aligned} U &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x + y = 0 \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} \right\} \\ &= N\left(\begin{bmatrix} 1 & 1 \end{bmatrix}\right). \end{aligned}$$

Therefore, by Theorem 2.7, U is a subspace of $\mathbb{R}^2$.

Example 2.9. Let

$$U = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + 2y = 3z \text{ and } 4z - 2x = y \right\}.$$

We have:

$$\begin{aligned} & x + 2y = 3z && \iff && \begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} \\ \text{and} \quad & 4z - 2x = y && \iff && \begin{bmatrix} -2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}. \end{aligned}$$

If follows that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in U \iff \begin{bmatrix} 1 & 2 & -3 \\ -2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Therefore,

$$U = N\left(\begin{bmatrix} 1 & 2 & -3 \\ -2 & -1 & 4 \end{bmatrix} \right),$$

by definition of nullspace. Thus, by Theorem 2.7, U is a subspace of $\mathbb{R}^3$.

Theorem 2.10. *Let $A \in \mathbb{R}^{m \times n}$. Then $C(A)$ is a subspace of $\mathbb{R}^m$.*

Proof. We prove properties (1), (2), and (3) of Definition 2.1.

(1) Since $\mathbf{0} = A\mathbf{0}$ we have $\mathbf{0} \in C(A)$, by defintion of $C(A)$.

(2) Let $\mathbf{b}_1$ and $\mathbf{b}_2$ be arbitrary elements of $C(A)$. Then, by definition of $C(A)$, there are elements $\mathbf{x}_1$ and $\mathbf{x}_2$ of $\mathbb{R}^n$ such that

$$A\mathbf{x}_1 = \mathbf{b}_1 \quad \text{and} \quad A\mathbf{x}_2 = \mathbf{b}_2.$$

Therefore,

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{b}_1 + \mathbf{b}_2.$$

Thus, $\mathbf{b}_1 + \mathbf{b}_2 \in C(A)$, by definition of $C(A)$.

(3) Let $\mathbf{b}$ be an arbitrary element of $C(A)$ and let s be an y real number. As $\mathbf{b} \in C(A)$, there is a vector $\mathbf{x}$ in $\mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$, by defintion of $C(A)$. Therefore,

$$A(s\mathbf{x}) = s(A\mathbf{x}) = s\mathbf{b}.$$

Thus, $s\mathbf{b} \in C(A)$, by definition of $C(A)$. □

Example 2.11. Let

$$U = \left\{ \begin{bmatrix} 2x + 3y \\ y - x \\ y \end{bmatrix} \in \mathbb{R}^3 : s, t \in \mathbb{R}, \right\}.$$

By the identity

$$\begin{bmatrix} 2x + 3y \\ y - x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

and the defintion of column space, we have

$$U = C\left(\begin{bmatrix} 2 & 3 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}\right).$$

Therefore, by Theorem 2.10, U is a subspace of $\mathbb{R}^3$.

Example 2.12. Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 12 & 11 & 12 \end{bmatrix}.$$

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

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = s\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}.$$

By definition of nullspace,

$$N(A) = \left\{ s\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\}$$

By the identity

$$s\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix}$$

and the definitions of nullspace and column space, it follows that

$$N(A) = C\left(\begin{bmatrix} 1 & 2 \\ -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

Remarks 2.13.

- Example 2.12 illustrates that "solving the system" $A\mathbf{x} = \mathbf{b}$, in the sense of MATH 211, amounts to expressing the nullspace of A as the column space of another matrix: the one whose columns are the "basic solutions" of $A\mathbf{x} = \mathbf{b}$.
- Later we will show how to express a column space as a nullspace.

Exercise 2.14. Prove that $\{\mathbf{0}\}$ and $\mathbb{R}^m$ are subspaces of $\mathbb{R}^m$.

Exercise 2.15. Let U be a subset of $\mathbb{R}^m$. Prove that U is a subspace of $\mathbb{R}^m$ if and only if U satisfies (1′), below, as well as properties (2), and (3) of Definition 2.1.

(1′) U is nonempty.

3. Linear combinations

Let

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \cdots \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

The theory of the equation $A\mathbf{x} = \mathbf{b}$ — an algebraic theory, so far — can be understood in terms of the geometry of the column vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of A and the vector $\mathbf{b}$. This geometric perspective is well worth developing: It lets us to apply algebraic methods to geometry. Even more valuable, perhaps, it gives us a framework for thinking geometrically (visually) about algebra. There is no free lunch, however: In developing a geometric point of view on linear algebra, we incur some overhead, mainly in the form of new terminology.

Definition 3.1. A *linear combination of the vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is an expression of the form

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n,$$

where $t_j \in \mathbb{R}$.

3.1. **Examples.**

Example-Definition 3.2. The zero vector can be expressed as a linear comination of any sequence of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$:

$$\mathbf{0} = 0\mathbf{a}_1 + \cdots + 0\mathbf{a}_n.$$

The expression on the right hand side is called the *trivial linear combination* of $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Example-Definition 3.3. Let $\mathbf{i}_j \in \mathbb{R}^m$ be the j-th column of the identity matrix $I \in \mathbb{R}^{m \times m}$:

$$\mathbf{i}_j := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow j\text{-th row.} \\ \\ \\ \\ \end{array}$$

It's called the *j-th standard basis vector of* $\mathbb{R}^m$.

Any vector $\mathbf{b} \in \mathbb{R}^m$ can be expressed as a linear combination of these:

$$\begin{aligned}
\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} &= \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ b_m \end{bmatrix} \\
&= b_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + b_m \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \\
&= b_1\mathbf{i}_1 + b_2\mathbf{i}_2 + \cdots + b_m\mathbf{i}_m
\end{aligned}$$

3.2. **Linear combinations are matrix-vector products.** Here's a trivial, yet crucial, observation: By the definition of matrix multiplication,

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = A\mathbf{x}. \tag{1}$$

Remember this principle:

Linear combinations are just matrix-vector products.

We apply it right away:

Theorem 3.4. *The column space $C(A)$ is the set of all linear combinations of the column vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of A:*

$$C(A) = \{x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n : x_1, \ldots, x_n \in \mathbb{R}\}.$$

Proof.

$$\begin{aligned} C(A) &= \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} && \text{(by definition of } C(A)) \\ &= \{x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n : x_1, \ldots, x_n \in \mathbb{R}\} && \text{(by (1)).} \end{aligned}$$ □

4. Span

Definitions 4.1. The *span* of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$, written $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$, is the set of their linear combinations:

$$\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle := \{x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n : x_1, \ldots, x_n \in \mathbb{R}\}.$$

Example 4.2. Let $\mathbf{i}_1, \ldots, \mathbf{i}_m$ be the standard basis of $\mathbb{R}^m$, as in Example-Definition 3.3. As every vector $\mathbf{b} \in \mathbb{R}^m$ can be expressed as a linear comnination of $\mathbf{i}_1, \ldots, \mathbf{i}_m$, we have

$$\langle \mathbf{i}_1, \ldots, \mathbf{i}_m \rangle = \mathbb{R}^m.$$

Example 4.3. By Key fact 5.3 and Definition 1.1, we have:

$$C(A) = \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle. \tag{2}$$

Theorem 4.4. *Let A be its matrix and let $R = \operatorname{rref} A$.*

(1) *The equation $A\mathbf{x} = \mathbf{b}$ has a solution for all vectors $\mathbf{b}$ if and only if every row of R has a leading one.*
(2) *The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, $\mathbf{x} = \mathbf{0}$, if and only if every column of R has a leading one.*

Theorem 4.4 suggests taking a closer look at the following two sets of vectors associated to a matrix $A \in \mathbb{R}^{m\times n}$:

Definitions 4.5.

(1) The set of all vectors $\mathbf{b} \in \mathbb{R}^m$ such $A\mathbf{x} = \mathbf{b}$ has a solution is called the *column space of* A and written $C(A)$:
$$C(A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}.$$
(2) The set of all vectors $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{0}$ is called the *nullspace of* A and written $N(A)$:
$$N(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

Corollary 4.6. *Let $A \in \mathbb{R}^{m\times n}$.*

(1) *If $A\mathbf{x} = \mathbf{b}$ has a solution for all $\mathbf{b} \in \mathbb{R}^m$, i.e., if*
$$C(A) = \mathbb{R}^m,$$
then $m \le n$.

(2) *If $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, i.e., if*
$$N(A) = \{\mathbf{0}\},$$
then $n \le m$.

Corollary 4.7. *The following are equivalent for a square matrix $A \in \mathbb{R}^{m\times m}$.*

(1) *$A\mathbf{x} = \mathbf{b}$ has a solution for all vectors $\mathbf{b}$.*
(2) $C(A) = \mathbb{R}^m$
(3) *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
(4) $N(A) = \{\mathbf{0}\}$
(5) $\operatorname{rref} A = I$
(6) *A is invertible.*

Later, we'll give an intrinsic definition of the following essential notion:

Provisional definition 4.8. A *subspace of* $\mathbb{R}^m$ is a subset of the form $C(A)$, where $A \in \mathbb{R}^{m\times n}$, or $N(B)$, where $B \in \mathbb{R}^{\ell\times m}$.

Row equivalence

Uniqueness of rref

Pivot columns

5. Linear combinations

Definition 5.1. A *linear combination* of vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ is an expression of the form
$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n,$$
where $t_j \in \mathbb{R}$.

Example-Definition 5.2. The zero vector can be written as a linear comination of any sequence of vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$:
$$\mathbf{0} = 0\mathbf{a}_1 + \cdots + 0\mathbf{a}_n.$$
The expression on the right hand side is called the *trivial linear combination* of $\mathbf{a}_1, \dots, \mathbf{a}_n$.

Here's a trivial, yet crucial, observation. Let
$$A = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix} \in \mathbb{R}^{m\times n}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$
Then, by the definition of matrix multiplication,
$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = A\mathbf{x}.$$

We record (a paraphrase of) this simple fact, prominently, for ease of reference.

Key fact 5.3. Linear combinations are just matrix-vector products.

6. Span

Definitions 6.1. The *span* of a sequence $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of vectors in $\mathbb{R}^m$, written $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$, is the set of all linear combinations of these vectors:

$$\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle := \{x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n : x_1, \ldots, x_n \in \mathbb{R}\}.$$

We say $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *span* $\mathbb{R}^m$ if every vector $\mathbf{b} \in \mathbb{R}^m$ can be written as such a linear combination, i.e., if

$$\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle = \mathbb{R}^m.$$

If

$$A = \begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{m \times n},$$

we abbreviate to $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$ to $C(A)$, and refer to this set as the *column space* of A:

$$C(A) := \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle.$$

Note that, by Key Fact 5.3,

$$C(A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}.$$

Theorem 6.2. *The following are equivalent for a matrix* $A \in \mathbb{R}^{m \times n}$.

(1) $C(A) = \mathbb{R}^m$
(2) *The equation* $A\mathbf{x} = \mathbf{b}$ *has a solution for all* $\mathbf{b}$.
(3) rref A *has a leading one in every row.*

Moreover, if any (hence, all) of these conditions hold, then $m \leq n$.

7. Linear independence

Definition 7.1. Vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ are *linearly independent* if the only linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ that equals $\mathbf{0}$ is the trivial one (see Example-Definition 5.2).

8. Reduced row echelon form

Definition 8.1. A matrix R is in *reduced row echelon form* if:

(1) All zero rows of R are at the bottom.
(2) Every nonzero row has a leading one.
(3) A leading one is the right of those in the rows above it.
(4) A leading one is the only nonzero entry in its column.

Definition 8.2. Let $A, B \in \mathbb{R}^{m \times n}$. We say that A *and* B *are row equivalent* or that A *is row equivalent to* B if there is an *invertible* matrix $\gamma \in \mathbb{R}^{m \times m}$ such that $\gamma A = B$.

Remarks 8.3.

(1) Since $\gamma A = B$ if and only if $\gamma^{-1}B = A$, the notion of row equivalence is symmetric in A and B: A is row equivalent to B if and only if B is row equivalent to A.
(2) A and B are row equivalent if and only if A can be transformed into B via a sequence of elementary row operations. The theory of elementary matrices connects this formulation of row equivalence with that of Definition 8.2.

Theorem 8.4. *A matrix A is row equivalent to a unique matrix in reduced row echelon form.*

9. Uniqueness of reduced row echelon form

(Under construction; do not read!)

Lemma 9.1. *Let $A = \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots \mathbf{a}_n\end{pmatrix}$ be an $m \times n$ matrix reduced row echelon form with pivot columns $j_1 < j_2 < \ldots < j_r$. The following are equivalent:*

(1) $\mathbf{a}_j$ *is a pivot column of A.*
(2) $\mathbf{a}_j \notin \operatorname{span}\{\mathbf{a}_{j_i} : j_i < j\}$.
(3) $\mathbf{a}_j \notin \operatorname{span}\{\mathbf{a}_k : k < j\}$.

Proof. Let $A_j = \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_j\end{pmatrix}$. Then A_j is in reduced row echelon form as A is (reference an exercise?) and $\{\mathbf{a}_{j_i} : j_i < j\}$ is a basis for $C(A_j)$ by Theorem **??**. In particular, if $1 \le p \le r$ then $\{\mathbf{a}_{j_i} : 1 \le i \le p\}$ is a basis of $C(A_{j_p})$. Since bases are linearly independent,

$$\mathbf{a}_{j_p} \notin \operatorname{span}\{\mathbf{a}_{j_i} : 1 \le i < p\},$$

proving that (1) implies (2).

Conversely, suppose $\mathbf{a}_j$ is not a pivot column of A. Now the column space of A_j is spanned by its pivot columns, namely, the $\mathbf{a}_{j_i}$ with $j_i < j$, so $\mathbf{a}_j \in \operatorname{span}\{\mathbf{a}_{j_i} : j_i < j\}$. This proves that (2) implies (1).

To see that (2) and (3) are equivalent, observe that $\operatorname{span}\{\mathbf{a}_k : k < j\}$ is by definition, the column space of A_j while $\operatorname{span}\{\mathbf{a}_{j_i} : j_i < j\}$ is the span of the pivot columns of A_j. But these spans are equal by Theorem **??**. $\square$

Exercise 9.2. Let A be an $m \times n$ matrix in row echelon form whose leading 1s lie in positions (i, j_i), $1 \le i \le r$. Suppose $j_i < j < j_{i+1}$. Prove that $a_{pj} = 0$ for $i < p \le m$.

Theorem 9.3. *Let A and B be row equivalent matrices in reduced row echelon form. Then $A = B$.*

Proof. Suppose not. Then there is a pair of $m \times n$ matrices (A, B) witnessing the failure of the theorem such that n minimal among all such witnesses. In other words:

(i) A and B are distinct, row equivalent matrices in reduced row echelon form.
(ii) If A' and B' is another such pair then A has fewer columns than A'.

Since A and B are row equivalent and in reduced row echelon form, so are A_{n-1} and B_{n-1}. (Why?) By the minimality of n, the theorem holds for A_{n-1} and B_{n-1}. Therefore, $A_{n-1} = B_{n-1}$. Thus, to prove the theorem, it suffices to show that $\mathbf{a}_n = \mathbf{b}_n$.

Suppose, first, that $\mathbf{a}_n$ is a pivot column of A. Then $\mathbf{a}_n \notin C(A_{n-1})$ by Lemma 9.1. Since A and B are row equivalent, $\mathbf{a}_n \in C(A_{n-1})$ if and only if $\mathbf{b}_n \in C(B_{n-1})$. (Row equivalence preserves linear relations between columns.) Therefore, $\mathbf{b}_n \notin C(B_{n-1})$ and, by Lemma 9.1 again, $\mathbf{b}_n$ is a pivot column of B. Let $r = \operatorname{rank} A_{n-1} = \operatorname{rank} B_{n-1}$. Then $\mathbf{a}_n$ and $\mathbf{b}_n$ must both equal $\mathbf{e}_{r+1}$ because A and B are in reduced row echelon form. Thus, $\mathbf{a}_n = \mathbf{b}_n$.

Now suppose that $\mathbf{a}_n$ is not a pivot column of A. Then $\mathbf{a}_n \in C(A_{n-1})$ by Lemma 9.1, so there is a vector $\mathbf{x} \in \mathbb{R}^{n-1}$ such that $\mathbf{a}_n = A_{n-1}\mathbf{x}$. Since A and B are row equivalent, $\mathbf{b}_n = B_{n-1}\mathbf{x}$. But $A_{n-1} = B_{n-1}$, so it follows once again that $\mathbf{a}_n = \mathbf{b}_n$. $\square$