CHAPTER 1

# Vector Spaces over a Field $\mathbb{K}$

## 1.1. Vector Spaces

The objects of interest in this chapter will be vector spaces over arbitrary fields.

**Basic Properties and Examples.**

DEFINITION 1.1. A *vector space over a field* $\mathbb{K}$ is a set $V$ equipped with two operations $(+)$ and $(\cdot)$ from $V \times V \to V$ and $\mathbb{K} \times V \to V$ having the following properties:

1. **Axioms for** $(+)$**:** For all $x, y \in V$ we have:
   COMMUTATIVE LAW: $x + y = y + x$
   ASSOCIATIVE LAW: $(x + y) + z = x + (y + z)$
   ZERO ELEMENT: There exists an element "0" in $V$ such that $0 + v = v$ for all $v$.
   ADDITIVE INVERSE: For every $x \in V$, there is an element $v \in V$ such that $v + x = 0$.
2. **Axioms for** $(\cdot)$**:** For all $v, w \in V$ and $a, b \in \mathbb{K}$ we have:
   IDENTITY LAW: $1 \cdot v = v$ ($1 =$ the multiplicative identity in $\mathbb{K}$)
   ASSOCIATIVE LAW: $(a \cdot b) \cdot v = a \cdot (b \cdot v)$
   DISTRIBUTIVE LAW: $a \cdot (v + w) = (a \cdot v) + (a \cdot w)$
   DISTRIBUTIVE LAW: $(a + b) \cdot v = (a \cdot v) + (b \cdot v)$

NOTATION: We generally write "0" for both the zero vector in $V$ and the scalar "0" in the field $\mathbb{K}$. Usually this causes no trouble, but when it becomes necessary to distinguish the two we will write $0_V$ for the zero element in $V$. The "additive inverse" of $x \in V$ is written "$-x$."

As a consequence we have:

LEMMA 1.2. *The zero element is unique: if* $0, 0' \in V$ *are elements such that* $0 + v = v$ *and* $0' + v = v$*, for all* $v \in V$*, then* $0' = 0$.

PROOF. $0 + 0' = 0'$ and $0 + 0' = 0$, so $0' = 0$. □

LEMMA 1.3. *The additive inverse is unique. That is, given* $v \in V$ *there is just one element* $u \in V$ *such that* $u + v = 0$.

PROOF. Suppose $v \in V$ and we are given $u$ and $u'$ with $u + v = 0$ and $u' + v = 0$. Look at the combination $u + v + u'$; by associativity we get

$$u' = 0 + u' = (u + v) + u' = u + (v + u') = u + 0 = u \ ,$$

so that $u' = u$. □

EXERCISE 1.4. From the axioms and previous results prove:

(a) $0 \cdot v = 0_V$
(b) $\lambda \cdot 0_V = 0_V$
(c) $\lambda \cdot v = v$ and $v \neq 0_V \Rightarrow \lambda = 1$

for all $v \in V$ and $\lambda \in \mathbb{K}$, where "0" is the zero element in the field of scalars $\mathbb{K}$.

EXERCISE 1.5. Prove that $-v = (-1) \cdot v$ where $-1$ is the negative of $1 \in \mathbb{K}$.

HINT: $1+(-1) = 0$ in $\mathbb{K}$ and $0 \cdot v = 0_V$. Remember: "$-v$" is the unique element that added to $v$ is $0_V$; prove that $(-1) \cdot v$ has this property and conclude that $(-1) \cdot v = -v$ by uniqueness of the additive inverse.

EXERCISE 1.6. Prove that $-(-v) = v$, for all $v \in V$.

HINT: Same as the previous exercise: what identity does $-(-v)$ satisfy?

EXERCISE 1.7 (Cancellation Laws). If $a + v = a + w$ for $a, v, w \in V$, prove that $v = w$. Then use this to prove:

(a) $\lambda \cdot v = 0_V$ and $v \neq 0_V$ implies that $\lambda = 0$ in $\mathbb{K}$ .
(b) $\lambda \cdot v = v$ and $v \neq 0_V$ implies $\lambda = 1$.

DEFINITION 1.8. *Coordinate space* over the field $\mathbb{K}$ consists of all ordered $n$-tuples $\mathbb{K}^n = \{\mathbf{x} = (x_1, \dots, x_n) : x_k \in \mathbb{K}\}$, equipped with the usual $(+)$ and $(\cdot)$ operations:

1. $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$,
2. $\lambda \cdot (x_1, \dots, x_n) = (\lambda \cdot x_1, \dots, \lambda \cdot x_n)$ for $\lambda \in \mathbb{K}$.

EXERCISE 1.9. explain why $(+)$ in $\mathbb{R}^2$ is described geometrically by the "parallelogram law" for vector addition shown in Figure 1.1.

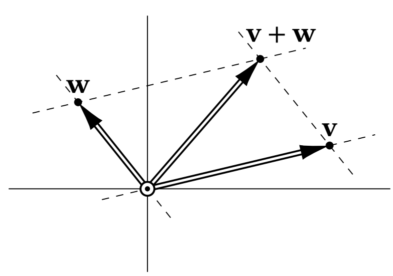


FIGURE 1.1. The parallelogram law for vector addition, illustrated in $\mathbb{R}^2$.

EXAMPLE 1.10. The space $\mathrm{M}(n \times m, \mathbb{K})$ of $n \times m$ matrices with entries in $\mathbb{K}$ becomes a vector space when equipped with the operations

ADDITION: $(A + B)_{ij} = A_{ij} + B_{ij}$
SCALING: $(\lambda \cdot A)_{ij} = \lambda A_{ij}$

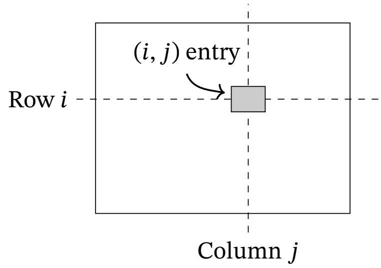


FIGURE 1.2. The entry in a matrix array with "address" $(i, j)$ is the one in Row $i$ and Column $j$.

The space of square matrices, with $m = n$, is denoted $\mathrm{M}(n, \mathbb{K})$.
NOTE: Matrix entry $A_{ij}$ is the one in the $i^{\text{th}}$ row and $j^{\text{th}}$ column. The pair $(i, j)$ is referred to as its "address"; see Figure 1.2.

There is also a matrix multiplication that makes $\mathrm{M}(n, \mathbb{K})$ an associative-multiplication that makes $\mathrm{M}(n, \mathbb{K})$ an associative algebra with identity, but the matrix product $AB$ can be defined more generally for nonsquare matrices as long as they are "compatible," with the number of columns in $A$ equal to the number of rows in $B$. Thus if $A$ is $m \times q$ and $B$ is $q \times n$ we get an $m \times n$ matrix $AB$ with entries

$$(AB)_{ij} = \sum_{k=1}^{q} A_{ik} B_{kj}.$$

The algebra $\mathrm{M}(n, \mathbb{K})$ of square matrices is not commutative unless $n = 1$.

EXAMPLE 1.11 (polynomial ring $\mathbb{K}[x]$). The set $\mathbb{K}[x]$ consists of all finite "formal sums" $a_0 + a_1 x + \cdots + a_n x^n + \cdots = \sum_{k \geq 0} a_k x^k$ with $a_i \in \mathbb{K}$, and $a_i = 0$ for all but a finite number of indices. These sums can have arbitrary length. They include the "constant polynomials," which have the form $c \cdot \mathbb{1}$ with $c \in \mathbb{K}$, where $\mathbb{1}$ is the particular constant polynomial $1 + 0 \cdot x + 0 \cdot x^2 + \cdots$; the zero polynomial $0 \cdot \mathbb{1}$ is written as "0", which might get confusing.

The algebraic operations in $\mathbb{K}[x]$ are

1. ADDITION:

$$\Big(\sum_{k \geq 0} a_k x^k\Big) + \Big(\sum_{k \geq 0} b_k x^k\Big) = \sum_{k \geq 0} (a_k + b_k) x^k$$

2. SCALING:

$$\lambda \cdot \Big(\sum_{k \geq 0} a_k x^k\Big) = \sum_{k \geq 0} (\lambda a_k) x^k$$

There is also a multiplication operation, obtained by multiplying terms in the formal sums and gathering together those of the same degree:

3. PRODUCT:

$$\Big(\sum_{k \geq 0} a_k x^k\Big) \times \Big(\sum_{l \geq 0} b_l x^l\Big) = \sum_{k,l \geq 0} a_k b_l \, x^{k+l} = \sum_{r \geq 0} \Big(\sum_{\substack{k,l \geq 0 \\ k+l=r}} a_k b_l\Big) \cdot x^r$$

(the sum being finite for each $r$). This makes $\mathbb{K}[x]$ into a commutative associative algebra over $\mathbb{K}$ with $\mathbb{1}$ as its multiplicative identity.

All information about a polynomial resides in the symbol string $(a_0, a_1, a_2, \dots)$ of coefficients, and the algebraic operations on $\mathbb{K}[x]$ can be performed as operations on symbol strings; the zero polynomial is represented by $(0, 0, \dots)$, the identity by $\mathbb{1} = (1, 0, \dots, 0)$, and $x$ by $x = (0, 1, 0, \dots.)$, etc.

EXERCISE 1.12. If $f(x) = 3 + 3x + x^2$ and $g(x) = 4x^2 - 2x^3 + x^5$, compute the sum $f + g$ and product $f \cdot g$.

The *degree* $\deg(f)$ of $f = \sum_{k\geq 0} a_k x^k$ is $n$ if $a_n \neq 0$ and $a_k = 0$ for all $k > n$. The degree of a constant polynomial $c\mathbb{1}$ is zero, except that no "degree" can be assigned to the zero polynomial 0. (For various reasons, the only possible assignment would be "$-\infty$".)

EXERCISE 1.13. If $f, g \neq 0$ in $\mathbb{K}[x]$, prove that $fg \neq 0$ and $\deg(fg) = \deg(f) + \deg(g)$.

EXERCISE 1.14. If $f, g \neq 0$ in $\mathbb{K}[x]$, what (if anything) can you say about $\deg(f + g)$?

**Multivariate Polynomials and Multi-Index Notation.**

EXAMPLE 1.15. (Polynomials in several unknowns). The polynomial ring $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \dots, x_n]$is handled using very efficient "multi-index notation." A *multi-index* is an element $\alpha = (\alpha_1, \dots, \alpha_n)$ of the Cartesian product set $\mathbb{Z}_+^n = \mathbb{Z}_+ \times \cdots \times \mathbb{Z}_+$ ($n$ factors). Each $\alpha$ determines a *monomial* $x^\alpha = x_1^{\alpha_1} \cdot \cdots \cdot x_n^{\alpha_n}$, in which we interpret $x_k^0 = 1$. Elements of $\mathbb{K}[x_1, \dots, x_n]$ are finite formal linear combinations of monomials

$$f(x_1, \dots, x_n) = \sum_{\alpha \in \mathbb{Z}_+^n} c_\alpha x^\alpha \quad (c_\alpha \in \mathbb{K}).$$

The monomial $x^{(0,\dots,0)}$ is the constant polynomial $\mathbb{1}$ in $\mathbb{K}[x_1, \dots, x_n]$. With these ideas in mind we define the following:

1. The *total degree* of a multi-index is $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and the degree of the corresponding monomial is $\deg(x^\alpha) = |\alpha|$. Many monomials can have the same total degree, for example $x^2y$ and $xy^2$.
2. The *degree of a polynomial* $f \in \mathbb{K}[\mathbf{x}]$ is

$$\deg(f) = \max\{|\alpha| : c_\alpha \neq 0\}.$$

Nonzero constant polynomials $c\mathbb{1}$ have degree zero; if $f$ is the zero polynomial (all coefficients $c_\alpha = 0$) $\deg(f)$ cannot be defined. The generators $f_k(\mathbf{x}) = x_k$ of the polynomial ring all have degree 1.

The following operations make $V = \mathbb{K}[x_1, \dots, x_n]$ a vector space and a commutative associative algebra with identity $\mathbb{1} = x^{(0,\dots,0)}$.

1. SUM: $f + g = \sum_\alpha (a_\alpha + b_\alpha)\, x^\alpha$
2. SCALING: $\lambda \cdot f = \sum_\alpha (\lambda a_\alpha) x^\alpha$

3. PRODUCT OPERATION:

$$f \cdot g = \Big(\sum_\alpha a_\alpha x^\alpha\Big) \cdot \Big(\sum_\beta b_\beta x^\beta\Big) = \sum_{\alpha,\beta\in\mathbb{Z}^n_+} a_\alpha b_\beta x^{\alpha+\beta}$$
$$= \sum_{\gamma\in\mathbb{Z}^n_+}\Big(\sum_{\alpha+\beta=\gamma} a_\alpha \cdot b_\beta\Big) \cdot x^\gamma$$

where we define $\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$.

As an example, the monomials of degree 2 in $\mathbb{K}[x_1, x_2, x_3]$ are

| multi-index $\alpha$ | monomial $x^\alpha$ | dictionary word |
|---|---|---|
| $(0, 0, 2)$ | $x_3^2$ | AAC |
| $(0, 1, 1)$ | $x_2x_3$ | ABA |
| $(0, 2, 0)$ | $x_2^2$ | ACA |
| $(1, 0, 1)$ | $x_1x_3$ | BAB |
| $(1, 1, 0)$ | $x_1x_2$ | BBA |
| $(2, 0, 0)$ | $x_1^2$ | CAA |

Here we have lined up the monomials in "lexicographic"or "dictionary" order (taking $A = 0$, $B = 1$, $C = 2$, ...), which is a useful way to manage them. This is a strict *linear* ordering of monomials; they are only partially ordered by their "total degree" $\deg(x^\alpha) = |\alpha|$. The system $\mathbb{K}[x_1, \dots, x_n]$ is a commutative associative algebra with identity element $\mathbb{1}$. Its properties are quite a bit more complicated than those of polynomials $\mathbb{K}[x]$ in one unknown, but they do share two crucial algebraic properties.

EXERCISE 1.16. (Hard, but try it.) If $f, g \neq 0$ in $\mathbb{K}[x_1, \dots, x_n]$ prove that:

1. DEGREE FORMULA: $\deg(f \cdot g) = \deg(f) + \deg(g)$ for all $f, g \neq 0$ in $\mathbb{K}[x_i, \dots, x_n]$.
2. NO ZERO DIVISORS: $f, g \neq 0$ in $\mathbb{K}[x_1, \dots, x_n] \Rightarrow f \cdot g \neq 0$. This implies we can perform "cancellation": if $f \neq 0$ and $f \cdot h_1 = f \cdot h_2$, then $h_1 = h_2$.

HINT: Try it first for $n = 1$. For $n = 2$ try lexicographic ordering of monomials in $\mathbb{K}[x, y]$.

NOTE: The maximum possible degree for a nonzero monomial in the product $fg$ is obviously $d = \deg(f) + \deg(g)$. The problem is that the coefficient $c_\gamma$ of such a monomial will be a sum of products $(\sum_{\alpha+\beta=\gamma} a_\alpha b_\beta)$, and not a simple product as it is when there is just one variable. Such sums could equal zero even if all terms are nonzero, so why couldn't these coefficients (sums) be zero for *all* monomials with the maximum possible degree $d$, making $\deg(fg) < \deg(f) + \deg(g)$?

A more complete discussion of the degree formula for $n \geq 2$, and especially its proof using lexicographic ordering of monomials, is provided in the appendix of this chapter. I-1.17

EXAMPLE 1.17 (Function Spaces). If $S$ is a set, $\mathcal{C}(S)$ = all scalar-valued functions $f : S \to \mathbb{K}$ becomes a vector space under the usual operations

$$(f+g)(x) = f(x) + g(x), \quad (\lambda \cdot f)(x) = \lambda f(x) \quad \forall x \in S.$$

There is also a pointwise multiplication operation

$$(f \cdot g)(x) = f(x) \cdot g(x),$$

which makes $\mathcal{C}(S)$ a commutative associative algebra over $\mathbb{K}$ with identity element $\mathbb{1}(x) = 1$ for all $x$ and zero element $0(x) = 0$ for $x \in S$.

EXAMPLE 1.18 (Polynomial Functions vs. Formal Sums). The *polynomial functions* $\mathcal{P}_{\mathbb{K}}$ with values in $\mathbb{K}$ are the functions $\phi_f : \mathbb{K} \to \mathbb{K}$ of the form

$$\phi_f(t) = [f(x)|_{x=t}] = \sum_{k \geq 0} a_k t^k \quad (t \in \mathbb{K})$$

for some $f(x) \in \mathbb{K}[x]$. (Thus, $\phi(t) = \sin(t)$ is not a polynomial function on $\mathbb{K}$.) Note carefully that the elements of $\mathcal{P}_{\mathbb{K}}$ are functions, while $\mathbb{K}[x]$ is made up of symbol strings or formal sums. They are not the same thing, though there is a close relation between them implemented by the surjective (= "onto") mapping $\Phi : \mathbb{K}[x] \to \mathcal{P}_{\mathbb{K}}$ such that

$$\Phi f(t) = \sum_{k \geq 0} a_k t^k \quad (t \in \mathbb{K})$$

if $f(x) = \sum_{k \geq 0} a_k x^k$ in $\mathbb{K}[x]$. This surjective map is a *homomorphism*: it preserves, or "intertwines," the algebraic operations in $\mathbb{K}[x]$ and in the "target space" $\mathcal{P}_{\mathbb{K}}$, so that

$$\Phi(\lambda \cdot f) = \lambda \cdot \Phi(f), \quad \Phi(f+g) = \Phi(f) + \Phi(g), \quad \Phi(f \cdot g) = \Phi(f) \cdot \Phi(g).$$

EXERCISE 1.19. If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, explain why $\Phi$ is a bijection, hence an "isomorphism" between commutative associative algebras. In fact, prove that this is so for polynomials over any *infinite* field $\mathbb{K}$.

HINT: $\Phi$ is linear, hence being one-to-one is equivalent to saying that $\Phi(f) = 0 \Rightarrow f = 0$ in $\mathbb{K}[x]$. If $f$ is nonzero in $\mathbb{K}[x]$, the corresponding polynomial function $\Phi(f) : \mathbb{K} \to \mathbb{K}$ can take on the value zero at no more than $n = \deg(f)$ points—i.e., the number of roots in $\mathbb{K}$ cannot exceed $\deg(f)$. Since $\mathbb{R}$ and $\mathbb{C}$ (and even $\mathbb{Q}$) are infinite, we cannot have $\Phi(f) \equiv 0$ on these fields unless $f$ is the zero polynomial.

The finite fields $\mathbb{Z}_p$ ($p$ a prime) are widely used in number theory, cryptography, image processing, etc. The one-to-one correspondence between formal polynomials and scalar-valued functions $\Phi f(t)$ defined on $\mathbb{K}$ breaks down for these fields. For example, if $\mathbb{K} = \mathbb{Z}_p$, the nonzero polynomial $f = x^p - x$ has value zero for every choice of $x \in \mathbb{Z}_p$, so $f$ has $p = \deg(f)$ distinct roots in $\mathbb{Z}_p$. (By a theorem of Fermat, if $p$ is a prime, then $t^{p-1} = 1$ for all nonzero $t$ in $\mathbb{Z}_p$. Thus $t^p - t$ is zero at *every* $t \in \mathbb{Z}_p$ and the function $\Phi(f)$ is the zero function on $\mathbb{Z}_p$.)

EXERCISE 1.20. For $p = 3$, verify that $t^3 - t = 0$ for each of the three elements $t = [0], [1], [2]$ in $\mathbb{Z}_3$. But the corresponding element of $\mathbb{Z}_3[x]$ is $f(x) = x^3 - x$, whose symbol string $(0, -1, 0, 1, 0, 0, \ldots)$ differs from that of the zero polynomial in $\mathbb{Z}_3[x]$.

**Row Space and Column Space of a Matrix.**

EXAMPLE 1.21 (Row Space and Column Space). Two vector spaces can be associated with any $n \times m$ matrix $A$ having entries in $\mathbb{K}$:

1. ROW SPACE: $\mathrm{RowSpace}(A) = \{$linear combinations $c_1R_1(A) + \cdots + c_nR_n(A) : c_i \in \mathbb{K}\}$ of the rows $R_i(A) = (a_{i1}, \ldots, a_{im})$, regarded as vectors in $\mathbb{K}^m$.
2. COLUMN SPACE: $\mathrm{ColSpace}(A) = \{$linear combinations $c_1C_1(A) + \cdots + c_mC_m(A) : c_i \in \mathbb{K}\}$ of the columns $C_i(A) = \mathrm{col}(a_{1i}, \ldots, a_{ni})$, regarded as vectors in $\mathbb{K}^n$.

These will play important roles in our study of matrix operations.

## 1.2. Vector Subspaces

DEFINITION 1.22. A nonempty subset $W$ of a vector space $V$ is a *vector subspace* if

1. $W$ is *closed under* $(+)$: $W + W \subseteq W$, so $w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W$.
2. $W$ is *closed under* $(\cdot)$: $\mathbb{K} \cdot W \subseteq W$, so $\lambda \in \mathbb{K}, w \in W \Rightarrow \lambda \cdot w \in W$.

The zero vector $0_V$ then lies in $W$, for if $w \in W$ then $-w = (-1) \cdot w \in W$ and then $0_V = w + (-w) \in W$. Thus $W$ becomes a vector space over $\mathbb{K}$ in its own right under the $(+)$ and $(\cdot)$ operations applied to elements of $W$.

Subspaces of $V$ include the trivial examples $W = (0)$ and $W = V$; all others are "proper" subspaces of $V$.

DEFINITION 1.23. Given a nonempty set $S$ of vectors in $V$, its *linear span* $\langle S \rangle = \mathbb{K}\text{-span}(S)$ is the smallest subspace $W \subseteq V$ such that $W$ contains $S$.

EXERCISE 1.24. If $\{W_\alpha : \alpha \in I\}$ is any family of subspaces in $V$, prove that their intersection $W = \bigcap_{\alpha \in I} W_\alpha$ is also a subspace.

Thus Definition 1.23 makes sense: Given $S$ there is at least one subspace containing $S$, namely $V$. If $E =$ intersection of all subspaces $W$ that contain $S$, then $E$ is a subspace and is obviously the smallest subspace containing $S$. Thus $\mathbb{K}\text{-span}(S)$ exists, even if $V$ is "infinite dimensional," for instance, if $V = \mathbb{K}[x]$.

This "top-down"definition has its uses, but an equivalent "bottom-up" version is often more informative.

LEMMA 1.25. *If $S \neq \varnothing$ in $V$, its linear span is the set of finite sums*

$$\left\{ \sum_{i=1}^{n} a_i v_i \ : \ a_i \in \mathbb{K},\ v_i \in S,\ n < \infty \right\}.$$

PROOF. Let $E = \{\sum_{i=1}^{n} c_i v_i \,:\, n < \infty,\ c_i \in \mathbb{K},\ v_i \in S\}$. Since $S \subseteq \mathbb{K}$-span$(S)$, every finite sum lies in this span, proving $E \subseteq \mathbb{K}$-span$(S)$. For $(\supseteq)$, it is clear that the family $E$ of finite linear combinations is closed under $(+)$ and $(\cdot)$ operations because a linear combination of linear combinations is just one big linear combination of elements of $S$. It is a subspace of $V$, and it contains $S$ because $1 \cdot s = s$ is a (trivial) linear combination. On the other hand, every subspace $W \supseteq S$ must contain all these linear sums, so $S \subseteq E \subseteq W$. Hence $E$ is the smallest subspace containing $S$ and $E = K$-span$(S)$. □

EXERCISE 1.26. If $K = \mathbb{R}$ and $V = \mathbb{R}^3$ show that

$$W = \{x \in \mathbb{R}^3 \,:\, 3x_1 + 2x_2 - x_3 = 0\}$$

is a subspace and $W' = \{x \in \mathbb{R}^3 \,:\, 3x_1 + 2x_2 - x_3 = 1\}$ is not a subspace.

HINT: For one thing the zero vector $0 = (0, 0, 0)$ is not in $W'$. The situation is shown in Figure 1.3.

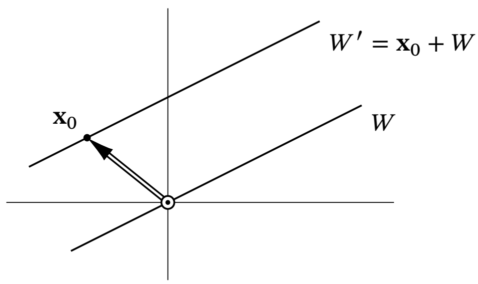


FIGURE 1.3. The subspace $W$ in Exercise 1.26 and a translate $W' = \mathbf{x}_0 + W$ by some $\mathbf{x}_0 \in V$ such that $3x_1^0 + 2x_2^0 - x_3^0 = 1$, for instance, $\mathbf{x}_0 = (0, 1, 1)$. The set $W'$ is not a subspace.

**Systems of Linear Equations and Their Solutions.** Systems of $n$ linear equations in $m$ unknowns are of two general types:

$$\text{Homogeneous}\begin{cases} a_{11}x_1 + \cdots + a_{1m}x_m = 0 \\ \quad\vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m = 0 \end{cases}$$

$$\text{Inhomogeneous}\begin{cases} a_{11}x_1 + \cdots + a_{1m}x_m = b_1 \\ \quad\vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m = b_n \end{cases}$$

with $a_{ij}$ and $b_k$ in $\mathbb{K}$.

EXERCISE 1.27. Verify that the solutions $\mathbf{x} = (x_1, \dots, x_m)$ of the homogeneous system form a vector subspace of $\mathbb{K}^m$. Explain why the solution set of an inhomogeneous system cannot be a vector subspace unless $\mathbf{b} = (b_1, \dots, b_n) = \mathbf{0}$ in $\mathbb{K}^n$.

If we regard vectors $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{K}^m$ as $m \times 1$ column matrices,

$$\mathbf{x} = \text{col}(x_1, \dots, x_m) = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix},$$

you will recognize that the solutions $\mathbf{x} \in \mathbb{K}^m$ of the homogeneous system of equations are precisely the solutions of the matrix equation

$$A\mathbf{x} = \mathbf{0} \quad \text{where the zero vector is} \quad \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_{n\times 1},$$

and for inhomogeneous systems we must solve

$$A\mathbf{x} = \mathbf{b} \quad \text{where } \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}_{n\times 1} \in \mathbb{K}^n.$$

The homogeneous system always has the zero vector $\mathbf{0} \in \mathbb{K}^m$ as a solution, and the solution set $\{\mathbf{x} \in \mathbb{K}^m : A\mathbf{x} = \mathbf{0}\}$ is a vector subspace in $\mathbb{K}^m$. If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, then the number of solutions is either 1 or $\infty$ for this system. An inhomogeneous system might not have any solutions at all; otherwise, it has just one solution or infinitely many.

If $A$ is an $n \times m$ matrix with entries in $\mathbb{K}$, we will find it useful to let $A$ act by left multiplication as an operator $L_A : \mathbb{K}^m \to \mathbb{K}^n$ on column vectors

$$\mathbf{y} = L_A(\mathbf{x}) = A \cdot \mathbf{x} \quad (\text{an } (n \times m) \cdot (m \times 1) \text{ matrix product})$$

for $\mathbf{x} \in \mathbb{K}^m$. This is a *linear operator* in the sense that

$$L_A(\mathbf{x} + \mathbf{y}) = L_A(\mathbf{x}) + L_A(\mathbf{y}) \quad \text{and} \quad L_A(\lambda \cdot \mathbf{x}) = \lambda \cdot L_A(\mathbf{x})$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{K}^m$ and $\lambda \in \mathbb{K}$. Solving a system of linear equations is then equivalent to finding solutions of $L_A(\mathbf{x}) = \mathbf{0}$ or $L_A(\mathbf{x}) = \mathbf{b}$ for $\mathbf{x} \in \mathbb{K}^m$. From this point of view, $A\mathbf{x} = \mathbf{b}$ has solutions if and only if $\mathbf{b}$ lies in the range $R(L_A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{K}^m\}$, which is a vector subspace in $\mathbb{K}^n$. If $\mathbf{b} = \mathbf{0}$ the "homogeneous" equation $A\mathbf{x} = \mathbf{0}$ always has the trivial solution $\mathbf{x} = \mathbf{0}$.

EXERCISE 1.28. If $A$ is an $n \times m$ matrix and $L_A : \mathbb{K}^m \to \mathbb{K}^n$ is defined as above, verify the following:

(a) The *range* $R(L_A) = L_A(\mathbb{K}^m) = \{A \cdot \mathbf{x} : \mathbf{x} \in \mathbb{K}^m\}$ of the operator $L_A$ is a vector subspace in $\mathbb{K}^n$.

(b) The *kernel* of $L_A$,

$$K(L_A) = \ker(L_A) = \{\mathbf{x} \in \mathbb{K}^m : L_A(\mathbf{x}) = A \cdot \mathbf{x} = \mathbf{0} \text{ in } \mathbb{K}^n\},$$

is a vector subspace in $\mathbb{K}^m$.

EXERCISE 1.29 (Inhomogeneous vs. Homogeneous Systems). Given a particular solution $\mathbf{x}_0$ of $A\mathbf{x} = \mathbf{b}$, the full solution set of this equation consists of the vectors $W_{\mathbf{b}} = \mathbf{x}_0 + W$, where $W = \{x \in \mathbb{K}^m : A\mathbf{x} = \mathbf{0}\}$ is the solution set for the homogeneous system $A \cdot \mathbf{x} = \mathbf{0}$. $W$ is a vector subspace of $\mathbb{K}^m$ because

$A \cdot \mathbf{x}_1 = A \cdot \mathbf{x}_2 = \mathbf{0}$ implies $A \cdot (\mathbf{x}_1 + \mathbf{x}_2) = A \cdot \mathbf{x}_1 + A \cdot \mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$ and $A(\lambda \cdot \mathbf{x}) = \lambda A \cdot \mathbf{x} = \lambda \cdot \mathbf{0} = \mathbf{0}$. But $\mathbf{x}_0 + W$ is not a subspace unless $\mathbf{x}_0 \in W$, because it does not contain $\mathbf{0}$.

NOTE: The converse is also true: in $\mathbb{K}^m$ every vector subspace is the solution set of some homogeneous system of linear equations $A\mathbf{x} = \mathbf{0}$, but we are not ready to prove that yet. The situation is shown in Figure 1.4.

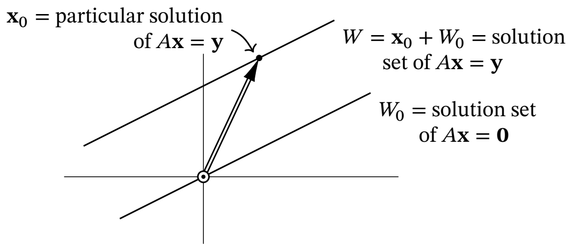


FIGURE 1.4. The subspace $W_0$ is the solution set for a homogeneous equation $A\mathbf{x} = \mathbf{0}$. If the inhomogeneous equation $A\mathbf{x} = \mathbf{b}$ has solutions and if $\mathbf{x}_0$ is a particular solution, so $A\mathbf{x}_0 = \mathbf{b}$, the full solution set $W = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ is the translate $W' = \mathbf{x}_0 + W$ of $W_0$.

This of course presumes that $A\mathbf{x} = \mathbf{b}$ has any solutions at all; if it does not, we say that the system is *inconsistent*. Geometrically, that means $\mathbf{b}$ does not lie in the range $R(L_A)$. Here is an example of an inconsistent inhomogeneous system:

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The corresponding system of linear equations

$$\begin{cases} x_1 + 0 \cdot x_2 = 0, \\ 2x_1 + 0 \cdot x_2 = 1, \end{cases}$$

implies that $x_1 = 0$ and $2x_1 = 1$, an obvious impossibility.

We will continue discussion of linear systems and their solutions via elementary row operations on $A$, or on the $n \times (m+1)$ augmented matrix $[A : \mathbf{b}]$, where $\mathbf{b}$ regarded as an $n \times 1$ column matrix $\mathbf{b} = \mathrm{col}(b_1, \ldots, b_n)$. But first we provide a few more examples of vector spaces we will encounter from time to time.

EXAMPLE 1.30 (Sequence Space $\ell^\infty$). Let $\ell^\infty$ = all bounded sequences of complex numbers $a = (a_1, a_2, \ldots)$ with $a + b = (a_1 + b_1, a_2 + b_2, \ldots)$ and $\lambda \cdot a = (\lambda a_1, \lambda a_2, \ldots)$ for $\lambda \in \mathbb{C}$. This infinite-dimensional space over $\mathbb{C}$ has the following subspaces:

1. $W_0 = \{$sequences such that $a_n \to 0$ as $n \to \infty\}$.
2. $W_n = \{$all sequences of the form $(a_1, \ldots, a_n, 0, 0, \ldots)\}$.
3. $\ell^1 = \{a : \sum_{n=1}^\infty |a_n| < \infty\}$.

**$\mathbf{M}(\mathbf{n}, \mathbb{R})$ and Other Spaces of Matrices.**

EXAMPLE 1.31. Matrix space $\mathrm{M}(n, \mathbb{K})$ contains a few significant subspaces.

1. SYMMETRIC MATRICES:

$$A^{\mathsf{T}} = A \quad \text{where } A^{\mathsf{T}} = (\text{transpose of } A) \text{ with } A^{\mathsf{T}}_{ij} = A_{ji}.$$

2. DIAGONAL MATRICES:

$$D = \begin{pmatrix} d_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & d_n \end{pmatrix}$$

3. BLOCK DIAGONAL MATRICES:

$$D_{m_1,\dots,m_r} = \begin{pmatrix} \boxed{B_{m_1 \times m_1}} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \boxed{B_{m_r \times m_r}} \end{pmatrix}$$

for fixed indices $m_1, \dots, m_r \geq 1$. (The square "blocks" are allowed to have arbitrary entries; all other entries are zero and $m_1 + \cdots + m_r = n$.)

4. UPPER TRIANGULAR and STRICTLY UPPER TRIANGULAR MATRICES:

$$\begin{pmatrix} * & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & * \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & * & & * \\ & \cdot & * & \\ & & \cdot & * \\ & & & \cdot & * \\ 0 & & & & 0 \end{pmatrix} \quad \text{with } * \text{ in } \mathbb{K}.$$

EXERCISE 1.32. Which of these four subspaces, if any, are closed under matrix multiplication as well as $(+)$ ?

EXERCISE 1.33. Show that the vector subspaces of upper triangular and strictly upper triangular matrices are closed under formation of the matrix product $AB$.

EXERCISE 1.34. Show that the vector subspaces of upper triangular (or of strictly upper triangular) matrices are *Lie algebras*: all *commutators* $[A, B] = AB - BA$ are (strictly) upper triangular if $A, B$ are.

EXERCISE 1.35. If an $n \times n$ matrix $A$ has the strictly upper triangular form shown in (a), prove that $A^2$ has the form in (b).

$$\text{(a)}\ A = \begin{pmatrix} 0 & * & & & * \\ & 0 & * & & * \\ & & \ddots & & \\ & & & 0 & * \\ 0 & & & & 0 \end{pmatrix} \qquad \text{(b)}\ A^2 = \begin{pmatrix} 0 & 0 & * & & & & * \\ & 0 & 0 & * & & & * \\ & & 0 & 0 & * & & * \\ & & & \ddots & \ddots & & \\ & & & & \ddots & \ddots & * \\ & & & & & 0 & 0 \\ 0 & & & & & & 0 \end{pmatrix}$$

NOTE: Further computations show that $A^3$ has three diagonal files of zeros so that $A$ is a *nilpotent matrix*, with $A^n = 0_{n\times n}$.

## 1.3. Solving Matrix Equation $A\mathbf{x} = \mathbf{b}$

We have seen that solving a system of $n$ linear equations in $m$ unknowns is equivalent to solving the matrix equation $A\mathbf{x} = \mathbf{b}$ where $A$ is the $n \times m$ array of coefficients in the system. Given such a system several questions arise:

1. *For which vectors* $\mathbf{b}$ *does the system have any solutions at all? This is equivalent to determining the range of the linear map* $L_A(\mathbf{x}) = A\mathbf{x}$ *from* $\mathbb{K}^m \to \mathbb{K}^n$.
2. *When the system* $A\mathbf{x} = \mathbf{b}$ *is consistent, so there are solutions, how can we find them all? What is the connection between the systems* $A\mathbf{x} = \mathbf{b}$ *and* $A\mathbf{x} = \mathbf{0}$*?*

Row operations on matrices are the main tool for resolving such questions. These can also be applied to solve the basic problems of interest in studying vector spaces.

3. *Given vectors* $\{u_1, \dots, u_r\} \subseteq V$ *and* $b \in V$*, decide whether there exist* $c_1, \dots, c_r \in \mathbb{K}$ *such that* $b = \sum_{i=1}^{r} c_i u_i$*—i.e., decide which* $b$ *are in the linear span* $\mathbb{K}\text{-span}\{u_1, \dots, u_r\}$.
4. *Are the vectors* $u_1, \dots, u_r$ *independent, or is there some dependency between them that would allow us to write one of them as a linear combination of the others?*

This last question is equivalent to asking whether some nontrivial linear combination is zero:

$4'$. $c_1u_1 + \cdots + c_ru_r = 0_V \quad (\text{not all } c_i = 0)$

In fact, if $c_i \neq 0$ in ($4'$) there is a nontrivial dependency $u_i = \sum_{j\neq i} -(c_j/c_i)\, v_j$; the converse is trivial.

Questions 3 and 4, posed for an arbitrary vector space $V$, can often be reduced to computations in coordinate space, in which $V = \mathbb{K}^m$ and $\mathbf{b} \in \mathbb{K}^n$, so we will focus on this special case.

**Elementary Operations and Echelon Form of $A\mathbf{x} = \mathbf{b}$.** Before proceeding farther we recall some basic facts about solving matrix equations using elementary row operations. These are based on observations with which you should already be familiar.[1] The simple (but important) verifications are left as exercises.

PROPOSITION 1.36. *The following elementary operations on the rows $R_i(A)$ of an $n \times m$ matrix $A$ do not change the set of solutions $\mathbf{x}$ of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.*

1. $R_i \leftrightarrow R_j$: *switch two rows*
2. $R_i \to \lambda R_i$: *scale* (*Row* $i$) *by some* $\lambda \neq 0$ *in* $\mathbb{K}$
3. $R_i \to R_i + \lambda R_j$: *for* $i \neq j$ *add any scalar multiple of* (*Row* $j$) *to* (*Row* $i$), *leaving* (*Row* $j$) *unaltered.*

*Furthermore, if row operations on a matrix $A$ yield the matrix $A'$, then:*

4. *Each row $R_i(A')$ is a linear combination of the rows of $A$, viewed as vectors in $\mathbb{K}^m$.*
5. *The linear span of the rows of $A'$ is the same as the linear span of the rows of $A$.*

*The same is obviously true if $A'$ results from a sequence of elementary row operations.*

SKETCH OF PROOF. This follows because each row operation is reversible, with $R_i \to R_i - \lambda R_j$ the inverse of $R_i \to R_i + \lambda R_j$. Applied to the $n \times (m+1)$ "augmented matrix" $[A : \mathbf{b}]$ associated with an inhomogeneous system $A\mathbf{x} = \mathbf{b}$, the matrix equation $A'\mathbf{x} = \mathbf{b}'$ corresponding to the modified matrix $[A' : \mathbf{b}']$ has the same solution set as $A\mathbf{x} = \mathbf{b}$ . □

Although row operations do not change the solution set they can greatly simplify the system of equations to be solved, or the task of determining the linear span of some set of vectors in $\mathbb{K}^m$, leading to easy systematic solution of matrix equations. For instance, when $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ it is always possible to find a sequence of row operations that put $A$ into *echelon form* $A'$,

$$\text{ECHELON FORM:}\quad A' = \left(\begin{array}{cccccccccc} \boxed{1} & * & . & . & . & & & & & * \\ & & & \boxed{1} & * & . & . & . & . & * \\ 0 & & & & & \boxed{1} & * & . & . & * \\ \hline & & & & \mathbf{0} & & & & & \end{array}\right).$$

[1]See chapters 1–4 of *Linear Algebra*, by Seymour Lipschutz, Schaum's Outline Series, 1991.

The same moves put the augmented matrix $[A : \mathbf{b}]$ into a similar form:

$$
\text{(1.1)} \qquad [A' : \mathbf{b}'] = \left(\begin{array}{cccccccccc|c}
\boxed{1} & * & . & 0 & * & 0 & & & & * & b'_1 \\
 & & & \boxed{1} & * & 0 & . & . & . & * & \vdots \\
0 & & & & & \boxed{1} & * & . & . & * & b'_r \\
\hline
 & & & & & & & & & & b'_{r+1} \\
 & & & & \mathbf{0} & & & & & & \vdots \\
 & & & & & & & & & & b'_m
\end{array}\right).
$$

However, when we apply the row operations that put $A$ in echelon form $A'$ to the augmented matrix $[A : \mathbf{b}]$, the last column of $[A' : \mathbf{b}']$ need not terminate in a string of zeros.

Solutions of the systems $A'\mathbf{x} = \mathbf{0}$, $A'\mathbf{x} = \mathbf{b}'$ are quickly found by "backsolving" (illustrated below). One could go further, forcing $A$ into even simpler form by knocking out all terms $*$ above the "step corners." These additional operations would of course affect $\mathbf{b}'$ in the augmented matrix yielding the *reduced echelon form*

$$
[A'' : \mathbf{b}''] = \left(\begin{array}{cccccccccc|c}
\boxed{1} & * & . & 0 & * & 0 & & & & * & b''_1 \\
 & & & \boxed{1} & * & 0 & . & . & . & * & \vdots \\
0 & & & & & \boxed{1} & * & . & . & * & b''_r \\
\hline
 & & & & & & & & & & b''_{r+1} \\
 & & & & \mathbf{0} & & & & & & \vdots \\
 & & & & & & & & & & b''_m
\end{array}\right).
$$

The step corners appearing in the echelon displays $A'$ and $A''$ are often referred to as *pivots*, and the columns in which they occur are the *pivot columns*.

The systems $A\mathbf{x} = \mathbf{b}$ and $A'\mathbf{x} = \mathbf{b}'$ have the same solutions. But solutions of $A'\mathbf{x} = \mathbf{b}'$ exist if and only if $b'_{r+1} = \cdots = b'_n = 0$, because the rows lying below the last step corner in the echelon form $[A' : \mathbf{b}']$ correspond to equations

$$b'_{r+1} = 0, \dots, b'_n = 0,$$

which cannot be satisfied if any of the right-hand entries $b'_{r+1}, \dots, b'_m$ in 1.1 are nonzero. (The variables $x_1, \dots, x_m$ don't appear!) To summarize, we have established the following:

> CONSISTENCY CONDITION: *A matrix equation* $A\mathbf{x} = \mathbf{b}$ *has solutions if and only if the echelon form* $[A' : \mathbf{b}']$ *of the augmented matrix* $[A : \mathbf{b}]$ *has only rows of zeros below the last step corner.*

Once the consistency conditions $b'_{r+1} = \cdots = b'_n = 0$ have been met (without which there are no solutions at all), the columns $C_i(A')$ in the echelon form of $A$ that DO NOT include a step corner correspond to "free variables" or "pivot variables" $x_i$ in the solutions of the equation $A'\mathbf{x} = \mathbf{0}$; they are also free variables in solutions of $A'\mathbf{x} = \mathbf{b}'$. If $I = \{1 \le i_1 < \cdots < i_r \le m\}$ are the indices labeling the pivot columns, the REMAINING indices correspond to free variables

$x_k$ ($k \notin I$) in the solutions to $A\mathbf{x} = \mathbf{b}$. Once values have been specified for the free variables, backsolving yields the values of the remaining "dependent" variables $x_k$ ($k \in I$). We get a unique solution of $A'\mathbf{x} = \mathbf{b}'$ for *every* choice of the free variables ($k \notin I$); different choices yield different solutions and all solutions are accounted for. By Proposition 1.36 these are also the solutions of the original equation $A\mathbf{x} = \mathbf{b}$. This concludes our review of row operations.

EXAMPLE 1.37 (A Case Study). Consider the vectors in $\mathbb{K}^3$

$$\mathbf{u}_1 = \begin{pmatrix}1\\2\\1\end{pmatrix},\ \mathbf{u}_2 = \begin{pmatrix}-2\\-4\\-2\end{pmatrix},\ \mathbf{u}_3 = \begin{pmatrix}0\\2\\3\end{pmatrix},\ \mathbf{u}_4 = \begin{pmatrix}2\\0\\-3\end{pmatrix},\ \mathbf{u}_5 = \begin{pmatrix}-3\\8\\16\end{pmatrix},$$

and let $A$ be the matrix with these vectors as its columns

$$A = \begin{pmatrix}1 & -2 & 0 & 2 & -3\\2 & -4 & 2 & 0 & 8\\1 & -2 & 3 & -3 & 16\end{pmatrix}.$$

We will show how to solve $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$.

DISCUSSION: By definition of row/column matrix multiplication, finding a solution $\mathbf{x} = \mathrm{col}(x_1, \dots, x_5) \in \mathbb{R}^5$ of the matrix equation $A\mathbf{x} = \mathbf{b}$ for an arbitrary vector $\mathbf{b} = \mathrm{col}(b_1, b_2, b_3)$ in $\mathbb{R}^3$ is the same as finding coefficients $x_1, \dots, x_5$ such that

$$\mathbf{b} = \sum_{i=1}^{5} x_i\mathbf{u}_i = x_1\begin{pmatrix}1\\2\\1\end{pmatrix} + x_2\begin{pmatrix}-2\\-4\\-2\end{pmatrix} + x_3\begin{pmatrix}0\\2\\3\end{pmatrix} + x_4\begin{pmatrix}2\\0\\-3\end{pmatrix} + x_5\begin{pmatrix}-3\\8\\16\end{pmatrix}$$

$$= \begin{pmatrix}1 & -2 & 0 & 2 & -3\\2 & -4 & 2 & 0 & 8\\1 & -2 & 3 & -3 & 16\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = A\mathbf{x}.$$

Clearly, $A\mathbf{x} = \mathbf{b}$ has solutions if and only if $\mathbf{b}$ is a linear combination of the columns of $A$. A similar calculation provides a solution to $A\mathbf{x} = \mathbf{b}$ for any finite set of vectors $\mathbf{u}_1, \dots, \mathbf{u}_m$ in coordinate space $\mathbb{K}^n$. Viewing the $\mathbf{u}_i$ as $n \times 1$ column vectors we may assemble them into an $n \times m$ matrix $A = [\mathbf{u}_1; \dots; \mathbf{u}_n]$ and apply the following general result, whose proof is a simple variant of the preceding discussion.

LEMMA 1.38. *If $A$ is any $n \times m$ matrix and $\mathbf{b} \in \mathbb{K}^n$, the system $A\mathbf{x} = \mathbf{b}$ has solutions if and only if $\mathbf{b}$ is in the linear span of the columns $\mathbf{u}_1, \dots, \mathbf{u}_m$ of $A$, regarded as vectors in $\mathbb{K}^n$.*

In particular, a system $A\mathbf{x} = \mathbf{b}$ is consistent precisely when $\mathbf{b}$ belongs to the column space $\mathrm{ColSpace}(A)$, so it is in the range of $L_A : \mathbb{K}^m \to \mathbb{K}^n$.

EXERCISE 1.39. Fill in the details needed to verify the claim in Lemma 1.38.

Continuing the discussion of Example 1.37, we now allow arbitrary $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{K}^3$ and will determine whether the system $A\mathbf{x} = \mathbf{b}$ has solutions $\mathbf{x} = \mathrm{col}(x_1, \dots, x_5)$, and when it does we will find them all. By allowing undetermined coefficients in $\mathbf{b}$ we will also be determining the linear span of $\mathbf{u}_1, \dots, \mathbf{u}_5$, which is just the set of $\mathbf{b}$ for which the system $A\mathbf{x} = \mathbf{b}$ is consistent.

**The Homogeneous Equation $A\mathbf{x} = \mathbf{0}$.** Start by putting the system $A\mathbf{x} = \mathbf{0}$ into echelon form by applying row operations to the augmented matrix

$$[A : \mathbf{0}] = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & 2 & -3 & 0 \\ 2 & -4 & 2 & 0 & 8 & 0 \\ 1 & -2 & 3 & -3 & 16 & 0 \end{array}\right).$$

By applying $R_2 \to R_2 - 2R_1$ and $R_3 \to R_3 - R_1$, this becomes

$$\left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 2 & -3 & 0 \\ 0 & 0 & 2 & -4 & 14 & 0 \\ 0 & 0 & 3 & -5 & 19 & 0 \end{array}\right).$$

Now apply $R_3 \to R_3 - \frac{3}{2}R_2$, $R_2 \to \frac{1}{2}R_2$, and then $R_3 \to R_3 - 3R_2$ to get

$$[A' : \mathbf{0}] = \left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 2 & -3 & 0 \\ 0 & 0 & \boxed{1} & -2 & 7 & 0 \\ 0 & 0 & 0 & \boxed{1} & -2 & 0 \end{array}\right).$$

This is the desired echelon form. The system is consistent since there are no nonzero rows below the last step corner. Some additional work, needless for most purposes, would yield the *reduced echelon form*,

$$A'' = \left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 0 & 1 & 0 \\ 0 & 0 & \boxed{1} & 0 & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} & -2 & 0 \end{array}\right).$$

By recursively backsolving the corresponding system of linear equations $A'\mathbf{x} = \mathbf{0}$ we find that

1. $x_2, x_5$ are free variables;
2. $x_4 - 2x_5 = 0 \Rightarrow x_4 = 2x_5$;
3. $x_3 - 2x_4 + 7x_5 = 0 \Rightarrow x_3 = -7x_5 + 2(2x_5) = -3x_5$;
4. $x_1 - 2x_2 + 2x_4 - 3x_5 = 0 \Rightarrow x_1 = 2x_2 - 2(2x_5) + 3x_5 = 2x_2 - x_5$.

The solutions of $A'\mathbf{x} = \mathbf{0}$ (which are also the solutions of $A\mathbf{x} = \mathbf{0}$) form a vector subspace in $\mathbb{R}^5$, each of whose points is uniquely labeled (parametrized) by the choice of the free variables $x_2, x_5$. Relabeling $x_2 = s$, $x_5 = t$ ($s, t \in \mathbb{K}$) we find that the solution set $W = \{\mathbf{x} \in \mathbb{K}^5 : A\mathbf{x} = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{K}^5 : A'\mathbf{x} = \mathbf{0}\}$ is equal to

$$W = \left\{ \begin{pmatrix} 2x_2 - x_5 \\ x_2 \\ -3x_5 \\ 2x_5 \\ x_5 \end{pmatrix} : x_2, x_5 \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} 2s - t \\ s \\ -3t \\ 2t \\ t \end{pmatrix} : s, t \in \mathbb{K} \right\}.$$

These solutions can be rewritten in a more instructive form

$$\mathbf{x} = \begin{pmatrix} 2s-t \\ s \\ -3t \\ 2t \\ t \end{pmatrix} = s\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ -3 \\ 2 \\ 1 \end{pmatrix} = s\mathbf{w}_1 + t\mathbf{w}_2 \quad (s, t \in \mathbb{K}),$$

which shows that every solution of $A\mathbf{x} = \mathbf{0}$ is a linear combination of two basic solutions

$$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{pmatrix} -1 \\ 0 \\ -3 \\ 2 \\ 1 \end{pmatrix}. \tag{1.2}$$

Thus the solution set of $A\mathbf{x} = \mathbf{0}$ is the linear span $W = \mathbb{K}\text{-span}\{\mathbf{w}_1, \mathbf{w}_2\}$ of a set of generators $\{\mathbf{w}_1, \mathbf{w}_2\}$. We will later observe that these vectors are a "basis" for the solution set $W$.

**Solving Inhomogeneous Equations $A\mathbf{x} = \mathbf{b}$.** The same elementary row operations that put $A$ into echelon form may be applied to the augmented matrix $[A \,:\, \mathbf{b}]$ where $\mathbf{b} = \text{col}(b_1, b_2, b_3)$ has indeterminate coefficients. We already know what happens to $A$; applying the same moves to $\mathbf{b}$, the operations $R_2 \to R_2 - 2R_1$ and $R_3 \to R_3 - R_1$ transform

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \to \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - b_1 \end{pmatrix}.$$

Then $R_3 \to R_3 - \frac{3}{2}R_2$, $R_2 \to \frac{1}{2}R_2$, and $R_3 \to R_3 - 3R_2$ yield

$$\to \begin{pmatrix} b_1 \\ \frac{1}{2}b_2 - b_1 \\ b_3 - b_1 - \frac{3}{2}(b_2 - 2b_1) \end{pmatrix} = \begin{pmatrix} b_1 \\ \frac{1}{2}b_2 - b_1 \\ b_3 - \frac{3}{2}b_2 + 2b_1 \end{pmatrix}$$

so the augmented matrix becomes

$$[A \,:\, \mathbf{b}\,] \to [A' \,:\, \mathbf{b}'] = \left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 2 & -3 & b_1 \\ 0 & 0 & \boxed{1} & -2 & 7 & \frac{1}{2}b_2 - b_1 \\ 0 & 0 & 0 & \boxed{1} & -2 & b_3 - \frac{3}{2}b_2 + 2b_1 \end{array}\right). \tag{1.3}$$

Again, $x_2, x_5$ are free variables. The general solution $\mathbf{x} = \text{col}(x_1, x_2, x_3, x_4, x_5)$ of $A\mathbf{x} = \mathbf{b}$ can be found by backsolving. We already know the general solutions of $A\mathbf{x} = \mathbf{0}$, so all we need is one particular solution $\mathbf{x_b}$, and the simplest way to find one is to set the free variables $x_2 = x_5 = 0$ and backsolve to get

$$x_2, x_5 = 0, \qquad x_4 = b_3 - \frac{3}{2}b_2 + 2b_1,$$

$$x_3 - 2x_4 = \frac{1}{2}b_2 - b_1 \Rightarrow x_3 = 2\left(b_3 - \frac{3}{2}b_2 + 2b_1\right) + \frac{1}{2}b_2 - b_1$$
$$= 2b_3 - \frac{5}{2}b_2 + 3b_1,$$

$$x_1 - 2\cdot 0 + 0 + 2x_4 + 0 = b_1 \Rightarrow$$
$$x_1 = b_1 - 2x_4 = -2\left(b_3 - \frac{3}{2}b_2 + 2b_1\right) + b_1 = -2b_3 + 3b_2 - 3b_1.$$

So

$$\mathbf{x_b} = \mathrm{col}\left(-2b_3 + 3b_2 - 3b_1, 0, 2b_3 - \frac{5}{2}b_2 + 3b_1, b_3 - \frac{3}{2}b_2 + 2b_1, 0\right)$$

is a particular solution and the full solution set is

$$W_{\mathbf{b}} = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} = \begin{pmatrix} -2b_3 + 3b_2 - 3b_1 \\ 0 \\ 2b_3 - \frac{5}{2}b_2 + 3b_1 \\ b_3 - \frac{3}{2}b_2 + 2b_1 \\ 0 \end{pmatrix} + \mathbb{K}\mathbf{w}_1 + \mathbb{K}\mathbf{w}_2$$
$$= \mathbf{x_b} + \mathbb{K}\mathbf{w}_1 + \mathbb{K}\mathbf{w}_2$$

where $\mathbf{w}_1$ and $\mathbf{w}_2$ are the basis vectors for the space $W = \{x : A\mathbf{x} = \mathbf{0}\}$ of homogeneous solutions determined previously. Writing $x_2 = \mathrm{s}$, $x_5 = t$ for the variable attached to $\mathbf{w}_1$, $\mathbf{w}_2$ we have found a *parametric description* of the solution set, in which each point in $W_{\mathbf{b}}$ is tagged by a unique pair $(s, t)$ in the parameter space $\mathbb{K}^2$.

If $\mathbf{b} = \mathrm{col}(2, 6, 8)$, a particular solution is $\mathbf{x_b} = \mathrm{col}(-4, 0, 7, 3, 0)$, and the full solution set for $A\mathbf{x} = \mathbf{b}$ is

$$W_{\mathbf{b}} = \begin{pmatrix} -4 + 2s - t \\ s \\ 7 - 3t \\ 3 + 2t \\ t \end{pmatrix} = \mathbf{x_b} + \mathbb{K}\mathbf{w}_1 + \mathbb{K}\mathbf{w}_2.$$

Inspection of the echelon form (1.3) reveals that there is no obstruction to backsolving to get valid solutions of $A'\mathbf{x} = \mathbf{b}'$ (and $A\mathbf{x} = \mathbf{b}$) no matter which $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{K}^3$ we consider. In fact, there will be a two-parameter family of solutions for any given $\mathbf{b}$ because two columns in $A'$ fail to meet a step corner.

That concludes the present discussion. We will have more to say about these issues in Section 1.4.

**Determining the Linear Span of a Set of Vectors.** The linear span $E = \mathbb{K}\text{-span}(S)$ of a set of $m$ column vectors $S = \mathbf{u}_1, \dots, \mathbf{u}_m$ in $\mathbb{K}^n$ consists of all finite linear combinations $\sum_{i=1}^m c_i\mathbf{u}_i$. In Lemma 1.38 we showed that $E$ is precisely the set of vectors $\mathbf{b} \in \mathbb{K}^n$ such that the linear system $A\mathbf{x} = \mathbf{b}$ has solutions, where $A$ is the $n \times m$ matrix $A = [\mathbf{u}_1; \dots; \mathbf{u}_m]$.

**Implicit and Parametric Descriptions of Vector Subspaces.** Although this characterization is useful in theoretical discussions, one would also like to have more explicit descriptions of the linear span $E$. There are two approaches.

1. IMPLICIT DESCRIPTION. Describe the vectors in $E$ as the solution set of some new homogeneous system of linear equations $C\mathbf{x} = \mathbf{0}$ where $\mathbf{x} \in \mathbb{K}^n$. This tells us how the linear span is carved out of coordinate space $\mathbb{K}^n$ in which the given vectors $\mathbf{u}_i$ live.
2. PARAMETRIC DESCRIPTION. Determine a *basis* for $E$, a set of vectors $\mathbf{b}_1, \dots, \mathbf{b}_r$ such that every $\mathbf{b} \in E$ is a unique linear combination of these vectors.

In Section 1.4 we will show how such descriptions can be obtained, but for the moment we simply show how basis vectors for $E$ can be found using row operations.

EXAMPLE 1.40. Given $n$ vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in $\mathbb{K}^m$ we can find a set of basis vectors for their linear span $W$ using row operations.

DISCUSSION: Let us regard the $\mathbf{v}_i$ as the *rows* of an $n \times m$ matrix

$$B = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix}_{m\times n} .$$

Now $E = \mathbb{K}\text{-span}\{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ coincides with the row space RowSpace$(B)$. Applying row operations to put $B$ into echelon form

$$B' = \left(\begin{array}{cccccccccc} \boxed{1} & * & \cdot & \cdot & \cdot & & & & & * \\ & & & \boxed{1} & * & \cdot & \cdot & \cdot & \cdot & * \\ 0 & & & & & \boxed{1} & * & \cdot & \cdot & * \\ \hline & & & & & & & & & \\ & & & & \mathbf{0} & & & & & \end{array}\right),$$

we know that the span of the rows is the same for $B$ and $B'$, and in particular the *nonzero* rows $R_1(B'), \dots, R_r(B')$ of $B'$ span $E$. They also have the desired independence property. In fact, if some vector $\mathbf{v} \in E$ had two different expansions

$$\mathbf{v} = \sum_{i=1}^{r} a_i R_i(B') = \sum_{i=1}^{r} b_i R_i(B'),$$

their difference, the zero vector $\mathbf{0} = \mathbf{v} - \mathbf{v}$, would have a nontrivial expansion

$$\mathbf{0} = \sum_{i=1}^{r} (a_i - b_i) R_i(B') \quad \text{(not all coefficients} = 0).$$

The echelon form of $B'$ prevents this from happening. In fact, if any coefficient in this sum is nonzero, let $R_i(B')$ be the first row vector $(1 \leq i \leq r)$ for which this happens. The previous rows $(j < i)$ cannot undo this because they all get multiplied by zero, and the later rows $(j > i)$ can't either because they all have zeros in the column that contains the step corner of $R_i(B')$. Therefore we have $a_i = b_i$ in the preceding sum and the coefficients are unique.

**More on Elementary Row and Column Operations.** Elementary row operations on a matrix $A$ do not change the linear span RowSpace($A$) of its rows. Note carefully, however, that row operations will mess up column space!

As for columns, there is an obvious family of *elementary column operations* on any matrix $A$.

1. $C_i \leftrightarrow C_j$.
2. $C_i \to \lambda C_i$ for $\lambda \neq 0$ in $\mathbb{K}$.
3. $C_i \to C_i + \lambda C_j$ for $i \neq j$ where $\lambda$ is any element in $\mathbb{K}$.

Column operations on $A$ have properties completely analogous to those listed for row operations in Proposition 1.36. In particular, column operations preserve the linear span of the columns of $A$. We won't spell all this out since we will work mostly with row operations. The properties of column operations can of course be verified by direct calculation, but they also follow by observing that row and column operations are related via the transpose operation $A \to A^{\mathsf{T}}$on matrices given by

$$(A^{\mathsf{T}})_{ij} = A_{ji}$$

(see Figure 1.5). Note that $(A^{\mathsf{T}})^{\mathsf{T}} = A$.

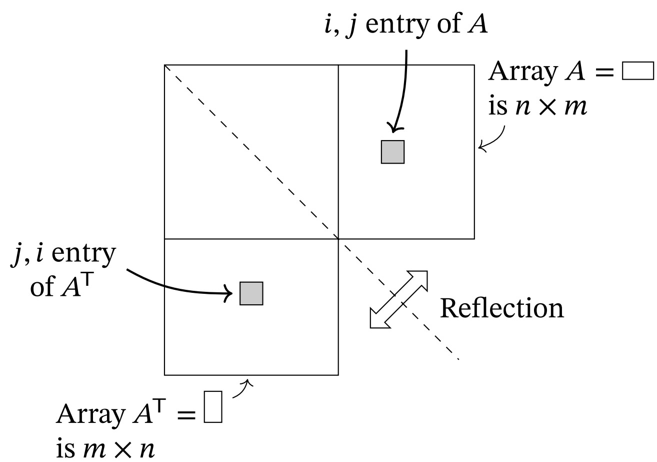


FIGURE 1.5. A matrix $A$ and its transpose $A^{\mathsf{T}}$ are related by a reflection that sends rows in $A$ to columns in $A^{\mathsf{T}}$, and columns to rows.

Obviously transposition takes rows of $A$ to columns of $A^{\mathsf{T}}$ and vice versa. Furthermore, elementary row operations on $A$ become the corresponding elementary operations on columns of $A^{\mathsf{T}}$, so that ColSpace($A$) = RowSpace($A^{\mathsf{T}}$). Invariance of ColSpace($A$) under column operations follows from invariance of RowSpace($A^{\mathsf{T}}$) under row operations.

EXERCISE 1.41. Verify that the transpose operation $A \to A^{\mathsf{T}}$ on matrices of any shape has the following algebraic properties:

(a) $(A^{\mathsf{T}})^{\mathsf{T}} = A$.
(b) If $A$ is square and $I$ is the $n \times n$ identity matrix, then $I^{\mathsf{T}} = I$.
(c) $(c \cdot A)^{\mathsf{T}} = c \cdot A^{\mathsf{T}}$ for any scalar $c \in \mathbb{K}$.
(d) $(A + B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}}$.

If $A$ is $m \times k$ and $B$ is $k \times n$ they have compatible shapes and their matrix product $AB$ is a well-defined $m \times n$ matrix.

(e) Prove that $(AB)^{\mathsf{T}} = B^{\mathsf{T}} \cdot A^{\mathsf{T}}$ for compatible matrices.

Note the reversal of order in (e).

EXERCISE 1.42. If $A$ is an $n \times m$ matrix, consider an elementary row operation $T_R$, and let $T_C$ be the corresponding column operation on $m \times n$ matrices. (For instance, $T_R : R_i \leftrightarrow R_j$ in $A$ corresponds to $T_C : C_i \leftrightarrow C_j$ in $A^{\mathsf{T}}$.) Verify that

$$\left(T_R(A)\right)^{\mathsf{T}} = T_C(A^{\mathsf{T}}) \quad \text{or equivalently} \quad \left(T_C(A^{\mathsf{T}})\right)^{\mathsf{T}} = T_R(A)$$

for all $n \times m$ matrices $A$.

EXERCISE 1.43. Use Exercise 1.42 to verify the following:

LEMMA 1.44. *If a row operation $T_R$ on $\mathrm{M}(n, \mathbb{K})$ corresponds to left multiplication $T_R : A \to EA$ by the elementary matrix $E$, the corresponding column operation is $T_C : A \to AE^{\mathsf{T}}$.*

## 1.4. Linear Span, Independence, and Bases

DEFINITION 1.45. A set of vectors $S = \{v_1, \dots, v_r\}$ in a vector space $V$ *spans* a subspace $W$ if

$$W = \mathbb{K}\text{-span}\{S\} = \left\{\sum_{i=1}^{r} c_i v_i \,:\, c_i \in \mathbb{K}\right\}.$$

The vectors are *linearly independent* if the only linear combination $\sum_i c_i v_i = \mathbf{0}$ adding up to zero in $V$ is the trivial combination with $c_1 = \cdots = c_r = 0$. The vectors are a *basis* for $W$ if they span $W$ and are independent, so every $w \in W$ has a unique representation as $\sum_{i=1}^{n} \lambda_i v_i$ ($\lambda_i \in \mathbb{K}$).

EXERCISE 1.46. If $\mathfrak{X} = \{v_1, \dots, v_n\}$ span $V$ and are independent, explain why every $v \in V$ has a unique representation as $\sum_{i=1}^{n} \lambda_i v_i$ ($\lambda_i \in K$), so $\mathfrak{X}$ is a basis for $V$.

**Existence and Construction of Bases.** The next result exhibits two ways to construct a basis in a vector space. One starts with a spanning set and then "prunes" it, deleting redundant vectors until we arrive at an independent subset with the same span as the original vectors. This yields a basis for $V$. The other constructs a basis recursively by adjoining "outside vectors" to an initial family of independent vectors in $V$. The initial family might consist of a single nonzero vector (obviously an independent set).

PROPOSITION 1.47. *Every finite set* $\{v_1, \dots, v_n\}$ *that spans a vector space* $V$ *can be made into a basis by deleting suitably chosen entries from the list.*

PROOF. We argue by induction on $n = \#(\text{vectors in list})$. There is nothing to prove if $n = 1$; then $V = \mathbb{K} \cdot v_1$ and $\{v_1\}$ is already a basis. The induction hypotheses (one for each index $n = 1, 2, \dots$) are:

> HYPOTHESIS $P(n)$: *If a set of n vectors spans a vector space* $V$*,*
> *we can delete vectors from the list to get a basis for* $V$*.*

We have proved this for $n = 1$. It will be true for all $n \in \mathbb{N}$ if we can prove $P(n+1)$ is true, using only the information that $P(n)$ is true—i.e., if we can verify that

$$P(n) \text{ true} \Rightarrow P(n+1) \text{ true} \qquad \text{for all } n = 1, 2, \dots .$$

(Remember: This is a *conditional* statement owing to the presence of the word "*if*". By itself, it does not assert that either statement $P(n)$ or $P(n+1)$ is actually true.)

So, assuming $P(n)$ true, consider a spanning set $\mathfrak{X} = \{v_1, \dots, v_n, v_{n+1}\}$ in $V$. If these vectors are already independent (which could be checked using row operations if $V = \mathbb{K}^m$), we already have a basis for $V$ without deleting any vectors. If $\mathfrak{X}$ is not independent there must be coefficients $c_1, \dots, c_{n+1} \in \mathbb{K}$ (not all equal to 0) such that $\sum_{i=1}^{n+1} c_i v_i = \mathbf{0}$. Relabeling, we may assume $c_{n+1} \neq 0$, and then ($\mathbb{K}$ being a field)

$$-c_{n+1} v_{n+1} = \sum_{i=1}^{n} c_i v_i \quad \text{and} \quad v_{n+1} = \sum_{i=1}^{n} -(c_i / c_{n+1}) \cdot v_i .$$

Thus $v_{n+1} \in \mathbb{K}\text{-span}\{v_1, \dots, v_n\}$ and $\mathbb{K}\text{-span}\{v_1, \dots, v_{n+1}\} = \mathbb{K}\text{-span}\{v_1, \dots, v_n\}$ is all of $V$. By the induction hypotheses we may thin out $\{v_1, \dots, v_{n+1}\}$ to get a basis for $V$. □

PROPOSITION 1.48. *If* $\{v_1, \dots, v_n\}$ *are independent in a vector space* $V$*, and* $v_{n+1}$ *is a vector not in* $W_0 = \mathbb{K}\text{-span}\{v_1, \dots, v_n\}$*, then*:

1. $\{v_1, \dots, v_n, v_{n+1}\}$ *are independent.*
2. $W_0 \subsetneqq W_1 = \mathbb{K}\text{-span}\{v_1, \dots, v_n, v_{n+1}\}$*.*
3. $\{v_1, \dots, v_n, v_{n+1}\}$ *is a basis for* $W_1$*.*

PROOF. If $v_1 \dots, v_{n+1}$ are not independent there would be $c_i \in \mathbb{K}$ (not all zero) such that $\sum_{i=1}^{n+1} c_i v_i = \mathbf{0}$. We can't have $c_{n+1} = 0$; otherwise $\sum_{i=1}^{n} c_i v_i = \mathbf{0}$ contrary to the assumed independence of $\{v_1, \dots, v_n\}$. Thus

$$v_{n+1} = \sum_{i=1}^{n} -(c_i / c_{n+1}) \cdot v_i$$

is in $W_0$, which contradicts the assumption $v_{n+1} \notin W_0$. We then conclude that $v_1, \dots, v_{n+1}$ are independent. It follows immediately that $\{v_1, \dots, v_{n+1}\}$ is a basis for $W_1 = \mathbb{K}\text{-span}\{v_1, \dots, v_{n+1}\}$.

NOTE: This is an example of a "proof by contradiction," in which the assumption that $v_1, \dots, v_n$ are *not* independent leads to an impossible conclusion. Thus the statement "$v_1, \dots, v_n$ are independent" must be true. □

IMPORTANT REMARK: This process of "adjoining an outside vector" can be iterated to construct larger and larger independent sets and subspaces

$$W_0 = \mathbb{K}\text{-span}\{v_1, \dots, v_n\} \subsetneqq W_1 = \mathbb{K}\text{-span}\{v_1, \dots, v_n, v_{n+1}\} \subsetneqq \cdots \subsetneqq W_r = \mathbb{K}\text{-span}\{v_1, \dots, v_{n+r}\}.$$

Since $\{v_1, \dots, v_n\}$ are independent, they are a basis for the initial space $W_0$, and by Proposition 1.48, $v_1, \dots, v_n, \dots, v_{n+r}$ will be a basis for $W_r$. If this process stops in finitely many steps (because $W_r = V$ and we can no longer find a vector outside $W_r$), we have produced a basis for $V$. If the process never stops, no finite subset of vectors can span $V$ and in this case we say $V$ is *infinite dimensional.* To begin the process we need an initial set of independent vectors, but if $V \neq (0)$ we could start with any $v_1 \neq 0$ and $W_0 = \mathbb{K} \cdot v_1$. Then apply Proposition 1.48 recursively as above.

DEFINITION 1.49. A vector space $V$ is *finite dimensional* if there is a finite set of vectors $S = \{v_1, \dots, v_n\}$ that span $V$. Otherwise $V$ is *infinite dimensional* which we indicate by writing $\dim(V) = \infty$.

Coordinate space $\mathbb{K}^n$ and matrix spaces $\mathrm{M}(m \times n, \mathbb{K})$ are finite dimensional; the spaces of polynomials $\mathbb{K}[x]$ and $\mathbb{K}[x_1, \dots, x_n]$ are infinite dimensional.

EXAMPLE 1.50. Coordinate space $\mathbb{K}^n$ is finite dimensional and is spanned by the *standard basis vectors* $\mathfrak{X} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$

$$\mathbf{e}_1 = (1, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \quad \dots \quad , \mathbf{e}_n = (0, \dots, 0, 1).$$

In fact, $\mathfrak{X}$ is a basis for $\mathbb{K}^n$.

EXAMPLE 1.51. Polynomial space $\mathbb{K}[x]$ is infinite dimensional. Given any finite set of nonzero vectors $\mathfrak{X} = \{f_1, \dots, f_r\}$, let $d_i = \deg(f_i)$. All coefficients of terms $c_j x^j$ in $f_i$ are zero if $j > N = \max\{d_1, \dots, d_r\}$, and the same is true for all linear combinations $\sum_{i=1}^{r} c_i f_i$. But then $\mathfrak{X}$ cannot span $\mathbb{K}[x]$ because $x^{N+1}$ is not in $\mathbb{K}\text{-span}\{f_1, \dots, f_r\}$.

However, the vectors $f_0 = 1, f_1 = x, f_2 = x^2, \dots$ *are* a basis for $\mathbb{K}[x]$. This (infinite) set of vectors clearly spans $\mathbb{K}[x]$, but it is also independent, for if $\sum_{i=0}^{r} c_i f_i = 0$ that means $c_0 + c_1 x + \cdots + c_r x^r = 0$ as a polynomial, so the symbol string $(c_0, \dots, c_r, 0, 0, \dots)$ is equal to $(0, 0, 0, \dots)$.

LEMMA 1.52. *Every finite-dimensional vector space has a basis.*

PROOF. If $\{v_1, \dots, v_r\}$ span $V$, then by Proposition 1.47 we may delete some of the vectors to get an independent set with the same linear span. □

PROPOSITION 1.53. *If $S \subseteq V$ is an independent set of vectors in a finite-dimensional vector space $V$, and $T$ a finite set of vectors that span $V$, we can adjoin certain vectors from $T$ to $S$ to get a basis for $V$ that contains the original set of independent vectors $S$.*

PROOF. Let $W = \mathbb{K}\text{-span}\{S\}$. If $W = V$, $S$ is already a basis. If $W \neq V$, there exists some $v_1 \in T$ such that $v_1 \notin W$ and then $S \cup \{v_1\}$ is an independent set, a basis for the larger space $W_1 = \mathbb{K}\text{-span}\{S \cup \{v_1\}\} \supsetneqq W$. Continuing, we get vectors $v_1, \dots, v_r$ in $T$ such that $W \subsetneqq W_1 \subsetneqq W_2 \subsetneqq \cdots \subsetneqq W_r$ for $0 \leq i \leq r$, where $W_i = \mathbb{K}\text{-span}\{v_1, \dots, v_i\}$. The process must terminate when no vector $v_{r+1} \in T$ can be found outside of $W_r$. Then $T \subseteq W_r$, so $\mathbb{K}\text{-span}\{T\} = V \subseteq W_r$ and $W_r = V$. Therefore $S \cup \{v_1, \dots, v_r\}$ is a basis for $V = W_r$ (and $S \cup \{v_1, \dots, v_k\}$ is a basis for $W_k$ for each $1 \leq k \leq r$). □

### The Dimension dim($V$) of a Vector Space.

THEOREM 1.54. *All bases in a finite-dimensional vector space have the same cardinality. More generally, if $V$ is finite dimensional, and $S$ is a finite spanning set (with $|S| = n$), every independent set of vectors $L \subseteq V$ has cardinality $|L| \leq |S|$. In other words, the size of any independent set is always less than or equal to that of any spanning set.*

PROOF. We can eliminate vectors from $S$ to get an independent spanning set $S' \subseteq S$, which is then a basis for $V$, and will show that $|L| \leq |S'| \leq |S|$. Let $S' = \{u_1, \dots, u_n\}$ and $L = \{v_1, \dots, v_m\}$. Every $v_i \in L$ can be written $v_i = \sum_{i=1}^{n} a_{ji} u_j$ since the $u_i \in S'$ are a basis for $V$. On the other hand, if $c_1, \dots, c_m$ are scalars such that $0 = \sum_{j=1}^{m} c_j v_j$, we must have $c_1 = \cdots = c_m = 0$ because the $v_j$ are independent. But the identity $\sum_{j=1}^{m} c_j v_j = 0$ can be written another way, as

$$0 = \sum_{i=1}^{m} c_i \Big( \sum_{j=1}^{n} a_{ji} u_j \Big) = \sum_{j=1}^{n} \Big( \sum_{i=1}^{m} a_{ji} c_i \Big) u_j .$$

Since the $u_j \in S'$ span $V$ and are independent, each sum $(\cdots)$ is equal to 0 so the coefficients $c_1, \dots, c_m$ satisfy the system of $n$ equations in $m$ unknowns

$$\sum_{i=1}^{m} a_{ji} c_i = 0 \quad \text{for } 1 \leq j \leq n. \tag{1.4}$$

Hence the $m \times 1$ matrix $\mathbf{c} = \text{col}(c_1, \dots, c_m)$ is a solution of the matrix equation $A\mathbf{c} = \mathbf{0}$.

A linear system such as (1.4) always has nontrivial solutions if the number of unknowns $m = |L|$ exceeds the number of equations $n = |S'|$, and from this it will follow that $|L| \leq |S'|$, as claimed. In fact, row operations on the coefficient matrix $A$ yield an echelon form $A'$ shown below. There are just $n$ columns in $A$

and $A'$, hence at most $n$ step corners in $A'$, and if $m > n$ there must be at least one column that fails to meet a step corner:

Echelon form when $m > n$:

$$A' = \left(\begin{array}{ccccccccc} \boxed{1} & * & . & . & . & & & & * \\ & & & . & . & . & . & . & . \\ & & & . & . & . & . & . & . \\ & & & \boxed{1} & * & . & . & . & * \\ 0 & & & & & \boxed{1} & * & . & * \\ \hline 0 & . & & . & & . & & . & 0 \end{array}\right)_{n\times m}$$

Hence there is at least one free variable and the system $A\mathbf{c} = \mathbf{0}$ has nontrivial solutions. But we showed above that $\mathbf{c} = \mathbf{0}$ is the only solution, so we obtain a contradiction unless $|L| \leq |S'| \leq |S|$. The theorem is proved. □

COROLLARY 1.55. *In a finite-dimensional vector space, all bases have the same cardinality, which we refer to hereafter as the dimension* $\dim_{\mathbb{K}}(V)$.

NOTATION: We will often simplify notation when the underlying ground field $\mathbb{K}$ is understood by writing $\dim(V)$ or even $|V|$ for the dimension of $V$.

EXAMPLE 1.56. We have already seen that $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$, with the standard basis vectors $\mathbf{e}_1 = (1, 0, \dots, 0), \dots, \mathbf{e}_n = (0, \dots, 0, 1)$. We may view $\mathbb{C}^n$ (or any vector space over $\mathbb{C}$) as a vector space over $\mathbb{R}$ by restricting scalars in $\lambda \cdot v$ to be real. As a vector space over $\mathbb{C}$ we have $\dim_{\mathbb{C}}(V) = n$, but as a vector space over $\mathbb{R}$ we have $\dim_{\mathbb{R}}(V) = 2n$.

DISCUSSION: In fact, any $v \in \mathbb{C}^n$ can be written as a sum with complex coefficients $v = \sum_{j=1}^{n} z_j\mathbf{e}_j$, and if $z_j = x_j + iy_j$ this becomes

$$v = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n + y_1(i\mathbf{e}_1) + \cdots + y_n(i\mathbf{e}_n) \quad \text{with } x_i, y_j \in \mathbb{R}.$$

Thus the vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_n, i\mathbf{e}_1, \dots, i\mathbf{e}_n\} \subseteq \mathbb{C}^n$ span $\mathbb{C}^n$ as a vector space over $\mathbb{R}$. They are also independent over $\mathbb{R}$, for if $a_i, b_j \in \mathbb{R}$ and

$$\mathbf{0} = \sum a_j\mathbf{e}_j + \sum b_j(i\mathbf{e}_j) = \sum(a_j + ib_j)\mathbf{e}_j\ ,$$

we must have $a_j + ib_j = 0$ and $a_j = b_j = 0$ because $\{\mathbf{e}_j\}$ is a basis over $\mathbb{C}$. Since elements of $\mathbb{C}$ are $\mathbb{R}$-linear combinations $z = a + ib = a1 + ib$, the elements $1, i$ are a basis for $\mathbb{C}$ when we regard it as a vector space over $\mathbb{K} = \mathbb{R}$. Thus $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ while $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.

EXERCISE 1.57. If $V$ is a finite-dimensional vector space and $W \subseteq V$ a subspace, explain why $W$ must also be finite dimensional.

NOTE: When $\mathbb{C}$ is regarded as a vector space over $\mathbb{K} = \mathbb{R}$ then $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ and the elements $v_1 = 1, v_2 = i$ are convenient basis vectors over $\mathbb{R}$ because $z = a + ib = a \cdot 1 + b \cdot i$.

EXERCISE 1.58. If $V_1, V_2$ are finite-dimensional vector spaces, prove the following:

(a) If $V_1 \subseteq V_2$ then $\dim(V_1) \leq \dim(V_2)$.
(b) If $\dim(V_1) = \dim(V_2)$ and $V_1 \subseteq V_2$, then $V_1 = V_2$ as sets.

EXERCISE 1.59. Explain why $W \subseteq V \Rightarrow \dim(W) \leq \dim(V)$, even if one or both of these spaces is infinite dimensional.

**Implicit and Parametric Description of Subspaces (Revisited).** How can a subspace $W$ in a vector space be specified? Every $V$ of dimension $n$ can be identified in a natural way with $\mathbb{K}^n$ once a basis $\{f_1, \dots, f_n\}$ in $V$ has been determined, so we may as well restrict attention to describing subspaces $W$ of coordinate space $\mathbb{K}^n$. Given a basis $\mathfrak{X} = \{f_i\}$ in $V$, the map $j_{\mathfrak{X}} : \mathbb{K}^n \to V$ given by

$$\mathbf{x} = (x_1, \dots, x_n) \mapsto j_{\mathfrak{X}}(\mathbf{x}) = \sum_{i=1}^{n} x_i f_i$$

is a bijection that respects all vector space operations in the sense that

$$j_{\mathfrak{X}}(\lambda \cdot \mathbf{x}) = \lambda \cdot j_{\mathfrak{X}}(\mathbf{x}) \quad \text{and} \quad j_{\mathfrak{X}}(\mathbf{x} + \mathbf{y}) = j_{\mathfrak{X}}(\mathbf{x}) + j_{\mathfrak{X}}(\mathbf{y}).$$

It is an *isomorphism* between $\mathbb{K}^n$ and $V$, by which properties of one space can be matched with those of the other.

DEFINITION 1.60. Subspaces $W \subseteq \mathbb{K}^n$ can be described in two ways.

1. By exhibiting a basis $\mathfrak{X} = \{f_1, \dots, f_r\}$ in $W$, so $W = \mathbb{K}\text{-span}\{\mathfrak{X}\}$ and $\dim_{\mathbb{K}}(W) = r$. This is a "parametric description" of $W$ since each $w$ in $W$ is labeled by a unique coordinate $r$-tuple $\mathbf{c} = (c_1, \dots, c_r)$.
2. By finding a set of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1m}x_m &= 0 \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m &= 0 \end{aligned}$$

   described by a matrix equation $A\mathbf{x} = \mathbf{0}$ (where $A = n \times m$, $\mathbf{0} = n \times 1$, $\mathbf{x} = m \times 1$) with $A_{ij} = a_{ij}$, whose solution set $\{\mathbf{x} \in \mathbb{K}^m : A\mathbf{x} = \mathbf{0}\}$ is equal to $W$. Such an "implicit description" may include redundant equations. When there are no redundant equations we will see that $W = \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}$ has dimension

$$\dim(W) = m - n = \dim(V) - \#(\textit{constraint equations})$$

We illustrate this with some computational examples.

EXAMPLE 1.61. Find a basis for the subspace $W = \mathbb{R}\text{-span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in $\mathbb{R}^3$ if

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix},$$

and determine $\dim_{\mathbb{R}}(W)$. Then describe $W$ as the solution set of a system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0 \\ &\vdots \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= 0 \end{aligned}$$

where $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.

SOLUTION: Write the vectors as the *rows* of the $3 \times 3$ matrix

$$A = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}.$$

The span RowSpace($A$) of the rows is unaffected by elementary row operations, which yield the echelon form

$$A \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & -2 & -4 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore $\mathbf{w}_1 = (1, 2, 3)$ and $\mathbf{w}_2 = (0, 1, 2)$ span $W$; they are also independent because $0 = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 = (c_1, 2c_1 + c_2, 3c_1 + 2c_2)$ implies

$$\begin{cases} c_1 = 0 \\ 2c_1 + c_2 = 0 \\ 3c_1 + 2c_2 = 0 \end{cases} \Rightarrow c_1 = c_2 = 0.$$

Thus $\mathfrak{X} = \{\mathbf{w}_1, \mathbf{w}_2\}$ is a basis and $\dim(W) = 2$. A typical vector in $W$ can be written (uniquely) as

$$s\mathbf{w}_1 + t\mathbf{w}_2 = (s, 2s + t, 3s + 2t) = (x_1, x_2, x_3) \quad \text{with } s, t \in \mathbb{R}.$$

To describe $W$ as the solution set of a system of equations in $x_1$, $x_2$, $x_3$ we need to "eliminate" $s$ and $t$ from this parametric description of $W$. This can be done by writing $x_1, x_2, x_3$ in terms of $s, t$

$$\begin{cases} x_1 = s & \Rightarrow s = x_1, \\ x_2 = 2s + t & \Rightarrow x_2 = 2s + t = 2x_1 + t \Rightarrow t = x_2 - 2x_1, \\ x_3 = 3s + 2t. & \end{cases}$$

The last equation yields the "constraint" identity that determines $W$:

$$x_3 = 3s + 2t = 3x_1 + 2(x_2 - 2x_1) = -x_1 + 2x_2$$

or $x_1 - 2x_2 + x_3 = 0$ (1 equation in 3 unknowns). Thus $W = \{\mathbf{x} \in \mathbb{R}^3 : x_1 - 2x_2 + x_3 = 0\}$, which has dimension $\dim(\mathbb{R}^3) - 1 = 2$.

EXAMPLE 1.62. Let $W \subseteq \mathbb{R}^4$ be the solution set of the system of linear equations

$$\begin{cases} x_1 + x_2 - x_3 + 2x_4 = 0, \\ 3x_1 - x_2 + x_4 = 0, \end{cases}$$

so $A\mathbf{x} = \mathbf{0}$ ($\mathbf{x} \in \mathbb{R}^4$) where

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

Find a basis for $W$ and determine $\dim_{\mathbb{R}}(W)$. Do the answers change if we replace $\mathbb{R}$ by $\mathbb{Q}$ or $\mathbb{C}$?

SOLUTION: Elementary row operations yield

$$A \to \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -4 & 3 & -6 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & -1 & 2 \\ 0 & \boxed{1} & -\frac{3}{4} & -\frac{3}{2} \end{pmatrix},$$

and for any solution of $A\mathbf{x} = 0$, $\mathbf{x} = \mathrm{col}(x_1, x_2, x_3, x_4)$ has $x_3$, $x_4$ as free variables. Backsolving yields the dependent variables

$$x_2 = \frac{3}{4}x_3 - \frac{3}{2}x_4,$$
$$x_1 = -x_2 + x_3 - 2x_4 = \left(-\frac{3}{4}x_3 + \frac{3}{2}x_4\right) + x_3 - 2x_4 = \frac{1}{4}x_3 - \frac{1}{2}x_4.$$

Thus solutions have the form

$$\mathbf{x} = \begin{pmatrix} \frac{1}{4}x_3 - \frac{1}{2}x_4 \\ \frac{3}{4}x_3 - \frac{3}{2}x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} \frac{1}{4} \\ \frac{3}{4} \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{1}{2} \\ -\frac{3}{2} \\ 0 \\ 1 \end{pmatrix} = x_3\mathbf{f}_1 + x_4\mathbf{f}_2$$

for every $x_3, x_4 \in \mathbb{K}$. After rescaling basis vectors to clear denominators, we see that the solution set is

$$\mathbb{R}\text{-span}\{\mathrm{col}(1,3,4,0), \mathrm{col}(3,1,0,-2)\} = \mathbb{R}\text{-span}\{\mathbf{f}_1, \mathbf{f}_2\}.$$

The vectors $\mathbf{f}_1, \mathbf{f}_2$ span the solution set $W$ but are also independent because

$$c_1(1,3,4,0) + c_2(3,1,0,-2) = (c_1 + 3c_2, 3c_1 + c_2, 4c_1, -2c_2) = (0,0,0,0)$$

implies that $c_1 = c_2 = 0$. Thus $\{\mathbf{f}_1, \mathbf{f}_2\}$ is a basis and $\dim_{\mathbb{R}}(W) = 2$. The result is the same if we replace the ground field $\mathbb{R}$ with $\mathbb{Q}$ or $\mathbb{C}$. As a rule of thumb, each constraint equation $a_{i1}x_1 + \cdots + a_{im}x_m = 0$ on $\mathbb{K}^m$ reduces the dimension of the solution set $W = \{\mathbf{x} \in \mathbb{K}^m : A\mathbf{x} = 0\}$ by 1, but this is not always the case.

EXERCISE 1.63. Consider the special case of one constraint equation,

$$W = \left\{\mathbf{x} : \sum_{i=1}^{n} c_i x_i = 0\right\} \quad \text{with } c_1, \dots, c_n \in \mathbb{K}.$$

(a) Under what condition on $\{c_1, \dots, c_n\}$ do we have $\dim_{\mathbb{K}}(W) = n - 1$?
(b) Explain why $\dim(W) < n - 1$ is impossible.

EXERCISE 1.64. Given vectors in $\mathbb{R}^4$

$$\mathbf{v}_1 = (1,0,-1,2), \quad \mathbf{v}_2 = (2,1,1,-3), \quad \mathbf{v}_3 = (-2,1,0,-2),$$

determine whether they are linearly independent, and then decide whether the vector $\mathbf{w} = (1, 2, 3, -4)$ lies in their linear span.

EXERCISE 1.65. Find an explicit basis for the subspace in $\mathbb{R}^n$ determined by the single constraint equation

$$x_1 + \cdots + x_n = 0.$$

What is its dimension?

**The Lagrange Interpolation Formula.**

EXAMPLE 1.66 (Lagrange Interpolation Formula). For any infinite field, such as $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, the problem of finding a polynomial $f \in \mathbb{K}[x]$ having specified values $f(p_j) = \lambda_j$ at a finite set of distinct points $p_1, \dots, p_n$ in $\mathbb{K}$ always has a solution. The solution is nonunique (the problem is underdetermined) unless we require that $\deg(f) = n-1$; there may be no solution if $\deg(f) < n-1$.

DISCUSSION: The product $h(x) = \prod_{j=1}^{n}(x - p_j)$ has degree equal to $n$ and is zero at each $p_j$ (and zero nowhere else), so the solution to the interpolation problem cannot be unique without restrictions on $f(x)$: one can add $h$ (or any scalar multiple thereof) to any proposed solution $f$. It is reasonable to ask for a solution $f(x)$ of minimal degree to reduce the ambiguity. The polynomial

$$f(x) = \sum_{i=1}^{n} \lambda_i \cdot \frac{\prod_{j \neq i}(x - p_j)}{\prod_{j \neq i}(p_i - p_j)} \tag{1.5}$$

has nonzero denominator, is equal to $\lambda_i$ at $p_i$ for each $i$, and has $\deg(f) = n-1$.

This is the *Lagrange interpolation formula*, determined by direct methods. It is a bit complicated to rewrite this sum of products in the form $f = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$. But the coefficients $c_0, \dots, c_{n-1}$ can also be found directly as the solution of a system of linear equations

$$\lambda_j = f(p_j) = \sum_{k=0}^{n-1} p_j^k c_k \quad \text{for } 1 \leq j \leq n-1,$$

which is equivalent to the matrix equation $A\mathbf{c} = \lambda$ in which

$$A = \begin{pmatrix} p_1^0 & \cdots & p_1^{n-1} \\ \vdots & \ddots & \vdots \\ p_n^0 & \cdots & p_n^{n-1} \end{pmatrix}_{n\times n} \quad \text{and} \quad \mathbf{c} = \begin{pmatrix} c_0 \\ \vdots \\ c_{n-1} \end{pmatrix}_{n\times 1}, \quad \lambda = \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_{n-1} \end{pmatrix}_{n\times 1}.$$

EXERCISE 1.67. (Another Dimension Formula) If $E, F$ are subspaces in a finite-dimensional vector space $V$ and $E + F = \{e + f \,:\, e \in E,\ f \in F\}$ is their linear span, prove that

$$\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F).$$

HINT: Choose appropriate bases related to $E, F$ and $E \cap F$.

**Rank of a Matrix: Row Rank vs. Column Rank.** If we view vectors in $\mathbb{K}^m$ as $m \times 1$ column vectors, an $n \times m$ matrix $A$ determines a map $L_A : \mathbb{K}^m \to \mathbb{K}^n$,

$$\mathbf{y} = L_A(\mathbf{x}) = A \cdot \mathbf{x} \quad \text{(matrix product)}$$

As an immediate consequence of this definition, $L_A$ is a *linear operator* from $\mathbb{K}^m$ to $\mathbb{K}^n$, one that preserves linear combinations of vectors in the sense that

$$L_A\Big(\sum_{i=1}^{r} \lambda_i \mathbf{x}_i\Big) = \sum_{i=1}^{r} \lambda_i \cdot L_A(\mathbf{x}_i) \quad (\mathbf{x}_i \in \mathbb{R}^m).$$

It follows that the *range* of the map $L_A$,

$$R(L_A) = L_A(\mathbb{K}^m) = \{A \cdot \mathbf{x} : \mathbf{x} \in \mathbb{K}^m\},$$

is a vector subspace in $\mathbb{K}^n$.

DEFINITION 1.68. We have defined the *rank* of a matrix $A$ to be the dimension dim $R(L_A)$ of the range of this operator,

$$\text{RANK OF A MATRIX:} \quad \text{rank}(A) = \dim(R(L_A)).$$

The problem of finding a basis for $R(L_A)$, and $\dim R(L_A)$, arises quite often in linear algebra. If $\mathbf{e}_1, \ldots, \mathbf{e}_m$ is the standard basis in $\mathbb{K}^m$, viewed as $m \times 1$ column vectors, then by definition of matrix multiplication the image vector $L_A(\mathbf{e}_i) = A \cdot \mathbf{e}_i$ is precisely the $i^{\text{th}}$ column $C_i(A)$. The linearity property of $L_A$ then implies that $R(L_A)$ is the linear span $\mathbb{K}\text{-span}\{C_1, \ldots, C_m\}$ of the columns of $A$.

DEFINITION 1.69. (Column Space/Row Space) The linear span of the rows of an $n \times m$ matrix $A$ is the *row space* of $A$, denoted RowSpace($A$), which can be viewed as a subspace of $\mathbb{K}^m$. Similarly, the linear span of the columns is the *column space* of $A$, ColSpace(A), which can be viewed as a vector subspace in $\mathbb{K}^n$. It is precisely the range of the map $L_A : \mathbb{K}^m \to \mathbb{K}^n$.

We now cite two important results addressing these issues, but defer the proofs until Chapter 4 since they involve *determinants*, which will be discussed in that chapter.

THEOREM 1.70. *If row reduction of an $n \times m$ matrix $A$ yields an echelon form $E$, the columns $C_{j_1}(A), \ldots, C_{j_r}(A)$ of $A$ that correspond to the pivot columns $C_{j_1}(E), \ldots, C_{j_r}(E)$ of $E$ (those columns in $E$ that contain a step corner) are a basis for* ColSpace($A$) *and the rank of $A$ is the number $r$ of pivot columns in $E$.*

This is unexpected because row operations do not interact well with column operations; for instance, if $A'$ is obtained from $A$ by an elementary operation on its rows, the column spaces of $A$ and $A'$ need not be the same subspace in $\mathbb{K}^n$. Nevertheless, a basis for ColSpace($A$), which would normally be determined by performing elementary *column operations* on $A$, can be found by inspection of the echelon matrix $E$ determined by performing *row operations* on $A$.

There is another important consequence of Theorem 1.70. As noted above, the rank $r = \dim(L_A(\mathbb{K}^m))$ of an $n \times m$ matrix is equal to its column rank,

$$\begin{aligned}\text{ColRank(A)} &= \dim\big(\mathbb{K}\text{-span}\{C_1(A), \dots, C_m(A)\}\big)\\ &= \#(\text{linearly independent columns in } A).\end{aligned}$$

The row rank of $A$ is defined similarly, as the dimension of the subspace in $\mathbb{K}^m$ spanned by the rows of $A$,

$$\begin{aligned}\text{RowRank}(A) &= \dim(\mathbb{K}\text{-span}\{R_1(A), \dots, R_n(A)\})\\ &= \#(\text{linearly independent rows in } A).\end{aligned}$$

These turn out to be equal.

COROLLARY 1.71. RowRank$(A)$ = ColRank$(A)$, *whether or not $A$ is a square matrix.*

EXERCISE 1.72. Prove Corollary 1.71 is a consequence of Theorem 1.70.

## 1.5. Quotient Spaces $V/W$

If $V$ is a vector space and $W$ a subspace, the *additive cosets* of a subspace $W$ are translates of $W$ by vectors in $V$. They are the subsets $x + W = \{x + w : w \in W\}$ for some $x \in V$, which we shall often denote by $[x]$ when the subspace $W$ is understood. In particular, $W$ itself is the "zero coset": $[0] = 0 + W = W$. The key observation is that the whole space $V$ gets partitioned into disjoint cosets that fill $V$. The collection of all cosets $[x]$ is the *quotient space* $V/W$. Observe that *points* in the space $V/W$ are at the same time *subsets* in $V$.

LEMMA 1.73. *If $W$ is a subspace in $V$ and $x, y \in V$, then:*

1. *Two cosets $x + W$ and $y + W$ either coincide or are disjoint; hence the distinct cosets of $W$ partition the space $V$.*
2. *An additive coset can have various representatives $x \in V$, so $y + W = x + W \Leftrightarrow$ there is some $w \in W$ such that $y = x + w$ (or $y - x \in W$).*
3. *If $y \in x + W$, then $y + W = x + W$.*

PROOF. We start by defining sums $A + B = \{a + b : a \in A,\ b \in B\}$ of sets $A, B \subseteq V$ that will be invoked repeatedly in what follows.

EXERCISE 1.74. If $W$ is a subspace of a vector space $V$ and $w \in W$, prove that:

(a) $w + W = W$ for all $w \in W$ and $(x + W) \cap W = \emptyset$ if $x \notin W$;
(b) $W + W = \{w_1 + w_2 : w_1, w_2 \in W\}$ is equal to $W$;
(c) $W - W = W$.

Resuming the proof of Lemma 1.73, if cosets $x + W$ and $y + W$ have a point $p$ in common, there are $w_1, w_2 \in W$ such that $x + w_1 = p = y + w_2$; hence $y = x + (w_1 - w_2)$. By Exercise 1.74 the cosets are equal:

$$y + W = (x + (w_1 - w_2)) + W = x + \big((w_1 - w_2) + W\big) = x + W$$

proving 1.73(a).

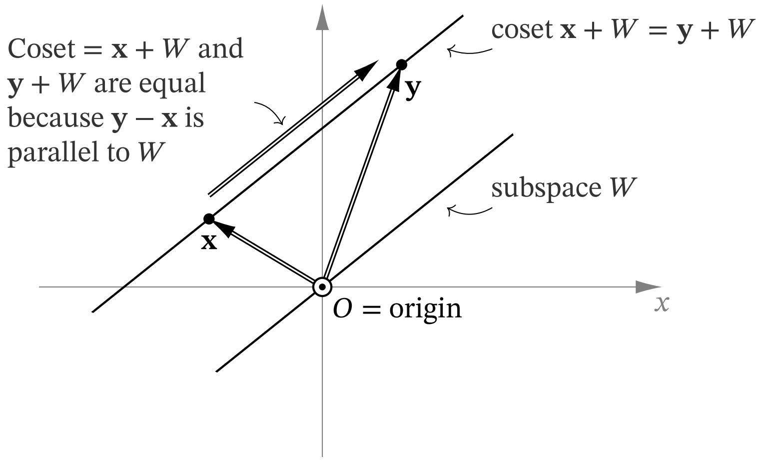


FIGURE 1.6. Additive cosets $\mathbf{x} + W$ of a subspace $W$ are a family of parallel "hyperplanes" in a vector space $V$. When $V = \mathbb{R}^2$ and $W$ is a line through the origin, *all* lines parallel to $W$ are cosets. Two vectors $\mathbf{x}, \mathbf{y}$ in the same coset yield the same translate of $W$: $\mathbf{x} + W = \mathbf{y} + W$ because $\mathbf{y} - \mathbf{x}$ is parallel to the subspace $W$.

For (b), $x+W = y+W \Rightarrow y = y+0 = x+w$ for some $w \in W$. Conversely, if $y = x+w$ for $w \in W$, then $y+W = x+(w+W) = x+W$ (again by Exercise 1.74). For (c), it follows from part (1.) in Lemma 1.73 that $y \in x+W \Rightarrow (y+W) \cap (x+W) \neq \varnothing \Rightarrow y+W = x+W$. □

As an example, if $V = \mathbb{R}^2$ and $W = \{(x, y) \in \mathbb{R}^2 : x = y\}$, the cosets of $W$ are precisely the distinct lines in the plane that make an angle of 45° with the positive $x$-axis. These lines are the "points" in the quotient space $V/W$; see Figure 1.6.

DEFINITION 1.75. There is a natural surjective *quotient map* $\pi : V \to V/W$, given by

$$\pi(x) = [x] = x + W.$$

If $C$ is a coset, any point $v \in C$ such that $C = [v] = v + W$ is called a *representative of the coset*. Part (b) of Lemma 1.73 tells us when two vectors $x, y$ represent the same coset.

**Algebraic Structure in $V/W$.** The quotient space $V/W$ inherits natural sum ($\oplus$) and scaling ($\odot$) operations from the overlying vector space $V$.

DEFINITION 1.76. For any $x, y \in V$ and $\lambda \in \mathbb{K}$ we define the following operations in $V/W$:

1. ADDITION: $[x] \oplus [y] = [x + y]$
2. SCALAR MULTIPLICATION: $\lambda \odot [x] = [\lambda \cdot x]$

To spell out what is involved, this definition tells us how to form the sum $X \oplus Y$ of two cosets $X, Y \in V/W$ via the following algorithm:

1. Pick representatives $x, y \in V$ such that $X = [x],\ Y = [y]$.
2. Add the representatives to get $x + y \in V$.
3. Form the coset $[x + y] = (x + y) + W$ and report the output: $X \oplus Y = [x + y]$.

But why should this make sense? The outcome depends on a choice of representatives for each coset $X, Y$, and if different choices yield different outputs, everything written above is nonsense. Fortuately, the outcome is independent of the choice of representatives and the operation $(\oplus)$ is *well-defined.* In fact, if $[x] = [x']$ and $[y] = [y']$ there must exist $w_1, w_2 \in W$ such that $x' = x + w_1,\ y' = y + w_2$, and then

$$\begin{aligned}[x' + y'] = (x' + y') + W &= (x + y) + ((w_1 + w_2) + W) \\ &= (x + y) + W = [x + y].\end{aligned}$$

Similarly, the scaling operation is well-defined: if $[x'] = [x]$ we have $x' = x + w$ for some $w \in W$, and then

$$[\lambda \cdot x'] = (\lambda \cdot x') + W = (\lambda \cdot x) + (\lambda w + W) = (\lambda \cdot x) + W = [\lambda \cdot x].$$

Once we know the operations $(\oplus)$ and $(\odot)$ make sense, direct calculations involving representatives show that all vector space axioms are satisfied by the system $(V/W, \oplus, \odot)$. For instance,

1. Associativity of $(\oplus)$ on $V/W$ follows from associativity of $(+)$ on $V$: since $x + (y + z) = (x + y) + z$ in $V$ we get

$$\begin{aligned}[x] \oplus ([y] \oplus [z]) &= [x] \oplus [y + z] = [x + (y + z)] \\ &= [(x + y) + z] = [x + y] \oplus [z] = ([x] \oplus [y]) \oplus [z].\end{aligned}$$

2. The zero element is $[0] = 0 + W = W$ because $[0] \oplus [x] = [0 + x] = [x]$.
3. The additive inverse $-[x]$ of $[x] = x + W$ is $[-x] = (-x) + W$ since $[x] \oplus [-x] = [x + (-x)] = [0]$.

EXERCISE 1.77. Verify the remaining vector space axioms for $(V/W, \oplus, \odot)$. Then show that the quotient map $\pi : V \to V/W$ with $\pi(x) = [x] = x + W$ "*intertwines*" the algebraic operations in $(V, +, \cdot)$ with those in $(V/W, \oplus, \odot)$ in the sense that, for any $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$, we have

(a) $\pi(v_1 + v_2) = \pi(v_1) \oplus \pi(v_2)$,
(b) $\pi(\lambda \cdot v_1) = \lambda \odot \pi(v_1)$.

Thus $\pi : V \to V/W$ is a *linear operator* that maps $V$ onto $V/W$.

When $W = (0)$ elements of the quotient are single points $[v] = v + W = \{v\}$, and $V/W$ has a natural identification with $V$ under the quotient map that is now a bijection. When $W = V$ there is just one coset, $v + W = v + V = V$; the quotient space reduces to a single point, the zero element $[0] = 0 + V = V$.

EXERCISE 1.78. Let $V = \mathbb{R}^3$ and $W = \{(x_1, x_2, x_3) : x_3 = 0\}$ = the $xy$-plane in 3-dimensional space. The cosets in $V/W$ are the distinct planes parallel to the $xy$-plane: if $v = (v_1, v_2, v_3)$, then

$$\begin{aligned} v + W &= \{v + w : w \in W\} \\ &= \{(v_1, v_2, v_3) + (w_1, w_2, 0) : w_1, w_2 \in \mathbb{R}\} \\ &= \{(v_1 + s, v_2 + t, v_3) : s, t \in \mathbb{R}\} \\ &= \{(x_1, x_2, x_3) : x_1, x_2 \in \mathbb{R}, x_3 = v_3\} \end{aligned}$$

(the plane parallel to $W$ passing through $(0, 0, v_3)$). Each value of $v_3 \in \mathbb{R}$ gives a different coset.

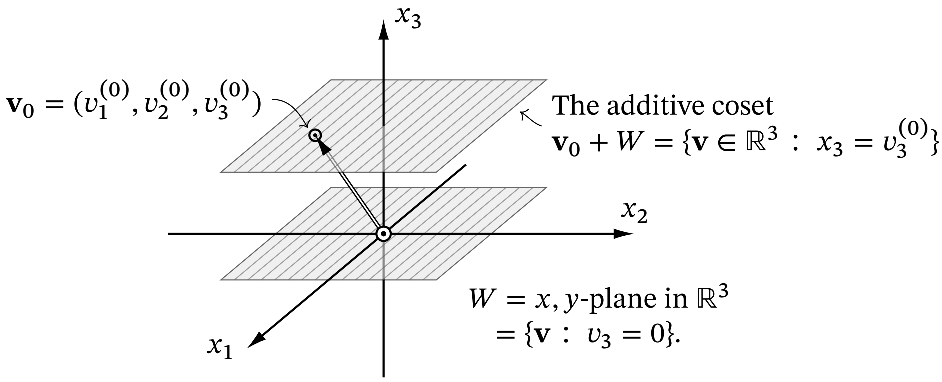


FIGURE 1.7. Additive cosets of $W = \{\mathbf{v} \in \mathbb{R}^3 : v_3 = 0\}$ are planes parallel to $W$ in $\mathbb{R}^3$. A typical coset $\mathbf{v}_0 + W$ is shown. These planes are the elements of the quotient space $\mathbb{R}^3/W$.

One important viewpoint is to think of the quotient map $\pi : v \to V/W$ as "erasing" inessential aspects of the original vector space, retaining only those relevant to the problem at hand. Whole "bunches" of vectors in $V$, the cosets $v + W$, collapse to single points in the target space $V/W$ (the planes in the last example become points in $V/W$). A lot of detail is lost in this collapse, but if $W$ is suitably chosen the quotient space will retain information that is buried in superfluous detail when we look at what is happening in the larger space $V$. We will soon encounter many examples of this once we start looking at the structure of "linear operators" between vector spaces. For the moment we assemble a few more basic facts about quotients of vector spaces.

**Finding Bases in $V/W$ and the Dimension Formula.**

THEOREM 1.79 (Dimension Theorem for Quotients). *If $V$ is finite dimensional and $W$ is a subspace. Then:*

1. $\dim(V/W) \le \dim(V) < \infty$,
2. $\dim(W) \le \dim(V) < \infty$,

*and*

$$\dim(V) = \dim(W) + \dim(V/W). \tag{1.6}$$

*By our notational convention of writing* $\dim_{\mathbb{K}}(V) = |V|$*, this identity can also be written in the abbreviated form* $|V| = |W| + |V/W|$.

PROOF. The quotient map $\pi : V \to V/W$ sends $V$ onto $V/W$ and preserves linear combinations in the sense that

$$\pi\Big(\sum_{i=1}^{m} \lambda_i v_i\Big) = \sum_{i=1}^{m} \lambda_i \pi(v_i).$$

(recall Exercise 1.77). Furthermore, $\pi(V) = V/W$, so if vectors $\{v_i\}$ span $V$ their images $\bar{v}_i = \pi(v_i)$ span $V/W$. That proves

$$\dim(V/W) \le \#\{\bar{v}_i\} \le \#\{v_i\} = \dim(V) < \infty$$

as claimed in (1.).

As for item (2.), we know $\dim(V) < \infty$ but have no a priori information about $W$; however, we showed earlier that no independent set in $V$ can have more than $\dim(V)$ elements, and a basis for $W$ would be such a set.

The identity (1.6) is proved by constructing a basis in $V/W$ aligned with a specially chosen basis in $V$. Since $\dim(W) < \infty$ there is a basis $\{w_1, \dots, w_m\}$ in $W$. If $W = V$ then $V/W$ is trivial and there is nothing more to do, but otherwise we can find an "outside vector" $v_{m+1} \notin W$ such that the larger set $\{w_1, \dots, w_m, v_{m+1}\}$ is independent, and hence a basis for

$$W_1 = \mathbb{K}\text{-span}\{w_1, \dots, w_m, v_{m+1}\} \supsetneqq W_0 = W.$$

If $W_1 \neq V$, we can adjoin one more vector $v_{m+2} \notin W_1$ to get an independent set $\{w_1, \dots, w_m, v_{m+1}, v_{m+2}\}$ with

$$W_0 \subsetneqq W_1 \subsetneqq W_2 = \mathbb{K}\text{-span}\{w_1, \dots, w_m, v_{m+1}, v_{m+2}\}.$$

This process must terminate; otherwise we would have arbitrary large independent sets in the finite-dimensional space $V$. When the construction terminates we get an independent spanning set $\{w_1, \dots, w_m, v_{m+1}, \dots, v_{m+k}\}$ in $W_k = V$. This is a basis for $V$ so $\dim(V) = m + k = \dim(W) + k$.

To conclude the proof we demonstrate that $k = \#(\text{outside vectors})$ is equal to $\dim(V/W)$ by showing that the $\pi$-images $\overline{v}_{m+1}, \dots, \overline{v}_{m+k} \in V/W$ of the outside vectors are a basis for $V/W$. Since $\pi$ is surjective the images $\pi(w_1), \dots, \pi(v_{m+k})$ span $V/W$. But $\pi$ "kills" all vectors in $W$, so

$$\pi(w_1) = \cdots = \pi(w_m) = [0] \quad \text{in } V/W,$$

and the remaining images $\overline{v}_{k+i} = \pi(v_{m+i})$ span $V/W$. They are also linearly independent, because if some linear combination $\sum_{i=1}^{k} c_{m+i}\overline{v}_{m+i} = [0]$ in $V/W$, then by linearity of the quotient map $\pi$ we get

$$[0] = \sum_{j=1}^{k} c_{m+j}\pi(v_{m+j}) = \pi\Big(\sum_{j=1}^{k} c_{m+j}v_{m+j}\Big).$$

But $\pi(v) = [0]$ for a vector $v \in V \Leftrightarrow [v] = v + W$ is equal to the zero coset $[0] = W$. Since $v + W = W \Leftrightarrow v \in W$, we can find coefficients $c_1, \dots, c_m$ such that

$$\sum_{i=1}^{m} c_i w_i = v = \sum_{j=1}^{k} c_{m+j} v_{m+j}$$

or

$$0 = \sum_{i=1}^{m} c_i w_i + \sum_{j=1}^{k} (-1) c_{m+j} v_{m+j} \quad \text{in } V.$$

Since $w_1, \dots, w_m, v_{m+1}, \dots v_{m+k}$ is a basis for $V$, this can only happen if all coefficients in this sum are zero, and in particular $c_{m+1}, \dots, c_{m+k} = 0$. Thus the $\{\overline{v}_i\}$ are independent and a basis for $V/W$, so $\dim(V/W) = k$ is equal to $\dim(V) - \dim(W)$. □

REMARK 1.80. The construction developed in proving Theorem 1.79 shows how to find bases in a quotient space $V/W$ and perform effective calculations with them. The key was to find representatives $v_i$ back in $V$ in order to transfer calculations involving cosets in $V/W$ to calculations involving actual vectors $v_i$ in $V$. This was achieved by finding a basis $\mathfrak{X} = \{w_1, \dots, w_r; u_{r+1}, \dots, u_n\}$ in $V$ such that $w_1, \dots, w_r$ are a basis for $W$. As a corollary" we can read the following observations out of that discussion.

COROLLARY 1.81. *If $W$ is a subspace in a finite-dimensional vector space $V$, and if $\mathfrak{X} = \{w_1, \dots, w_r; u_{r+1}, \dots, u_n\}$ is a basis in $V$ such that $\{w_1, \dots, w_r\}$ is a basis for $W$, then*:

1. *The cosets $v_{r+1} + W, \dots, v_n + W$ in $V/W$ corresponding to the "outside vectors" in $\mathfrak{X}$ are a basis for the quotient space $V/W$.*
2. *The linear span $H = \mathbb{K}\text{-span}\{v_{r+1}, \dots, v_n\}$ in $V$ is a subspace of $V$ that "cross-sections" the cosets in $V/W$: each coset $x + W$ meets $H$ in exactly one point $y \in V$ so that $H \cap (x + W) = \{y\}$.*

*Thus each point in $V/W$ is labeled by a unique point in $H$. The first statement can be read directly out of the proof of Theorem* 1.79. *Proof of* (2.) *is left as an exercise.*

EXERCISE 1.82. If $H$ is the linear span of the outside vectors $u_{r+1}, \dots, u_n$ in Corollary 1.81, explain why:

(a) Each coset $x + W \in V/W$ meets $H$ in a single point $y \in H$.
(b) The quotient map $\pi : V \to V/W$ restricts to a bijection $\pi|_H$ between $H$ and $V/W$.

EXERCISE 1.83. Consider the subspace $W = \{\mathbf{x} \in \mathbb{R}^n : x_1 + \cdots + x_n = 0\}$ in $V = \mathbb{R}^n$.

(a) What is its dimension?
(b) Prove that the following vectors are a basis for $W$:

$$(1, -1, 0 \dots, 0),\ (0, 1, -1, 0 \dots 0),\ \dots\ , (0, \dots, 1, -1),\ (-1, 0, \dots, 0, 1)$$

(c) Find a vector subspace $H \subseteq \mathbb{R}^n$ that cross-sections the cosets $\mathbf{x}+W$ in the quotient space $V/W$. Explain why $H$ must be a one-dimensional subspace in $V$.

EXERCISE 1.84. Find explicit bases for the following quotient spaces $V/W$.

(a) $V = \mathbb{R}^3$, $W = \mathbb{R}\mathbf{e}_1 + \mathbb{R}\mathbf{e}_2$

(b) $V = \mathbb{R}^3$, $W = \mathbb{R}\text{-span}\{\mathbf{w}_1 = (1,2,3), \mathbf{w}_2 = (0,1,-1)\}$

(c) $V = \mathbb{C}^4$,

$$W = \mathbb{C}\text{-span}\{\mathbf{z}_1 = (1, 1+i, 3-2i, -i),\ \mathbf{z}_2 = (4-i, 0, -1, 1+i)\}$$

(d) $V = \mathbb{R}^4$,

$$W = \{\mathbf{x} \in \mathbb{R}^4 : x_1 + x_2 - x_3 + x_4 = 0 \text{ and } 4x_1 - 3x_2 + 2x_3 + x_4 = 0\}.$$

Here is a simple example involving bases in a quotient space $V/W$.

EXAMPLE 1.85. Let $V = \mathbb{R}^4$ and $W = \{\mathbf{x} \in \mathbb{R}^4 : 2x_1 - x_2 + x_4 = 0\}$. The subspace $W$ is the solution set of the matrix equation

$$A\mathbf{x} = \mathbf{0} \quad \text{where} \quad A = [\,2, -1, 0, 1\,]_{1\times 4}$$

that imposes a single linear constraint on $\mathbb{R}^4$. Find a basis for $V/W$.

SOLUTION: Row operations yield

$$A \to A' = \left[\,\boxed{1}\,, -\frac{1}{2}\,, 0\,, \frac{1}{2}\,\right]$$

The free variables are $x_2, x_3, x_4$ and $x_1 = \frac{1}{2}x_2 - \frac{1}{2}x_4$, so the solutions have the form

$$\mathbf{x} = \begin{pmatrix} \frac{1}{2}x_2 - \frac{1}{2}x_4 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

for $x_2, x_3, x_4 \in \mathbb{R}$. Thus the solution set for $A\mathbf{x} = \mathbf{0}$ is the linear span of the column vectors

$$\mathbf{u}_1 = \text{col}(1,2,0,0), \quad \mathbf{u}_2 = \text{col}(0,0,1,0), \quad \mathbf{u}_3 = \text{col}(-1,0,0,2),$$

which are a basis for $W$ since they are easily seen to be linearly independent by row reducing the $3 \times 4$ matrix $M$ that has these vectors as its rows

$$M = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix}$$

to see if you get a row of zeros. You do not. Therefore $\dim(\text{RowSpace}(M)) = 3$ and the vectors are independent.

Since $\dim(V) = \dim(W) + \dim(V/W)$ and $\dim(W) = 3$, we need only find one outside vector $\mathbf{u}_4 \notin W$ to complete a basis for $V = \mathbb{R}^4$; then the coset $\pi(\mathbf{u}_4) = \mathbf{u}_4 + W$ will be nonzero, and a basis vector for the 1-dimensional quotient space. The vector $\mathbf{u}_4 = \mathbf{e}_4 = (0,0,0,1)$ is not in $W$ because it fails to satisfy the constraint equation $2x_1 - x_2 + x_4 = 0$. Thus the single vector $[\mathbf{e}_4] = \pi(\mathbf{e}_4) = \mathbf{e}_4 + W$ is a basis for $V/W$, and $\dim(V/W) = 1$.

EXERCISE 1.86. Let $W$ be the vector subspace in $\mathbb{R}^4$ determined by the equations

$$\begin{aligned} x_1 - \phantom{3}x_2 + x_3 + x_4 &= 0 \\ 4x_1 - 3x_2 \phantom{+ x_3} - x_4 &= 0 \end{aligned}$$

(a) Find a basis $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ for $W$. What are the dimensions of $W$ and $V/W$?
(b) Find additional vectors $\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_4\}$ lying outside of $W$ such that

$$\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_4 \text{ is a basis for } \mathbb{R}^4.$$

(c) Find a system $A\mathbf{x} = \mathbf{0}$ of $4 - r$ linear equations in the unknowns $x_1, \dots, x_4$ whose solution set is $H = \mathbb{R}\text{-span}\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_4\}$, the linear span of the outside vectors found in (b).

HINT: Some trial-and-error may be needed to find the outside vectors requested in (b). A vector lies outside of $W$ if it fails to satisfy the equations that define $W$. The outside vectors you seek are not unique.

Thus the points in $H$ uniquely label the cosets in $V/W$.

## Additional Exercises

### Section 1.1. Vector Spaces

**1.** TRUE/FALSE QUESTIONS: ("True" if the statement is *always* true.)
(a) In any vector space $c \cdot x = c \cdot y$ implies $x = y$.
(b) In any vector space $c \cdot x = c' \cdot x$ implies $c' = c$.
(c) If $W$ is a subspace of a finite-dimensional vector space $V$ and $\dim V = \dim W$, then $W = V$.
(d) $\deg(f + h) = \deg(f) + \deg(h)$ for nonzero polynomials $f, h$ in $\mathbb{K}[x]$.
(e) $\deg(f + h) = \max\{\deg(f), \deg(h)\}$ for nonzero polynomials $f, h$ in $\mathbb{K}[x]$,
(f) For $f \in \mathbb{K}[x]$ and $\lambda \in \mathbb{K}$, $\deg(\lambda \cdot f) = \deg(f)$.
(g) A "constant polynomial" $c\mathbb{1}$ in $\mathbb{K}[x]$ has degree zero.
(h) $\deg(f \cdot h) = \deg(f) + \deg(h)$ for all nonzero polynomials $f, h$ in $\mathbb{K}[x]$.
(i) $\deg(f \cdot h) = \deg(f) + \deg(h)$ for all nonzero polynomials $f, h$ in $\mathbb{K}[x_1, \dots, x_n]$.

**2.** For $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$ in $\mathbb{R}^2$ define operations $(*)$ and $(+)$, letting

$$\text{(i)} \ \ \mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2),$$

$$\text{(ii)} \ \ \lambda * \mathbf{a} = \begin{cases} (0,0) & \text{if } \lambda = 0, \\ \left(\lambda a_1, \frac{1}{\lambda} a_2\right) & \text{if } \lambda \neq 0. \end{cases}$$

Is $V$ a vector space when equipped with these operations? If not, exhibit a vector space axiom not satisfied by the system $(V, +, *)$.

**3.** In $\mathbb{R}^2$ the parallelogram law for addition of two vectors $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$ based at the origin was illustrated in Figure 1.3. Explain why this geomeangle/tric procedure for adding vectors is equivalent to the algebraic rule

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2)$$

when vector addition is described in Cartesian coordinates.

**Section 1.2. Vector Subspaces**

**1.** TRUE/FALSE QUESTIONS: ("True" if the statement is *always* true.)
(a) The set $W = \{(x_1, x_2, x_3) : x_1 = x_3 + 2 \text{ and } x_1 - x_2 + x_3 = 0\}$ is a vector subspace in $\mathbb{R}^3$.
(b) A nonzero vector space $V$ contains a "proper" subspace $W$ (one such that $\{0\} \subsetneqq W \subsetneqq V$.)
(c) The union $V \cup W$ of two vector subspaces of $V$ is a subspace.
(d) The set of odd polynomials (coefficients $c_n = 0$ for all even $n$) is a vector subspace of $\mathbb{K}[x]$.
(e) The set of polynomials with coefficients $c_n = 0$ for all indices $n \geq$ some cutoff $N$ is a vector subspace of $\mathbb{K}[x]$.
(f) The number of nonzero entries in an upper-triangular $n \times n$ matrix is $\leq \frac{1}{2}n^2$.
(g) The set $W$ consisting of the zero vector $\mathbf{0}$ and all *nonconstant* polynomials is a vector subspace of $\mathbb{K}[x]$.

**2.** Fix some integer $r \geq 1$ and let $E_r$ be all polynomials of the form $f = \sum_{n \geq r} c_n x^n$ (finite sums). Show that:
(a) $E_r$ is a vector subspace of $\mathbb{K}[x]$.
(b) $E_r$ is also a *subalgebra* of $\mathbb{K}[x]$, so $f, h \in E \Rightarrow f \cdot h$ is also in $E$.

NOTE: $E_r$ consists of all polynomials that have a "*zero of order at least* $(r-1)$ *at the origin.*"

**3.** If $W_1$ and $W_2$ are subspaces of a vector space $V$, prove that $S = W_1 \cup W_2$ is a subspace $\Leftrightarrow W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

**4.** The *sum* $S_1 + S_2$ of two nonempty subsets in a vector space $V$ is defined to be

$$S_1 + S_2 = \{a + b : a \in S_1 \text{ and } b \in S_2\}$$

(a) If $S_1$ and $S_2$ are both vector subspaces, prove that their sum $S_1 + S_2$ is also a vector subspace.
(b) This is no longer true if just one of the sets is a vector subspace. In $\mathbb{R}^2$ describe the sum of sets $S_1 + S_2$ when
(i) $S_1 = \{(x_1, x_2) : x_1 - x_2 = 0\}$ (a line through the origin).
(ii) $S_2$ is the open disc $S_2 = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$.
Make a sketch of the set $S_1 + S_2$.

**5.** The transpose $A^{\mathsf{T}}$ of a matrix $A$ is obtained by reflecting all entries in the array across the diagonal, so that $(A^{\mathsf{T}})_{ij} = A_{ji}$. Obviously $(A^{\mathsf{T}})^{\mathsf{T}} = A$. A matrix in $\mathrm{M}(n, \mathbb{K})$ is *symmetric* if $A_{ji} = A_{ij}$ for all indices $1 \leq i, j \leq n$ and is *antisymmetric* if $A_{ji} = -A_{ij}$.

(a) Prove that the sets

$$\mathcal{S}_n = (\text{symmetric matrices})$$
$$\mathcal{A}_n = (\text{antisymmetric matrices})$$

are both vector subspaces of matrix space $\mathrm{M}(n, \mathbb{K})$.

(b) If $\mathbb{K}$ is a field in which $1_{\mathbb{K}} + 1_{\mathbb{K}} \neq 0_{\mathbb{K}}$, explain why every square matrix $A$ is a sum $A = A_s + A_a$ of a symmetric and an antisymmetric matrix, so the sum of sets $\mathcal{S}_n + \mathcal{A}_n$ is all of $\mathrm{M}(m, \mathbb{K})$.

(c) If $\mathbb{K}$ is a field in which $1_{\mathbb{K}} + 1_{\mathbb{K}} \neq 0_{\mathbb{K}}$, prove that the decomposition $A = A_s + A_a$ is *unique*.

HINTS: If the decomposition in (c) fails to be unique for some matrix $A$ there would be a related matrix $B$ that is both symmetric and antisymmetric. That is impossible unless we have $1 + 1 = 0$ in the ground field. (This can happen, for instance, when $\mathbb{K} = \mathbb{Z}_2$, but not in fields such as $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$.)

## Section 1.3. Solving Systems of Linear Equations

**1.** TRUE/FALSE QUESTIONS: ("True" if the statement is *always* true.)

(a) The linear span of the empty set $\varnothing$ is the empty set.

(b) The linear span of a nonempty subset $S$ in a vector space $V$ is the intersection of all vector subspaces $W \subseteq V$ that contain $S$.

(c) In solving a system of linear equations, the solution set is unaffected if you multiply all terms in one of the equations by the same constant $c \in \mathbb{K}$.

(d) In solving a system of linear equations, the solution set is unaffected if you multiply each equation by a different randomly chosen nonzero constant.

(e) In solving a system of linear equations, the solution set is unaffected if you add a multiple of one equation to a different equation.

(f) Every system of linear equations in $n$ unknowns has at least one solution.

(g) If $A$ is an $n \times m$ matrix, $\mathbf{x} \in \mathbb{K}^m$, and $\mathbf{b} \in \mathbb{K}^n$, the solution set of the matrix equation $A\mathbf{x} = \mathbf{b}$ is a vector subspace of $\mathbb{K}^n$.

**2.** A homogeneous system $A\mathbf{x} = \mathbf{0}$ of $n$ linear equations in $m$ unknowns $x_1, \dots, x_m$ is *overdetermined* if there are more constraint equations than unknowns (so $n > m$). Do such systems have $x_1 = \cdots = x_n = 0$ as their only solution? Explain.

**3.** Solve the following systems of equations using row operations:

(a) $\begin{cases} x_1 + 2x_2 + 2x_3 = 2 \\ x_1 + 2x_2 + 8x_3 + 5x_4 = -16 \\ x_1 + x_2 + 5x_3 + 5x_4 = 3 \end{cases}$ (b) $\begin{cases} x_1 + 2x_2 + 6x_3 = -1 \\ 2x_1 + x_2 + x_3 = 8 \\ 3x_1 + x_2 - x_3 = 15 \\ x_1 + 3x_2 + 10x_3 = -5 \end{cases}$

**4.** Determine the intersection of the following subspaces in $\mathbb{R}^3$:

(i) $W_1 = \{(a_1, a_2, a_3) : a_1 = 3a_2 \text{ and } a_3 = -a_2\}$
(ii) $W_2 = \{(a_1, a_2, a_3) : 2a_1 - 7a_2 + a_3 = 0\}$

by finding a set of basis vectors for $W_1 \cap W_2$. What is the dimension of this intersection?

HINT: The intersection is the solution set of all the equations defining $W_1$ and $W_2$.

**5.** For each list of vectors given below decide whether the first vector can be written as a linear combination of the other two.

(a) In $\mathbb{R}^3$: $(1, 2, 3)$; $(-3, 2, 1), (2, -1, -1)$
(b) In $\mathbb{K}[x]$: $-2x^3 - 11x^2 + 3x + 2$; $x^3 - 2x^2 + 3x - 1$, $2x^3 + x^2 + 3x - 2$

**6.** If $\mathbf{b} \in \mathbb{R}^n$, $A$ is a matrix in $\mathrm{M}(n, \mathbb{R})$, and $\widetilde{A} = \lambda \cdot A$ is a scalar multiple with $\lambda \neq 0$, how are the solution sets of the following matrix equations related?

(a) $A\mathbf{x} = \mathbf{0}$ and $\widetilde{A}(\mathbf{x}) = \mathbf{0}$
(b) $A\mathbf{x} = \mathbf{b}$ and $\widetilde{A}(\mathbf{x}) = \lambda \cdot \mathbf{b}$

If $H = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ and $\widetilde{H} = \{\mathbf{x} : \widetilde{A}\mathbf{x} = \mathbf{b}\}$,

(c) show that the solution set $\widetilde{H}$ is equal to $\frac{1}{\lambda} \cdot H$.

**7.** A "line" in $\mathbb{R}^3$ is any translate $L = \mathbf{x}_0 + E$ of some 1-dimensional subspace $E$; a "hyperplane" is any translate $H = \mathbf{x}_0 + E$ of some 2-dimensional vector subspace $E$. Answer the questions below about the following set of vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = (2, -5, 1), \quad \mathbf{v}_2 = (0, 4, 6), \quad \mathbf{v}_3 = (-3, 7, 1)$$

(a) Are these vectors *collinear*? Do they lie in some line $L \subseteq \mathbb{R}^3$ (not necessarily one passing through the origin)?
(b) Prove that they lie in some hyperplane $H$ by exhibiting (i) a 2-dimensional subspace $E$ (determined by exhibiting basis vectors $\mathbf{f}_1, \mathbf{f}_2$), and (ii) a vector $\mathbf{x}_0 \in \mathbb{R}^3$ such that the given vectors all lie in the hyperplane $H = \mathbf{x}_0 + E$.
(c) Use (b) to provide an explicit "parametric" description of $H$ as the range of a map $(x_1, x_2, x_3) = \phi(s_1, s_2)$ from $\mathbb{R}^2$ into $\mathbb{R}^3$ that sends $\mathbb{R}^2$ bijectively to the hyperplane $H$.
(d) Find a $3 \times 2$ matrix $A$ and a vector $\mathbf{b} \in \mathbb{R}^3$ such that $H$ is the solution set of the inhomogeneous linear system $A\mathbf{x} = \mathbf{b}$.

HINT: Start by considering some difference vectors $\mathbf{v}_i - \mathbf{v}_j$.

NOTE: The vector subspace $E$ is uniquely determined, but the translation vector $\mathbf{x}_0$ is not, since it can be replaced by $\mathbf{y}_0 = \mathbf{x}_0 + \mathbf{e}$ where $\mathbf{e} \in E$. However, all such choices of the translation vector yield the same hyperplane $\mathbf{x}_0 + E$ in $\mathbb{R}^3$ because $\mathbf{e} + E = E$ when $\mathbf{e} \in E$ (Figure 1.4 suggests what is happening.)

**Section 1.4. Linear Span, Independence, and Bases**

**1.** TRUE/FALSE QUESTIONS: ("True" if the statement is *always* true.)

(a) If $S = \{v_1, \dots, v_r\}$ is a linearly dependent set containing at least two vectors, then $v_1$ is a linear combination of the *other* vectors in $S$.

(b) A set $S$ that contains the zero vector cannot be independent.

(c) A nonempty subset of a set $S$ whose vectors are linearly independent is itself linearly independent.

(d) A nonempty subset of a set $S$ whose vectors are linearly dependent is always linearly dependent.

(e) If $c_1v_1 + \cdots + c_mv_m = 0$ in a vector space $V$ and the $v_i$ are independent, then all coefficients $c_i$ must be zero.

(f) A nonempty set $S$ of linearly dependent vectors always contains a subset $S'$ that is independent and has the same linear span as $S$.

(g) If a linearly independent set $S$ contains at least two vectors and we remove one vector, the remaining set $S'$ is independent but $\mathbb{K}\text{-span}(S') \subsetneqq \mathbb{K}\text{-span}(S)$.

(h) For $n > 2$, the set of rank-2 matrices in $\mathrm{M}(n, \mathbb{K})$ is a vector subspace of $\mathrm{M}(n, \mathbb{K})$.

**2.** Let $E = \mathbb{R}\text{-span}\{M_1, M_2, , M_3, M_4\}$ be the linear span of the $2 \times 2$ matrices in $\mathrm{M}(2, \mathbb{R})$ listed below.

$$M_1 = \begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Use row operations to do the following:

(a) Verify that the vectors $\{M_1, \dots, M_4\}$ are not linearly independent.

(b) Find a set of matrices $A_1, \dots, A_r$ in $\mathrm{M}(2, \mathbb{R})$ that are a basis for $E$. What is $\dim(E)$?

(c) Find a proper subset $S \subseteq \{M_1, \dots, M_4\}$ that is independent and spans $E$, and hence is a basis for $E$.

HINT: You can identify $2 \times 2$ matrices with elements of $\mathbb{R}^4$ by listing the matrix entries as 4-tuples $(a_{11}, a_{12}, a_{21}, a_{22})$. $\mathrm{M}(2, \mathbb{R})$ and $\mathbb{R}^4$ are isomorphic as vector spaces.

**3.** Consider the following vectors in $\mathbb{K}^3$

$$\mathbf{u}_1 = (1, 2, 3), \quad \mathbf{u}_2 = (2, 3, 4), \quad \mathbf{u}_3 = (2, 3, 4)$$

(a) Are these vectors independent? Do they span $\mathbb{K}^3$?

(b) Describe their linear span $W$ by identifying a set of basis vectors.

(c) Does the vector $\mathbf{b} = (2, -1, -1)$ lie in the subspace $W$?

**4.** In matrix space $\mathrm{M}(n,\mathbb{K})$ the "matrix units" $E_{ij}$ $(1 \le i,j \le n)$ are defined by

$$E_{ij} = \text{ the matrix with entries } (E_{ij})_{rs} = \begin{cases} 1 & \text{if } r = i \text{ and } s = j \\ 0 & \text{for all other } (r,s) \end{cases}$$

so the only nonzero entry in $E_{ij}$ is a "1" in the $(i,j)$ spot.

(a) Write down the four matrix units $E_{ij}$ in $\mathrm{M}(2,\mathbb{K})$.

(b) Explain why the matrix units $\{E_{ij} : 1 \le i,j \le n\}$ are a basis for $\mathrm{M}(n,\mathbb{K})$.

(c) Prove that products of these units are given by the simple law

$$E_{ij} \cdot E_{k\ell} = \delta_{i,k} \cdot \delta_{j,\ell} \cdot E_{i\ell}$$

where $\delta_{pq}$ is the *Kronecker delta symbol*, which is equal to 1 when $p = q$ and to 0 otherwise.

NOTE: Property (c) is what makes the basis $E_{ij}$ so handy in matrix calculations. Once it has been verified, (c) eliminates the need to carry out tedious row/column multiplications in computing matrix products.

**5.** If nonzero polynomials $f_1, \dots, f_n$ in $\mathbb{K}[x]$ have different degrees explain why this set of polynomials is linearly independent.

**6.** Show that the matrices

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

are a basis for the subspace $\mathcal{S}_2$ of symmetric matrices in $\mathrm{M}(2,\mathbb{K})$.

**7.** What is the dimension of the space of symmetric matrices $\mathcal{S}_n$ that are in $\mathrm{M}(n,\mathbb{C})$?

**8.** The *trace* of a matrix $A \in \mathrm{M}(n,\mathbb{K})$ is $\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}$ (sum of the diagonal entries). Verify that $E = \{A \in \mathrm{M}(n,\mathbb{K}) : \mathrm{Tr}(A) = 0\}$ is a vector subspace and find a basis. What is $\dim(E)$?

**9.** Explain why the polynomials $1, x, x^2, x^3, \dots, x^n, \dots$ are a basis for the infinite-dimensional vector space $\mathbb{K}[x]$.

**10.** Let $E$ be the linear span of the following vectors in $\mathbb{R}^4$

$$\mathbf{v}_1 = (1,0,-1,2), \quad \mathbf{v}_2 = (4,5,1,-9), \quad \mathbf{v}_3 = (2,1,1,-3),$$
$$\mathbf{v}_4 = (-2,1,0,-2), \quad \mathbf{v}_5 = (-1,1,-1,-1)$$

(a) Find a basis for $E$. What is $\dim(E)$?

(b) A basis can also be found by deleting certain vectors from the original list. Exhibit such a basis.

**11.** If $S_1 \subseteq S_2$ are nonempty sets in a vector space $V$,

(a) Explain why $\mathbb{K}\text{-span}(S_1) \subseteq \mathbb{K}\text{-span}(S_2)$.

(b) Give an example of sets in $\mathbb{K}^n$ such that

$$S_1 \subseteq S_2, \; S_1 \ne S_2, \text{ and yet } \mathbb{K}\text{-span}(S_1) = \mathbb{K}\text{-span}(S_2).$$

**12.** Verify that $W = \{\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{C}^n : z_1 + \cdots + z_n = 0\}$ is a subspace of the complex vector space $V = \mathbb{C}^n$. Find a $\mathbb{C}$-basis and the dimension of $E$ as a vector space over $\mathbb{C}$.

**13.** Let $E$ be the linear span in $\mathbb{K}^4$ of the vectors

$$\mathbf{v}_1 = (1,2,3,4), \quad \mathbf{v}_2 = (2,0,4,5), \quad \mathbf{v}_3 = (-1,1,-3,-4), \quad \mathbf{v}_4 = (2,3,4,5)$$

(a) Give a parametric description of $E$ by finding a basis. What is $\dim(E)$?

(b) Give an implicit description: find a system of equations $A\mathbf{x} = \mathbf{0}$ whose solution set is equal to $E$.

**14.** Suppose a finite-dimensional vector space $V$ contains subspaces $E_1, E_2$ and that $\dim V = 10$, $\dim E_1 = 7$, $\dim E_2 = 5$. Let $E_1 + E_2$ be the vector subspace spanned by $E_1$ and $E_2$.

(a) What is the *maximum* possible dimension for $E_1 + E_2$? Give an example of subspaces in $\mathbb{R}^{10}$ for which maximum dimension is achieved.

(b) What is the *minimum* possible dimension for $E_1 + E_2$? Give an example of subspaces in $\mathbb{R}^{10}$ for which the minimum is achieved.

(c) Can $V$ be a direct sum $V = E_1 \oplus E_2$ of these subspaces?

(d) What is the maximum possible dimension for $E_1 \cap E_2$?

(e) If $E_1, E_2 \subseteq \mathbb{R}^{10}$, find the minimum possible dimension of $E_1 \cap E_2$. What if they lie in $\mathbb{R}^{15}$?

**15.** In $V = \mathbb{R}^4$ consider the subspaces $E_1 = \mathbb{R}\text{-span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $E_2 = \mathbb{R}\text{-span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where

$$\mathbf{v}_1 = (1,0,-1,2), \qquad \mathbf{u}_1 = (3,2,4,12),$$
$$\mathbf{v}_2 = (2,1,1,8), \qquad \mathbf{u}_2 = (1,1,3,4),$$

which we may regard as the rows of a $4 \times 4$ matrix

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 2 & 1 & 1 & 8 \\ 3 & 2 & 4 & 12 \\ 1 & 1 & 3 & 4 \end{pmatrix}$$

Suppose elementary row operations on $A$ and its transpose $A^{\mathsf{T}}$ yield the following echelon forms:

$$A \xrightarrow[\text{ops}]{\text{row}} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad A^{\mathsf{T}} \xrightarrow[\text{ops}]{\text{row}} \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Use this information to answer the following questions.

(a) What are the column ranks of $A$ and $A^{\mathsf{T}}$?

(b) Find a basis in $\mathbb{R}^4$ for the linear span $E_1 + E_2$. What is the dimension of $E_1 + E_2$?

(c) What is $\dim(E_1 \cap E_2)$?

HINT: For (c) recall Exercise 1.67. Explain your answers.

**16.** For the spaces $E_1, E_2$ in Additional Exercise 15:
(a) Find the dimension of the intersection $E_1 \cap E_2$.
(b) Find a basis for $E_1 \cap E_2$.

NOTE: Part (a) can be answered from the data given in Additional Exercise 15 using the dimension formula of Exercise 1.67. The usual efforts with row operations do not yield an answer in part (b); a solution requires direct calculations involving the basis vectors $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{v}_1, \mathbf{v}_2$.

HINTS: Find implicit descriptions of the $E_i$ as solutions of matrix equations $A_i\mathbf{x} = \mathbf{0}$ for $\mathbf{x} \in \mathbb{R}^4$. To be in $E_1 \cap E_2$ a vector $\mathbf{x}$ must be a solution for each system. Combining these systems yields a system of 4 equations in 4 unknowns whose solution set is precisely $E_1 \cap E_2$. Finding a basis for the solution set of the combined system is then a routine matter.

**17.** In Example 1.37 we determined the linear span $E$ of particular vectors $\mathbf{u}_1, \dots, \mathbf{u}_5$ in $\mathbb{K}^3$. Regarding them as column vectors we formed the $3 \times 5$ matrix $A = \text{col}(\mathbf{u}_1; \dots; \mathbf{u}_5)$, and then found $E$ using the fact that the span is precisely the set of vectors $\mathbf{b} \in \mathbb{K}^3$ for which the system $A\mathbf{x} = \mathbf{b}$ has solutions.

In Example 1.40 we showed that $E$ could be calculated more directly using row operations on the $5 \times 3$ transpose matrix $B = A^{\mathrm{T}}$ whose rows are the $\mathbf{u}_i$ regarded as row vectors, so $R_i(B) = \mathbf{u}_i$. Carry out this approach, using row operations to find basis vectors for $E$.

NOTE: Both approaches must yield sets of vectors that span $E$, but a vector space can have many different bases.

## Section 1.5. Quotient Spaces.

**1.** TRUE/FALSE QUESTIONS: ("True" if statement is *always* true.)
(a) If $W$ is a one-dimensional subspace in $V = \mathbb{R}^3$ (a line through the origin), the quotient space $V/W$ can be interpreted as the set of all lines in $\mathbb{R}^3$ that are parallel to $W$.
(b) If $V \supseteq W$ and both vector spaces are infinite dimensional, the quotient space $V/W$ must also be infinite dimensional.
(c) If $E \supseteq V \supseteq W$ are subspaces in a finite-dimensional vector space, then $\dim E/W = \dim E/V + \dim V/W$.

**2.** How are the quotient space $V/W$ and the (linear) quotient map $\pi : V \to V/W$ with $\pi(x) = x + W$ in $V/W$ to be interpreted when:
(a) $W = \{0\}$ (the trivial subspace in $V$)? (b) When $W = V$?

**3.** Find the dimension $\dim(V/W)$ and an explicit basis for this quotient space when $V = \mathbb{R}^4$ and

$$W = \{\mathbf{x} \in \mathbb{R}^4 : x_1 + x_2 - x_3 - x_4 = 0 \;\text{ and }\; 4x_1 - 3x_2 - x_4 = 0\}$$

It will suffice to give coset representatives $\mathbf{v} \in \mathbb{R}^4$ for the basis vectors $\mathbf{v} + W$ in $V/W$.

**4.** Let $W$ be the linear span of the following vectors in $V = \mathbb{C}^5$:

$$\mathbf{v}_1 = (1,1,2,0,3), \quad \mathbf{v}_2 = (3,2,1,5,-1), \quad \mathbf{v}_3 = (2,1,2,0,-1)$$

(a) Find vectors $\mathbf{e}_{i_1}, \mathbf{e}_{i_2}$ selected from the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_5$ in $V = \mathbb{C}^5$ such that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_{i_1}, \mathbf{e}_{i_2}\}$ is a basis for $V$.
(b) Prove that the cosets $\overline{\mathbf{e}}_{i_1}, \overline{\mathbf{e}}_{i_2}$ are a basis for the quotient space $V/W$.

**5.** What is $\dim(E_1 + E_2)/(E_1 \cap E_2)$ in Additional Exercise 14 for Section 1.4?

**6.** If $E_1, E_2$ are subspaces of $V$, find formulas for calculating the dimensions of the following quotient spaces in terms of $|V|, |E_1|, |E_2|$, and $|E_1 \cap E_2|$:
(a) $(E_1 + E_2)/E_1$
(b) $(E_1 + E_2)/(E_1 \cap E_2)$

## Appendix: The Degree Formula for $\mathbb{K}[\mathbf{x}_1, \dots, \mathbf{x}_N]$

Let $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \dots, x_N]$ be the algebra of polynomials with coefficients in a field $\mathbb{K}$ and let $1$ (constant polynomial $\equiv 1$) be its multiplicative identity element. Using the multi-index notation introduced in Section 1.1 we can write any such polynomial as a finite sum (finitely many nonzero coefficients)

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{Z}_+^N} a_\alpha \, x^\alpha \quad (x^\alpha = x_1^{\alpha_1} \cdot \dots \cdot x_N^{\alpha_N}, \; c_\alpha \in \mathbb{K}) \tag{1}$$

The degree of a monomial $x^\alpha$ is $|\alpha| = \alpha_1 + \cdots + \alpha_N$ and if $f \in \mathbb{K}[\mathbf{x}]$ is not the zero polynomial (all $a_\alpha = 0$), its degree is defined to be

$$m = \deg(f) = \max\{|\alpha| : c_\alpha \neq 0\} >$$

When $N > 1$ several different monomials $x^\alpha$ of the same maximum total degree $|\alpha| = m$ may have nonzero coefficients.

Let $f, g \neq 0$ in $\mathbb{K}[\mathbf{x}]$ with degrees $m = \deg(f), n = \deg(g)$. Their product is

$$\begin{aligned}(f \cdot g)(\mathbf{x}) = \Big(\sum_\alpha a_\alpha \, x^\alpha\Big) \cdot \Big(\sum_\beta b_\beta \, x^\beta\Big) &= \sum_{\alpha,\beta} a_\alpha b_\beta x^{\alpha+\beta} \\ &= \sum_\gamma \Big(\sum_{\alpha+\beta=\gamma} a_\alpha b_\beta\Big) \, x^\gamma = \sum_\gamma c_\gamma \, x^\gamma\end{aligned}$$

where $\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_N + \beta_n)$. If $a_\alpha b_\beta x^{\alpha+\beta} \neq 0$ in (1) we must have $|\alpha| \leq m$ and $|\beta| \leq n$, so that $|\alpha + \beta| \leq m + n$; consequently, $\deg(f \cdot g) \leq \deg(f) + \deg(g)$.

Let us split off the monomials of maximum degree, writing

$$f(\mathbf{x}) = \sum_{|\alpha|=m} a_\alpha\, x^\alpha + \cdots,$$

$$g(\mathbf{x}) = \sum_{|\beta|=n} b_\beta\, x^\beta + \cdots,$$

$$(f \cdot g)(\mathbf{x}) = \sum_{|\gamma|=m+n} c_\gamma x^\gamma + \cdots,$$

where $(\cdots)$ are terms of lower degree. To prove the degree formula,

(2) DEGREE FORMULA: $\deg(f \cdot g) = \deg(f) + \deg(g)$

for $f, g \neq 0$ in $\mathbb{K}[\mathbf{x}]$, it suffices to show

(3) There is at least one monomial $x^{\gamma_0}$ having the maximal degree $m + n$ whose coefficient $c_{\gamma_0} = \sum_{\alpha+\beta=\gamma_0} a_\alpha b_\beta$ is nonzero.

This is trivial for $N = 1$, but problematic when $N \geq 2$ because this sum of products can be zero if there is more than one term, even if the individual terms are nonzero. On the other hand, the degree formula (2) follows immediately if we can prove the following:

(4) There exists some monomial $x^\gamma$ of maximal degree $m + n$ for which the sum (3) consists of a single nonzero term.

The key to proving (4) is to introduce a ranking of the monomials $x^\gamma$, $\gamma$ in $\mathbb{Z}^N_+$, more refined than ranking by total degree $\deg(x^\gamma) = |\gamma|$, which cannot distinguish between the various monomials of the same degree. The tool for doing this is *lexicographic* or *dictionary* ordering of the indices in $\mathbb{Z}^N_+$, an idea that has proved useful in many parts of mathematics.

DEFINITION A.1 (Lexicographic order). For multi-indices $\alpha = (\alpha_1, \dots, \alpha_N)$, $\beta = (\beta_1, \dots, \beta_N)$ in $\mathbb{Z}^N_+$, we define the relation $\alpha \succ \beta$ to mean

$$\alpha_i > \beta_i \text{ at the first index } i = 1, 2, \dots, N \text{ at which } \alpha_i \text{ differs from } \beta_i$$

Thus

$$\alpha_1 = \beta_1, \dots, \alpha_{i-1} = \beta_{i-1} \text{ and } \alpha_i > \beta_i \quad \text{(other entries in } \alpha, \beta \text{ are irrelevant).}$$

This is a *linear ordering* of multi-indices: given $\alpha, \beta$ exactly one of the possibilities

$$\alpha \succ \beta \qquad \alpha = \beta \qquad \beta \succ \alpha$$

holds. We write $\alpha \succeq \beta$ when the possibility $\alpha = \beta$ is allowed.

Obviously $\alpha = (0, \dots, 0)$ is the lowest multi-index in lexicographic order, and any finite set of multi-indices has a unique highest element. Note carefully that $\alpha \succ \beta$ does not imply that $|\alpha| \geq |\beta|$. For instance we have

$$\alpha = (1, 0, 0) \succ \beta = (0, 2, 2) \text{ in lexicographic order, but } |\beta| = 4 > |\alpha| = 1\,.$$

Other elementary properties of lexicographic order are easily verified once you understand the definitions.

EXERCISE A.2. For lexicographic order in $\mathbb{Z}_+^N$ verify that we have:

(a) LINEAR ORDERING. For any pair $\alpha, \beta$ we have exactly one of the possibilities $\alpha \succ \beta$, $\alpha = \beta$, $\beta \succ \alpha$.
(b) TRANSITIVITY OF ORDER. If $\alpha \succ \beta$ and $\beta \succ \gamma$ then $\alpha \succ \gamma$.
(c) If $\alpha \succ \alpha'$, then $\alpha + \beta \succ \alpha' + \beta$ for all indices $\beta$.
(d) If $\alpha \succ \alpha'$ and $\beta \succ \beta'$, then $\alpha + \beta \succ \alpha' + \beta'$.

HINT: It might help to make diagrams showing how the various $N$-tuples are related. You will have to do some "casework" in (c).

We now outline a proof of the crucial fact (4), leaving the final details as an exercise for the reader. If $f \neq 0$ with $m = \deg(f)$, so $f = \sum_{|\alpha| \leq m} a_\alpha\, x^\alpha$, there may be several monomials having maximal degree $m$ with $a_\alpha \neq 0$, but just one of them is maximal with respect to lexicographic order, namely,

$$\alpha_0 = \max_{\succ}\{\alpha \,:\, |\alpha| = m \text{ and } a_\alpha \neq 0.\}$$

Likewise, there is a unique index $\beta_0$ for $h = \sum_{|\beta| \leq n} b_\beta x^\beta$ such that $|\beta| = \deg(h) = n$ and

$$\beta_0 = \max_{\succ}\{\beta \,:\, |\beta| = n \text{ and } b_\beta \neq 0\}.$$

The multi-index $\gamma_0 = \alpha_0 + \beta_0$ has $|\gamma_0| = m + n$ and is a likely candidate for the solution to (4); note that $a_{\alpha_0} b_{\beta_0} \neq 0$ *by definition*. We leave the reader to verify a few simple properties of this particular multi-index.

EXERCISE A.3. Explain why $\alpha_0 = \max_{\succ}\{\alpha \,:\, |\alpha| = m \text{ and } a_\alpha \neq 0\}$ might not be the same as $\alpha_0' = \max_{\succ}\{\alpha \,:\, a_\alpha \neq 0\}$. Is there any reason to expect $\alpha_0'$ to have maximal degree $|\alpha_0'| = m$?

EXERCISE A.4. In $\gamma_0 = \alpha_0 + \beta_0$ we have $|\alpha_0| = m$, $|\beta_0| = n$, and $a_{\alpha_0} b_{\beta_0} \neq 0$, by definition. If $\alpha, \beta$ are any indices such that

$$|\alpha + \beta| = |\alpha_0 + \beta_0| = m + n \quad \text{and} \quad a_\alpha b_\beta \neq 0$$

prove that we must have $|\alpha| = |\alpha_0| = m$ and $|\beta| = |\beta_0| = n$.

Defining $\alpha_0, \beta_0, \gamma_0 = \alpha_0 + \beta_0$ as above, we make the following claim:

CLAIM: If $\alpha + \beta = \alpha_0 + \beta_0$ and $a_\alpha b_\beta \neq 0$, then $\alpha = \alpha_0$ and $\beta = \beta_0$. Hence the sum

$$c_{\gamma_0} = \sum_{\alpha + \beta = \gamma_0} a_\alpha b_\beta \tag{5}$$

reduces to the single nonzero term $a_{\gamma_0} b_{\beta_0}$.

EXERCISE A.5. Prove the claim made in (5) using the facts assembled in the preceding discussion.

That will complete the proof of the degree formula.