Chapter 1

Vector Spaces

1.1 Groups and fields

Vector spaces are defined over fields, and the definitions of both fields and vector spaces depend on the notion of a commutative group. Groups and fields are reviewed in more detail in the appendix but, for convenience, we include here their definitions along with some basic examples.

1.1.1 Groups.

DEFINITION: A *group* is a pair $(G,*)$, where G is a set and $*$ is a binary operation $(x,y) \mapsto x*y$, defined for all pairs $(x,y) \in G \times G$, taking values in G, and satisfying the following conditions:

G-1 The operation is associative: For all $x,y,z \in G$,

$$(x*y)*z = x*(y*z).$$

G-2 There exists a unique element $e \in G$, called the *identity element* or the *unit* of G, such that $e*x = x*e = x$ for all $x \in G$.

G-3 For every $x \in G$ there exists a unique element x^{-1}, called *the inverse* of x, such that $x^{-1}*x = x*x^{-1} = e$.

A group $(G,*)$ is *abelian*, or *commutative*, if $x*y = y*x$ for all x and y. The group operation in a commutative group is often written and referred to as *addition*, in which case the identity element is written as 0, and the inverse of x as $-x$.

When the group operation is written as multiplication, the operation symbol $*$ is typically written as a dot (i.e., $x \cdot y$ rather than $x*y$)

and is often omitted altogether. We also simplify the notation by referring to the group, when the binary operation is assumed known, as G, rather than $(G,*)$.

EXAMPLES:

a. $(\mathbb{Z},+)$, the integers with standard addition.

b. $(\mathbb{R}\setminus\{0\},\cdot)$, the nonzero real numbers, with standard multiplication.

c. $\mathsf{S}_n = (\mathsf{S}_n,\cdot)$, the *symmetric group on* $[1,\dots,n]$; n a positive integer. The elements of S_n are all the permutations σ of the set $[1,\dots,n]$, i.e., the set of the bijections (1-1 maps) of $[1,\dots,n]$ onto itself.

The group operation is *composition*: for $\sigma,\tau\in\mathsf{S}_n$ we define $\tau\cdot\sigma$ by: $(\tau\cdot\sigma)(j)=\tau(\sigma(j))$ for all j in $[1,n]$.

The first two examples are commutative; the third is not if $n>2$.

1.1.2 Fields.

DEFINITION: A *field* $(\mathbb{F},+,\cdot)$ is a set $\mathbb{F}$ endowed with two binary operations, *addition*, $(a,b)\mapsto a+b$, and *multiplication*, $(a,b)\mapsto a\cdot b$, (usually written simply as ab), such that:

F-1 $(\mathbb{F},+)$ is a commutative group, whose identity element is denoted by 0.

F-2 $(\mathbb{F}\setminus\{0\},\cdot)$ is a commutative group, whose identity element is denoted by 1. It is *the multiplicative group of* $\mathbb{F}$.

F-3 Addition and multiplication are related by the *distributive law*:

$$a(b+c)=ab+ac.$$

EXAMPLES:

a. $\mathbb{Q}$, the field of rational numbers.

b. $\mathbb{R}$, the field of real numbers.

c. $\mathbb{C}$, the field of complex numbers.

d. $\mathbb{Z}_2$ denotes the field consisting of the two elements 0, 1, with addition and multiplication defined mod 2 (so that $1+1=0$).

More generally, if p is a prime, the set $\mathbb{Z}_p$ of residue classes mod p, with addition and multiplication mod p, is a field. (See exercise **ex1.1.4**.)

The fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are familiar, and are the most commonly used. Less familiar, yet very useful, are the finite fields $\mathbb{Z}_p$, mentioned above. Also important are *extensions* of a given field, see A.5.5.

EXERCISES FOR SECTION 1.1

ex1.1.1 Verify that S_n is not commutative if $n > 2$.

ex1.1.2 Show that if $\mathbb{F}$ is a field, then $0 \cdot a = 0$ for all $a \in \mathbb{F}$, and if $ab = 0$, then $a = 0$ or $b = 0$.

ex1.1.3 Verify that $\mathbb{Z}_3 = \{0,1,2\}$ is a field if addition and multiplication are defined mod 3, i.e., we add and multiply as usual, and if the result is ≥ 3, subtract 3; thus $1+1=2$ but $2 \times 2 = 1$.

Why is $\mathbb{Z}_4$, defined similarly—as the set $\{0,1,2,3\}$ with addition and multiplication defined mod 4—not a field?

ex1.1.4 Let $p > 1$ be a positive integer. Recall that two integers m, n are *congruent mod* p, written $n \equiv m \pmod{p}$, if $n - m$ is divisible by p. This is an equivalence relation (see appendix A.1). For $m \in \mathbb{Z}$, denote by $\tilde{m}$ the coset (equivalence class) of m, that is, the set of all integers n such that $n \equiv m \pmod{p}$.

a. Every integer is congruent mod p to one of the numbers $[0,1,\dots,p-1]$. In other words, there is a 1-1 correspondence between $\mathbb{Z}_p$, the set of cosets $\pmod{p}$, and the integers $[0,1,\dots,p-1]$.

b. As in subsection 1.2.4 above, we define the *quotient ring* $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ (both notations are common) as the space whose elements are the cosets $\pmod{p}$ in $\mathbb{Z}$, and define addition and multiplication by: $\tilde{m} + \tilde{n} = \widetilde{(m+n)}$ and $\tilde{m}\cdot\tilde{n} = \widetilde{m\cdot n}$. Prove that the addition and multiplication so defined are associative, commutative, and satisfy the distributive law.

c. Prove that $\mathbb{Z}_p$, endowed with these operations, is a field if and only if p is prime.

Hint: You may use the following fact: if p is a prime, and both n and m are *not* divisible by p, then nm is not divisible by p. Show that this implies that if $\tilde{n} \neq 0$ in $\mathbb{Z}_p$, then $\{\tilde{n}\tilde{m} : \tilde{m} \in \mathbb{Z}_p\}$ covers all of $\mathbb{Z}_p$.

1.2 Vector spaces

1.2.1 DEFINITION. A *vector space* $\mathscr{V}$ *over a field* $\mathbb{F}$ is an abelian group $(\mathscr{V}, +)$, for which a binary product, $(a, v) \mapsto av$, of $\mathbb{F} \times \mathscr{V}$ into $\mathscr{V}$ is defined, satisfying the following axioms for all $a, b \in \mathbb{F}$ and $u, v \in \mathscr{V}$:

VS 1. $1\,v = v$,

VS 2. $(ab)v = a(bv)$,

VS 3. $(a+b)v = av + bv$ and $a(v+u) = av + au$.

Other familiar properties may be derived from these. For example, for every $v \in \mathscr{V}$, $0v = (1-1)v = v - v = 0$.

The elements of $\mathscr{V}$ are usually referred to as *vectors*; the elements of the underlying field as *scalars*.

Observe that in the equality $(ab)v = a(bv)$ the multiplication (ab) is within $\mathbb{F}$; the others are products of a vector by a scalar. In $(a+b)v = av + bv$, the addition on the left is the addition in $\mathbb{F}$, while that on the right is the addition in $\mathscr{V}$.

Most of the notions and results we discuss are valid for vector spaces over arbitrary fields. When the underlying field does not need to be specified, we denote it by the generic $\mathbb{F}$.

Results that apply to vector spaces over specific fields, or to vector spaces over fields satisfying some additional conditions, will be stated explicitly in terms of the appropriate field or the additional conditions.

If the underlying field is $\mathbb{R}$ or $\mathbb{C}$, then the vector space is called a *real vector space* or a *complex vector space*, respectively.

Vector spaces may also have additional structures: geometric, such as *inner-product*, which we study in Chapter 6; or algebraic, such as multiplication, as discussed in A.5.6.

EXAMPLES: The following sets are all vector spaces over the indicated fields.

a. $\mathbb{F}^n$, the space of all $\mathbb{F}$-valued n-tuples $[a_1,\dots,a_n]$ with addition and multiplication by scalars defined by

$$[a_1,\dots,a_n]+[b_1,\dots,b_n]=[a_1+b_1,\dots,a_n+b_n],$$
$$c[a_1,\dots,a_n]=[ca_1,\dots,ca_n].$$

We may write the n-tuples as rows, as we did here, or as columns. If we want to specify that the vectors are written as columns or as rows, we write $\mathbb{F}^n_{\mathbf{c}}$ or $\mathbb{F}^n_{\mathbf{r}}$, respectively.

b. $\mathscr{M}(n,m;\mathbb{F})$, the space of all $\mathbb{F}$-valued $n\times m$ matrices; that is, arrays

$$A=\begin{bmatrix} a_{11} & \dots & a_{1m}\\ a_{21} & \dots & a_{2m}\\ \vdots & \dots & \vdots\\ a_{n1} & \dots & a_{nm}\end{bmatrix}$$

with entries from $\mathbb{F}$. We sometimes write $A=[a_{ij}]$ when the dimensions of the matrix are assumed known, to save space.

The addition and multiplication by scalars are again done entry by entry. As a vector space, $\mathscr{M}(n,m;\mathbb{F})$ is virtually identical with $\mathbb{F}^{mn}$.

We write $\mathscr{M}(n;\mathbb{F})$ instead of $\mathscr{M}(n,n;\mathbb{F})$, and if the underlying field is either known implicitly, or assumed explicitly, then we often write simply $\mathscr{M}(n,m)$ or $\mathscr{M}(n)$, as the case may be.

c. $\mathbb{F}[x]$, the space[1] of all polynomials $\sum a_n x^n$ with coefficients from $\mathbb{F}$. Addition and multiplication by scalars are defined formally either as the standard addition and multiplication of functions (of the variable x), or by adding (and multiplying by scalars) the corresponding coefficients. The two ways define the same operations.

More generally, the set $\mathbb{F}[x_1,\dots,x_k]$ of all polynomials in k variables over $\mathbb{F}$ is a vector space (in fact—an algebra).

[1] $\mathbb{F}[x]$ is an algebra over $\mathbb{F}$, i.e., a vector space with the additional operation of multiplication. See A.5.6.

d. Let X be a finite set. $C(X)$ denotes the set of all complex-valued functions on X, with the standard addition of functions, and multiplication of functions by scalars.

e. The set $C_{\mathbb{R}}([0,1])$ of all continuous real-valued functions f on $[0,1]$, and the set $C([0,1])$ of all continuous complex-valued functions f on $[0,1]$, with the standard operations of addition of functions and of multiplication of functions by scalars. $C_{\mathbb{R}}([0,1])$ is a real vector space. $C([0,1])$ is naturally a complex vector space, but becomes a real vector space if we limit the allowable scalars to real numbers only.

f. The set $C^{\infty}([-1,1])$ of all infinitely differentiable real-valued functions f on $[-1,1]$, with the standard operations on functions is a real vector space.

g. The set $\mathscr{T}_N$ of 2π-periodic trigonometric polynomials of degree bounded by N; that is, the functions of the form $\sum_{|n|\le N} a_n e^{inx}$. The underlying field may be $\mathbb{C}$ or $\mathbb{R}$, and the operations are the standard addition of functions and multiplication of functions by scalars.

Similarly, the space $\mathscr{T}_{N,M}$ of trigonometric polynomials in two variables of the form $\sum_{|n|\le N,\,|m|\le M} e^{i(nx+my)}$, with the same operations and the same underlying field(s).

h. The set of complex-valued functions f on $\mathbb{R}$ that satisfy the differential equation

$$3f'''(x) - \sin x f''(x) + 2f(x) = 0,$$

with the standard operations on functions. If we are interested in real-valued functions only, the underlying field is naturally $\mathbb{R}$. If we allow complex-valued functions we may choose either $\mathbb{C}$ or $\mathbb{R}$ as the underlying field.

1.2.2 Isomorphism. The expression "virtually identical" in the comparison above of $\mathscr{M}(n,m;\mathbb{F})$ with $\mathbb{F}^{mn}$ is not a proper mathematical term. The proper term here is *isomorphic*.

Recall (or see section A.2 in the appendix) that a map $\varphi\colon X \mapsto Y$ from a set X to a set Y is *bijective* if for every $y \in Y$ there is precisely one $x \in X$ such that $y = \varphi(x)$. Bijective maps are *invertible*—the inverse map is defined by:

$$\varphi^{-1}(y) = x \quad \text{if} \quad y = \varphi(x). \tag{1.2.1}$$

Let $\mathscr{V}$ and $\mathscr{W}$ be vector spaces (over the same field).

DEFINITION: A map $\varphi\colon \mathscr{V} \mapsto \mathscr{W}$ is *linear* if for all scalars a, b and vectors $v_1, v_2 \in \mathscr{V}$

$$\varphi(av_1 + bv_2) = a\varphi(v_1) + b\varphi(v_2). \tag{1.2.2}$$

A map $\varphi\colon \mathscr{V} \mapsto \mathscr{W}$ is *an isomorphism* if it is *both bijective and linear.* An *automorphism* is an isomorphism of a vector space $\mathscr{V}$ onto itself.

If φ is an isomorphism of $\mathscr{V}$ onto $\mathscr{W}$, then φ^{-1} is an isomorphism of $\mathscr{W}$ onto $\mathscr{V}$. This can be seen as follows: by (1.2.2),

$$\begin{aligned}\varphi^{-1}(a_1\varphi v_1 + a_2\varphi v_2) &= \varphi^{-1}(\varphi(a_1 v_1 + a_2 v_2)) = a_1 v_1 + a_2 v_2 \\ &= a_1\varphi^{-1}(\varphi v_1) + a_2\varphi^{-1}(\varphi v_2),\end{aligned} \tag{1.2.3}$$

and, as φ is surjective, every vector in $\mathscr{W}$ is equal to φv for some $v \in \mathscr{V}$.

We say that *$\mathscr{V}$ and $\mathscr{W}$ are isomorphic* if there is an isomorphism φ of $\mathscr{V}$ onto $\mathscr{W}$, and the fact that the inverse φ^{-1} is also an isomorphism guarantees that the relation of being isomorphic is symmetric.

The identity map $\varphi(x) = x$ shows that the relation is reflexive and, since the composition of isomorphisms is an isomorphism, see exercise **ex1.2.2**, the relation is also transitive. In other words, the relations of being isomorphic is *an equivalence relation* (see A.1.2).

1.2.3 Subspaces. A (vector) *subspace* of a vector space $\mathscr{V}$ is a subset that is closed under the operations of addition and multiplication by scalars inherited from $\mathscr{V}$. In other words, $\mathscr{W} \subset \mathscr{V}$ is a subspace if for all scalars a_j and vectors $w_j \in \mathscr{W}$, $j = 1, 2$, the vectors $a_1 w_1 + a_2 w_2$ are in $\mathscr{W}$.

EXAMPLES:

a. *Solution-set of a system of homogeneous linear equations.*

Here $\mathscr{V} = \mathbb{F}^n$. Given the scalars a_{ij}, $1 \le i \le k$, $1 \le j \le n$, we consider the solution-set of the system of k homogeneous linear equations

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = 1, \dots, k. \tag{1.2.4}$$

This is the set of all n-tuples, $[x_1, \dots, x_n] \in \mathbb{F}^n$ (thought of as vectors), for which all k equations are satisfied. If both $[x_1, \dots, x_n]$ and $[y_1, \dots, y_n]$ are solutions of (1.2.4), and a and b are scalars, then for each i,

$$\sum_{j=1}^{n} a_{ij}(ax_j + by_j) = a\sum_{j=1}^{n} a_{ij}x_j + b\sum_{j=1}^{n} a_{ij}y_j = 0.$$

It follows that the solution-set of (1.2.4) is a subspace of $\mathbb{F}^n$.

b. In $\mathbb{F}[x]$, the space $\mathbb{F}_N[x]$ of polynomials $\sum_0^N a_n x^n$ of degree $\le N$. While $\mathbb{F}[x]$ is an algebra, $\mathbb{F}_N[x]$ is not an algebra; why?

c. In the space $C^\infty(\mathbb{R})$ of all infinitely differentiable, real-valued functions f on $\mathbb{R}$ with the standard operations, the set of functions f that satisfy the differential equation

$$f'''(x) - 5f''(x) + 2f'(x) - f(x) = 0.$$

If we consider complex-valued solutions, then the field of scalars can be $\mathbb{R}$ or $\mathbb{C}$.

d. Subspaces of $\mathscr{M}(n)$:

The set of *diagonal matrices*—the $n \times n$ matrices with zero entries off the main diagonal, i.e., $a_{ij} = 0$ for $i \ne j$.

The set of *symmetric matrices*—the $n \times n$ matrices whose entries satisfy $a_{ij} = a_{ji}$.

The set of *skew-symmetric matrices*—the $n \times n$ matrices whose entries satisfy $a_{ij} = -a_{ji}$.

The set of *lower triangular matrices*—the $n \times n$ matrices with zero entries above the main diagonal, i.e., $a_{ij} = 0$ for $i < j$.

Similarly, the set of *upper triangular matrices*, i.e., those for which $a_{ij} = 0$ for $i > j$.

Remark: In view of the isomorphism between $\mathscr{M}(n)$ and $\mathbb{F}^{n^2}$, these subspaces can be viewed as special cases of example ***a.***

e. *Intersection of subspaces:* If $\mathscr{W}_j$ are subspaces of a space $\mathscr{V}$, $j \in J$, (the index set J can be finite or infinite), then $\bigcap \mathscr{W}_j$ is a subspace of $\mathscr{V}$.

f. *The sum of subspaces:* If $\mathscr{W}_j$, $j \in J$, are subspaces of $\mathscr{V}$, their sum is the set:

$$\sum \mathscr{W}_j = \bigcup_{J_1 \subset J} \{v : \ v = \sum_{j \in J_1} v_j, \ \ v_j \in \mathscr{W}_j\},$$

where the union extends over the collection of *finite* subsets of J.

Don't confuse the *sum of subspaces* with *the union of subspaces*, which is seldom a subspace, see exercises **ex1.2.5** and **ex1.2.4**.

g. *The span of a subset:* Given $E \subset \mathscr{V}$, a *linear combination* of elements of E is *a finite sum* of the form $\sum a_j v_j$, $\ a_j \in \mathbb{F}$, $v_j \in E$. The set of all the linear combinations of elements of E is a subspace of $\mathscr{V}$, called *the span* of E and denoted by $\mathrm{span}[E]$.

1.2.4 Quotient spaces. A subspace $\mathscr{W}$ of a vector space $\mathscr{V}$ defines an equivalence relation (see the appendix, section A.1) in $\mathscr{V}$:

$$x \equiv y \pmod{\mathscr{W}} \qquad \text{if} \quad x - y \in \mathscr{W}. \tag{1.2.5}$$

In order to establish that this is indeed an equivalence relation, we need to verify that it is reflexive, symmetric, and transitive:

a. *reflexive:* $x \equiv x$, since $x - x = 0 \in \mathscr{W}$,

b. *symmetric:* $x \equiv y \iff y \equiv x$, since $x - y \in \mathscr{W}$ if and only if $y - x = -(x - y) \in \mathscr{W}$,

c. *transitive:* If both $x \equiv y$ and $y \equiv z$, then $x \equiv z$. This follows from: if both $(x-y)$ and $(y-z)$ are in $\mathscr{W}$, then so is $x-z = (x-y)+(y-z)$.

This equivalence relation partitions $\mathscr{V}$ into equivalence classes, called the *cosets* of $\mathscr{W}$ in $\mathscr{V}$. For $x \in \mathscr{V}$, the coset of $\mathscr{W}$ that contains x is the set $\tilde{x} = x + \mathscr{W} = \{v = x + w : w \in \mathscr{W}\}$—the "translate" of $\mathscr{W}$ by x.

We define the *quotient space* $\mathscr{V}/\mathscr{W}$ to be the space whose elements are the cosets of $\mathscr{W}$ in $\mathscr{V}$, and the operations of addition and multiplication by scalars are defined as follows. If $\tilde{x} = x + \mathscr{W}$ and $\tilde{y} = y + \mathscr{W}$ are cosets, and $a \in \mathbb{F}$, then

$$\tilde{x} + \tilde{y} = x + y + \mathscr{W} = \widetilde{x+y} \quad \text{and} \quad a\tilde{x} = \widetilde{ax}. \tag{1.2.6}$$

The definition needs justification. We defined the sum of two cosets by taking a representative element from each, and taking the coset that contains their sum as the sum of the cosets. We need to show that the result is well defined, i.e., that it does not depend on the choice of the representatives in the cosets. In other words, we need to verify that if $x \equiv x_1 \pmod{\mathscr{W}}$ and $y \equiv y_1 \pmod{\mathscr{W}}$, then $x + y \equiv x_1 + y_1 \pmod{\mathscr{W}}$. Now, if $x = x_1 + w$, $y = y_1 + w'$ with $w, w' \in \mathscr{W}$, then $x + y = x_1 + w + y_1 + w' = x_1 + y_1 + w + w'$, and since $w + w' \in \mathscr{W}$, we have $x + y \equiv x_1 + y_1 \pmod{\mathscr{W}}$.

The definition of $a\tilde{x}$ is justified similarly: if $x \equiv x_1 \pmod{\mathscr{W}}$, then $ax - ax_1 = a(x - x_1)$, and since $\mathscr{W}$ is a *subspace*, it is closed under multiplication by scalars, $a(x - x_1) \in \mathscr{W}$, and $ax \equiv ax_1 \pmod{\mathscr{W}}$.

1.2.5 Direct sums. If $\mathscr{V}_1, \dots, \mathscr{V}_k$ are vector spaces over $\mathbb{F}$, *the (formal) direct sum*

$$\bigoplus_1^k \mathscr{V}_j = \mathscr{V}_1 \oplus \cdots \oplus \mathscr{V}_k$$

is the set $\{(v_1,\dots,v_k) : v_j \in \mathscr{V}_j\}$ in which addition and multiplication by scalars are defined by:

$$(v_1,\dots,v_k)+(u_1,\dots,u_k) = (v_1+u_1,\dots,v_k+u_k),$$
$$a(v_1,\dots,v_k) = (av_1,\dots,av_k).$$

The direct sum of vector spaces is clearly a vector space.

In the case that $\mathscr{V}_1,\dots,\mathscr{V}_k$ are all subspaces of the same vector space $\mathscr{V}$, we defined their *sum*, $\sum\mathscr{V}_j = \{v : v = \sum_{j=1}^k v_j,\ v_j \in \mathscr{V}_j\}$, in example ***f*** of 1.2.3.

The "natural" map of $\mathscr{V}_1 \oplus\cdots\oplus \mathscr{V}_k$ into $\mathscr{V}_1+\cdots+\mathscr{V}_k$, defined by

$$\Phi((v_1,\dots,v_k)) = \sum_1^k v_j, \tag{1.2.7}$$

is clearly linear and surjective. It is an isomorphism when the subspaces are *independent*.

DEFINITION: The subspaces $\mathscr{V}_j$, $j = 1,\dots,k$, of a vector space $\mathscr{V}$ are *independent* if $\sum v_j = 0$ with $v_j \in \mathscr{V}_j$ implies that $v_j = 0$ for all j.[2]

Proposition. *Let $\mathscr{V}_j$ be subspaces of $\mathscr{V}$. The map Φ defined by* (1.2.7) *is an isomorphism if and only if the subspaces are independent.*

PROOF: Φ is clearly linear and surjective. It is injective if and only if every vector in the range has a unique preimage, that is, if

$$v'_j, v''_j \in \mathscr{V}_j \qquad \text{and} \qquad v''_1+\cdots+v''_k = v'_1+\cdots+v'_k \tag{1.2.8}$$

implies that $v''_j = v'_j$ for every j. Subtracting and writing $v_j = v''_j - v'_j$, (1.2.8) is equivalent to: $\sum v_j = 0$ with $v_j \in \mathscr{V}_j$. The subspaces are independent if and only if this implies that $v_j = 0$ for all j. ◀

In view of the proposition, we refer to the sum $\sum\mathscr{W}_j$ of *independent* subspaces of a vector space as their *direct sum*, and write $\bigoplus\mathscr{W}_j$ instead of $\sum\mathscr{W}_j$.

If $\mathscr{V} = \mathscr{U}\oplus\mathscr{W}$, we refer to $\mathscr{U}$ as *a complement* of $\mathscr{W}$ in $\mathscr{V}$, and vice versa.

[2] Properly speaking: the set $\{\mathscr{V}_j\}$ is independent.

1.2.6 Tensor products. Let $\mathscr{V}$ and $\mathscr{U}$ be vector spaces over $\mathbb{F}$. Let $\widetilde{\mathscr{V}\otimes\mathscr{U}}$ be the set of all the (finite) formal sums $\sum a_j v_j \otimes u_j$, where $a_j \in \mathbb{F}$, $v_j \in \mathscr{V}$ and $u_j \in \mathscr{U}$. We define addition formally by

$$\sum_{j\in J_1} a_j v_j \otimes u_j + \sum_{j\in J_2} a_j v_j \otimes u_j = \sum_{j\in J_1\cup J_2} a_j v_j \otimes u_j,$$

and we define multiplication by scalars by

$$a\sum a_j v_j \otimes u_j = \sum (aa_j)\, v_j \otimes u_j.$$

With these definitions, $\widetilde{\mathscr{V}\otimes\mathscr{U}}$ is a vector space over $\mathbb{F}$.

The *tensor product* $\mathscr{V}\otimes\mathscr{U}$ is, by definition, the quotient of $\widetilde{\mathscr{V}\otimes\mathscr{U}}$ by the subspace $[\widetilde{\mathscr{V}\otimes\mathscr{U}}]_0$ spanned by the elements of the form

$$\text{(1.2.9)}\qquad \begin{aligned} &\textbf{a}. \quad (v_1+v_2)\otimes u - (v_1\otimes u + v_2\otimes u),\\ &\textbf{b}. \quad v\otimes(u_1+u_2) - (v\otimes u_1 + v\otimes u_2),\\ &\textbf{c}. \quad a(v\otimes u) - (av)\otimes u, \quad (av)\otimes u - v\otimes(au), \end{aligned}$$

for all $v, v_j \in \mathscr{V}$, $u, u_j \in \mathscr{U}$, and $a \in \mathbb{F}$.

In other words, $\mathscr{V}\otimes\mathscr{U}$ is the space of formal sums $\sum a_j v_j \otimes u_j$ *modulo the equivalence relation* generated by:

$$\text{(1.2.10)}\qquad \begin{aligned} &\textbf{a}. \quad (v_1+v_2)\otimes u \equiv v_1\otimes u + v_2\otimes u,\\ &\textbf{b}. \quad v\otimes(u_1+u_2) \equiv v\otimes u_1 + v\otimes u_2,\\ &\textbf{c}. \quad a(v\otimes u) \equiv (av)\otimes u \equiv v\otimes(au). \end{aligned}$$

EXAMPLE: If $\mathscr{V} = \mathbb{F}[x]$ and $\mathscr{U} = \mathbb{F}[y]$, and we define the map Φ of $\widetilde{\mathscr{V}\otimes\mathscr{U}}$ into $\mathbb{F}[x,y]$ by:

$$\text{(1.2.11)}\qquad \Phi\colon \sum_j a_j p_j(x)\otimes q_j(y) \mapsto \sum_j a_j p_j(x) q_j(y) \in \mathbb{F}[x,y],$$

then all the elements of $[\widetilde{\mathscr{V}\otimes\mathscr{U}}]_0$ are mapped to zero, so that all the elements in an equivalence class modulo $[\widetilde{\mathscr{V}\otimes\mathscr{U}}]_0$ are mapped to the same polynomial. For example, every formal sum in $\widetilde{\mathscr{V}\otimes\mathscr{U}}$ that is equivalent to $p(x)\otimes q(y)$ is mapped to $p(x)q(y)$. It follows that Φ induces a map of the quotient space $\mathscr{V}\otimes\mathscr{U}$ onto $\mathbb{F}[x,y]$

$$\overline{\Phi}\colon \sum_j p_j(x)\otimes q_j(y)\mapsto \sum_j p_j(x)q_j(y). \tag{1.2.12}$$

$\overline{\Phi}$ is an isomorphism, and the spaces $\mathscr{V}\otimes\mathscr{U}$ and $\mathbb{F}[x,y]$ are isomorphic.

EXERCISES FOR SECTION 1.2

ex1.2.1 Verify that $\mathbb{R}$ is a vector space over $\mathbb{Q}$, and that $\mathbb{C}$ is a vector space over either $\mathbb{Q}$ or $\mathbb{R}$.

ex1.2.2 Let $\mathscr{V}_j$, $j=1,2,3$, be vector spaces over the same field $\mathbb{F}$. Let $\varphi_1:\mathscr{V}_1\mapsto\mathscr{V}_2$ and $\varphi_2:\mathscr{V}_2\mapsto\mathscr{V}_3$ be isomorphisms. Prove that $\varphi_2\varphi_1$ is an isomorphism of $\mathscr{V}_1$ onto $\mathscr{V}_3$. Conclude that *isomorphism* is an equivalence relation for vector spaces (defined over the same field $\mathbb{F}$).

ex1.2.3 Verify that the intersection of subspaces is a subspace.

ex1.2.4 Let $\mathscr{U}$ and $\mathscr{W}$ be subspaces of a vector space $\mathscr{V}$, and neither of them contains the other. Show that $\mathscr{U}\cup\mathscr{W}$ is *not* a subspace.
Hint: Take $u\in\mathscr{U}\setminus\mathscr{W}$, $w\in\mathscr{W}\setminus\mathscr{U}$ and consider $u+w$.

ex1.2.5 Verify that the sum of subspaces is a subspace, and prove that

$$\sum\mathscr{W}_j=\mathsf{span}[\bigcup\mathscr{W}_j].$$

ex1.2.6 Check that for every $E\subset\mathscr{V}$, $\mathsf{span}[E]$ is a subspace of $\mathscr{V}$, and is contained in every subspace that contains E.

ex1.2.7 If $\mathscr{V}_1$ is a subspace of $\mathscr{V}$ and φ is an isomorphism of $\mathscr{V}$ onto $\mathscr{W}$, then $\varphi\mathscr{V}_1$ is a subspace of $\mathscr{W}$.

ex1.2.8 Describe all the complements in $\mathbb{R}^2$ of the subspace $X=\{(x,0):x\in\mathbb{R}\}$.

ex1.2.9 Prove that two subspaces $\mathscr{V}_1$ and $\mathscr{V}_2$ in a vector space are independent if $\mathscr{V}_1\cap\mathscr{V}_2=\{0\}$.

ex1.2.10 Prove that the subspaces $\mathscr{W}_j\subset\mathscr{V}$, $j=1,\dots,N$ are independent if and only if $\mathscr{W}_j\cap\sum_{l\neq j}\mathscr{W}_l=\{0\}$ for all j.

ex1.2.11 Let $\mathscr{U}$ and $\mathscr{W}$ be subspaces of a vector space $\mathscr{V}$; then $\mathscr{U}$ is a subspace of $\mathscr{U}+\mathscr{W}$, and $\mathscr{U}\cap\mathscr{W}$ is a subspace of $\mathscr{W}$. Prove that the quotient spaces

$$(\mathscr{U}+\mathscr{W})/\mathscr{U} \quad \text{and} \quad \mathscr{W}/(\mathscr{U}\cap\mathscr{W}) \tag{1.2.13}$$

are isomorphic.

Hint: A coset of $\mathscr{U}$ in $(\mathscr{U}+\mathscr{W})$ has the form $w+\mathscr{U}$, and $w_1+\mathscr{U}=w_2+\mathscr{U}$ if and only if $w_1-w_2\in\mathscr{U}$.

$*$**ex1.2.12** Assuming that $\mathbb{F}$ is infinite,[3] and $\mathscr{V}$ is a vector space over $\mathbb{F}$, show that the union of a finite number of subspaces of $\mathscr{V}$, none of which contains all the others, is not a subspace.

Hint: Let $\mathscr{V}_j$, $j=1,\dots,k$ be the subspaces in question. Show that there is no loss in generality in assuming that their union spans $\mathscr{V}$. Now you need to show that $\bigcup\mathscr{V}_j$ is not all of $\mathscr{V}$. Show that there is no loss of generality in assuming that $\mathscr{V}_1$ is not contained in the union of the others. Now take $v_1\in\mathscr{V}_1\setminus\bigcup_{j\neq 1}\mathscr{V}_j$, and $w\notin\mathscr{V}_1$; show that $av_1+w\in\bigcup\mathscr{V}_j$, $a\in\mathbb{F}$, for no more than k values of a.

1.3 Linear dependence, bases, and dimension

Let $\mathscr{V}$ be a vector space. A *linear combination of vectors* $v_1,\dots,v_k$ is a sum of the form $v=\sum a_j v_j$ with scalar coefficients a_j.

A linear combination is *nontrivial* if at least one of the coefficients is not zero; otherwise it is *trivial.*

1.3.1 Recall that if A is a subset of a vector space $\mathscr{V}$, then $\mathrm{span}[A]$, the set of all linear combinations of elements in A, is a subspace of $\mathscr{V}$, (see example ***f*** of 1.2.3).

DEFINITION: The set $A\subset\mathscr{V}$ is a *spanning set* if $\mathrm{span}[A]=\mathscr{V}$.

1.3.2 DEFINITION. The set $A\subset\mathscr{V}$ is *linearly independent* if for every sequence $\{v_1,\dots,v_l\}$ of distinct vectors in A, the only vanishing linear combination of the v_j's is trivial; that is, if $\sum a_j v_j=0$, then $a_j=0$ for all j.[4]

[3] If $\mathbb{F}$ is finite, then $\mathbb{F}^n=\cup_{v\in\mathbb{F}^n}\mathrm{span}[v]$ is a finite union of subspaces.

[4] Independence is a property of the set A; however, we often say, by abuse of language, that *the vectors (in A) are independent.*

If the set A is finite, we enumerate its elements as $v_1, \dots, v_m$ and write the elements in its span as $\sum a_j v_j$. By definition, independence of A means that the representation of $v = 0$ as a linear combination of elements from A is unique. Notice, however, that this implies that the representation of *every* vector in $\mathsf{span}[A]$ is unique. In fact, the equality $\sum_1^l a_j v_j = \sum_1^l b_j v_j$ implies $\sum_1^l (a_j - b_j) v_j = 0$ so that $a_j = b_j$ for all j.

A vector v is *linearly dependent on* a set A if it can be represented as a linear combination of vectors from A, that is, if $v \in \mathsf{span}[A]$.

1.3.3 A *minimal spanning set* is a spanning set such that no proper subset thereof is spanning.

A *maximal independent set* is an independent set such that no set that contains it properly is independent.

Lemma.

a. *A minimal spanning set is independent.*

b. *A maximal independent set is spanning.*

PROOF: **a.** Let A be a minimal spanning set. If $\sum a_j v_j = 0$, with distinct $v_j \in A$, and for some k, $a_k \neq 0$, then $v_k = -a_k^{-1} \sum_{j \neq k} a_j v_j$. This permits the substitution of v_k in any linear combination by the combination of the other v_j's, and shows that v_k is redundant: the span of $\{v_j : j \neq k\}$ is the same as the original span, contradicting the minimality assumption.

b. If B is independent and u is not in $\mathsf{span}[B]$, then the union $\{u\} \cup B$ is independent: otherwise there would exist $\{v_1, \dots, v_l\} \subset B$ and coefficients d and c_j, not all zero, such that $du + \sum c_j v_j = 0$.

If $d \neq 0$, then $u = -d^{-1} \sum c_j v_j$ and u would be in $\mathsf{span}[v_1, \dots, v_l] \subset \mathsf{span}[B]$, contradicting the assumption $u \notin \mathsf{span}[B]$.

If $d = 0$, we have $\sum c_j v_j = 0$ with some nonvanishing coefficients, contradicting the assumption that B is independent. It follows that if B is maximal independent, then $u \in \mathsf{span}[B]$ for every $u \in \mathscr{V}$; in other words: B is spanning. ◀

DEFINITION: A *basis for* $\mathscr{V}$ is a set $B \subset \mathscr{V}$ which is both spanning and independent. A vector space is *finite-dimensional* if it has a finite basis.

Thus, $\{v_1,\dots,v_n\}$ is a basis for $\mathscr{V}$ if and only if every $v\in\mathscr{V}$ has a unique representation as a linear combination of $\{v_1,\dots,v_n\}$. This representation, $v=\sum a_jv_j$, is the *expansion of v relative to the basis* $\{v_1,\dots,v_n\}$.

By the lemma, a minimal spanning set is a basis, and a maximal independent set is a basis.

1.3.4 Proposition. *If $\mathscr{V}$ is finite-dimensional, then:*

a. *Every finite spanning set can be trimmed to a basis.*

b. *Every independent set can be expanded to a basis.*

PROOF: **a.** Let $\{v_j\}_{j=1}^N$ be a spanning set for $\mathscr{V}$. Call a vector v_l *inessential* if it is linearly dependent on $\{v_j\}_{j=1}^{l-1}$, and *essential* otherwise. *This classification clearly depends on the order in which the vectors are enumerated.* Observe that an inessential v_l is linearly dependent on the *essential* vectors preceding it.

Remove the inessential vectors. Since every v_j is either essential or linearly dependent on the preceding essential vectors, the essential vectors span $\mathscr{V}$ and are independent, hence they form a basis.

b. Let $\{u_j\}_{j=1}^k$ be independent, and let $\{e_j\}_{j=1}^n$ be a basis for $\mathscr{V}$. Write $w_j=u_j$ for $j=1,\dots,k$, and $w_{k+j}=e_j$ for $j=1,\dots,n$. The sequence $\{w_j\}$ contains the basis $\{e_j\}$ and is therefore spanning. Now remove, as in part **a**, the inessential vectors to obtain a basis, and observe that the first k vectors, namely $\{u_j\}_{j=1}^k$, are all essential, and hence form part of the basis. ◀

Remarks: The statement of **a** and its proof are valid for infinite spanning sequences as well. See also 1.3.8.

The argument of **b** is refined somewhat in the following subsection.

EXAMPLES:

a. *The standard basis* for $\mathbb{F}^n$: we write e_j for the vector whose j'th entry is equal to 1 and all the other entries are zero. Then $\{e_1,\dots,e_n\}$

is a basis for $\mathbb{F}^n$, and the unique representation of $v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ in terms of this basis is $v = \sum a_j e_j$.

b. *The standard basis* for $\mathscr{M}(n,m)$: let E_{ij} denote the $n \times m$ matrix whose ij'th entry is 1 and all the other entries are zero. Then $\{E_{ij}\}$ is a basis for $\mathscr{M}(n,m)$, and the expansion of $A = [a_{ij}] \in \mathscr{M}(n,m)$ is $A = \sum a_{ij} E_{ij}$.

c. The space $\mathbb{F}[x]$ does not have a finite basis. The infinite sequence $\{x^n\}_{n=0}^{\infty}$ is both linearly independent and spanning, that is, a basis. As we see in the following subsection, the existence of an infinite basis, even of an infinite independent set, precludes a finite basis and the space is *infinite-dimensional*.

Notice, however, that the subspace $\mathbb{F}_N[x]$, consisting of all the polynomials of degree at most N, is finite-dimensional since the set of $N+1$ vectors, $\{x^n\}_{n=0}^{N}$, is a basis.

1.3.5 Steinitz' lemma and the definition of dimension.

Lemma (Steinitz). *If* $\operatorname{span}[v_1, \dots, v_n] = \mathscr{V}$ *and* $\{u_1, \dots, u_m\}$ *is linearly independent in* $\mathscr{V}$*, then the vectors* v_j *can be (re)ordered so that, for every* $k = 1, \dots, m$*, the sequence* $\{u_1, \dots, u_k, v_{k+1}, \dots, v_n\}$ *spans* $\mathscr{V}$*. In particular,* $m \le n$*.*

PROOF: Write $u_1 = \sum a_j v_j$; this is possible since $\operatorname{span}[v_1, \dots, v_n] = \mathscr{V}$. Reorder the v_j's, if necessary, to guarantee that $a_1 \neq 0$.

Now $v_1 = a_1^{-1}(u_1 - \sum_{j=2}^{n} a_j v_j)$, so that $\operatorname{span}[u_1, v_2, \dots, v_n]$ contains every v_j and hence is equal to $\mathscr{V}$.

Continue inductively: assume that, having reordered v_j's, if necessary, we have $\{u_1, \dots, u_k, v_{k+1}, \dots, v_n\}$ spans $\mathscr{V}$.

If $k = m$ we are done. If $k < m$ write

$$u_{k+1} = \sum_{j=1}^{k} a_j u_j + \sum_{j=k+1}^{n} b_j v_j. \tag{1.3.1}$$

Since $\{u_1,\dots,u_m\}$ is linearly independent, at least one of the coefficients b_j is not zero. Reordering, if necessary, the remaining v_j's, we may assume that $b_{k+1}\neq 0$. Now rewrite (1.3.1):

$$v_{k+1}=-b_{k+1}^{-1}\Big(\sum_{j=1}^{k}a_ju_j-u_{k+1}+\sum_{j=k+2}^{n}b_jv_j\Big), \tag{1.3.2}$$

so that $v_{k+1}\in\mathrm{span}[u_1,\dots,u_{k+1},v_{k+2},\dots,v_n]$, and, once again, the span is $\mathscr{V}$. Repeating the step m times proves the lemma. ◀

Theorem. *If $\{v_1,\dots,v_n\}$ and $\{u_1,\dots,u_m\}$ are both bases, then $m=n$.*

PROOF: Since $\{v_1,\dots,v_n\}$ is spanning and $\{u_1,\dots,u_m\}$ is independent, we have $m\le n$. Reversing the roles we have $n\le m$. ◀

Part of Steinitz' lemma repeats part **b** of proposition 1.3.4: in a finite-dimensional vector space, every independent set can be expanded to a basis by adding, if necessary, elements from any given spanning set. The additional information here, that any spanning set has at least as many elements as any independent set, is the basis for the current theorem, and enables the definition of *dimension*.

Recall that a vector space $\mathscr{V}$ is *finite-dimensional* if it has a finite basis.

DEFINITION: The *dimension* of a finite-dimensional vector space $\mathscr{V}$, denoted $\dim\mathscr{V}$, is the number of elements in any basis for $\mathscr{V}$. The definition is unambiguous, since all bases have the same cardinality.

As you are asked to check in exercise **ex1.3.6** below, a subspace $\mathscr{W}$ of a finite-dimensional space $\mathscr{V}$ is finite-dimensional and, unless $\mathscr{W}=\mathscr{V}$, the dimension $\dim\mathscr{W}$ of $\mathscr{W}$ is strictly lower than $\dim\mathscr{V}$.

The *codimension* of a subspace $\mathscr{W}$ in a finite-dimensional space $\mathscr{V}$ is, by definition, $\dim\mathscr{V}-\dim\mathscr{W}$.

1.3.6 Assume that $\mathscr{V}$ and $\mathscr{W}$ are finite-dimensional vector spaces over $\mathbb{F}$, and $\dim\mathscr{V}=\dim\mathscr{W}=d$. Let $\{v_j\}_{j=1}^{d}$ and $\{w_j\}_{j=1}^{d}$ be bases for $\mathscr{V}$ and $\mathscr{W}$ respectively.

Every element $v \in \mathscr{V}$ has a unique representation as a linear combination of the basis elements, so the map $\varphi \colon \mathscr{V} \mapsto \mathscr{W}$ defined by:

$$\varphi\Big(\sum_1^d a_j v_j\Big) = \sum_1^d a_j w_j, \tag{1.3.3}$$

is unambiguous. Since $\{w_j\}$ is a basis of $\mathscr{W}$, the map φ is a bijection onto $\mathscr{W}$ and it clearly satisfies condition (1.2.2); in other words, *it is an isomorphism*. Conversely, if $\mathscr{V}$ is finite-dimensional and φ is an isomorphism of $\mathscr{V}$ onto a vector space $\mathscr{W}$, then the φ-image of a basis of $\mathscr{V}$ is a basis of $\mathscr{W}$. This proves the following theorem.

Theorem. *Two finite-dimensional vector spaces over the same field are isomorphic if and only if they have the same dimension.*

Remark: The definition (1.3.3) of the map φ can be viewed as a two-step process. The first step assigns to each basis element v_j its image w_j; the second "completes the definition by linearity": if φ is to be linear and $\varphi(v_j) = w_j$, then linearity forces, and in turn is guaranteed by, (1.3.3). The fact that $\{v_j\}$ is a basis guarantees that φ is well defined, and this independently of whether $\{w_j\}$ is a basis or not. Adding the assumption that $\{w_j\}$ is a basis for $\mathscr{W}$ guarantees that φ is bijective, and hence an isomorphism.

1.3.7 The following observation is sometimes useful.

Proposition. *Let $\mathscr{U}$ and $\mathscr{W}$ be subspaces of an n-dimensional space $\mathscr{V}$, and assume that* $\dim \mathscr{U} + \dim \mathscr{W} > n$. *Then* $\mathscr{U} \cap \mathscr{W} \neq \{0\}$.

PROOF: Let $\{u_j\}_{j=1}^l$ be a basis for $\mathscr{U}$ and $\{w_j\}_{j=1}^m$ be a basis for $\mathscr{W}$. Since $l + m > n$, the set $\{u_j\}_{j=1}^l \cup \{w_j\}_{j=1}^m$ is linearly dependent, i.e., there exists a nontrivial vanishing linear combination

$$\sum c_j u_j + \sum d_j w_j = 0.$$

If all the coefficients c_j were zero, we would have a vanishing nontrivial combination of the basis elements $\{w_j\}_{j=1}^m$, which is ruled out. Similarly not all the d_j's vanish. We now have in $\mathscr{U} \cap \mathscr{W}$ the nontrivial vector $v = \sum c_j u_j = -\sum d_j w_j$. ◀

1.3.8 Infinite-dimensional vector spaces. Many examples of vector spaces, in many areas of mathematics, are *infinite-dimensional,* i.e., they do not have finite bases. Assuming *the axiom of choice*, it can be shown that every spanning set in an arbitrary vector space $\mathscr{V}$ can be trimmed down to a *minimal spanning set*, i.e., one that does not contain a proper subset that is spanning. Likewise, every independent set is contained in a *maximal independent set*, i.e., one that is not properly contained in a larger linearly independent set.

A subset $\mathscr{B}$ which is either a maximal independent set or a minimal spanning set is a basis in the sense that every $v \in \mathscr{V}$ is equal to a unique *finite* linear combination of elements of $\mathscr{B}$. A basis, so defined, is called a *Hamel basis* for an infinite-dimensional space.

In a typical study involving infinite-dimensional vector spaces this is irrelevant! Infinite-dimensional spaces (of interest) usually come with a *topology*, which allows one to introduce the notion of *convergence.* In this context, *convergent infinite sums* are allowed, and bases are defined accordingly. The field devoted to the study of topological vector spaces is *Functional Analysis.*

EXERCISES FOR SECTION 1.3

ex1.3.1 Show that the set $\{v_j : 1 \le j \le k\}$ is linearly dependent if and only if $v_1 = 0$, or there exists $l \in [2, k]$ such that v_l is a linear combination of vectors in $\{v_j : 1 \le j \le l-1\}$.

ex1.3.2 Let $\mathscr{V}$ be a vector space, and $\mathscr{W} \subset \mathscr{V}$ a subspace[5]. Let $v, u \in \mathscr{V} \setminus \mathscr{W}$, and assume that $u \in \operatorname{span}[\mathscr{W}, v]$. Prove that $v \in \operatorname{span}[\mathscr{W}, u]$.

ex1.3.3 What is the dimension of $\mathbb{C}^5$ considered as a vector space over $\mathbb{R}$?

ex1.3.4 Is $\mathbb{R}$ finite-dimensional over $\mathbb{Q}$?

ex1.3.5 Let $\mathscr{U}$, $\mathscr{W}$ be subspaces of a vector space $\mathscr{V}$, with $\mathscr{U} \cap \mathscr{W} = \{0\}$. Assume that $\{u_1, \dots, u_k\} \subset \mathscr{U}$ and $\{w_1, \dots, w_l\} \subset \mathscr{W}$ are linearly independent sets. Prove that $\{u_1, \dots, u_k\} \cup \{w_1, \dots, w_l\}$ is linearly independent.

ex1.3.6 Let $\mathscr{V}$ be finite-dimensional. Prove that every subspace $\mathscr{W} \subset \mathscr{V}$ is finite-dimensional, and that $\dim \mathscr{W} \le \dim \mathscr{V}$ with equality only if $\mathscr{W} = \mathscr{V}$.

[5] $\mathscr{V} \setminus \mathscr{W}$ denotes the difference set $\{v : v \in \mathscr{V} \text{ and } v \notin \mathscr{W}\}$.

ex1.3.7 Show that if $\mathscr{V}$ is finite-dimensional, then every subspace $\mathscr{W} \subset \mathscr{V}$ is a direct summand, i.e., there is a subspace $\mathscr{W}' \subset \mathscr{V}$ such that $\mathscr{V} = \mathscr{W} \oplus \mathscr{W}'$. Show by example that $\mathscr{W}'$ need not be unique (see exercise **ex1.2.8**).

ex1.3.8 Let $\mathscr{V}$ be a finite-dimensional vector space, and $\mathscr{A}$ a collection of subspaces of $\mathscr{V}$. Prove that there are one or more *minimal elements* in $\mathscr{A}$, that is $\mathscr{W} \in \mathscr{A}$ such that no element in $\mathscr{A}$ is a proper subspace of $\mathscr{W}$.

Similarly, show that $\mathscr{A}$ has one or more *maximal elements*, i.e., elements that are not contained in any other element of $\mathscr{A}$.

Remark: If $\mathscr{V}_j$ are subspaces of $\mathscr{V}$ such that none is contained in another, then every $\mathscr{V}_j$ is both minimal and maximal.

ex1.3.9 Assume that $\mathscr{V}$ is n-dimensional, and let $\mathscr{W}_j, j = 1, \dots, k$ be subspaces of $\mathscr{V}$ such that, for $1 \le j < k$, $\mathscr{W}_{j+1}$ is a proper subspace of $\mathscr{W}_j$. Prove that $k \le n$.

ex1.3.10 Let $\mathscr{V}$ and $\mathscr{W}$ be finite-dimensional subspaces of a vector space. Prove that $\mathscr{V} + \mathscr{W}$ and $\mathscr{V} \cap \mathscr{W}$ are finite-dimensional and that

$$\dim(\mathscr{V} \cap \mathscr{W}) + \dim(\mathscr{V} + \mathscr{W}) = \dim \mathscr{V} + \dim \mathscr{W}. \tag{1.3.4}$$

ex1.3.11 Repeat exercise **ex1.2.11** in the context of finite-dimensional spaces as follows:

a. Choose a basis $\{e_j\}$ for $\mathscr{U} \cap \mathscr{W}$;

b. Complete it to a basis for $\mathscr{U}$ by adding the vectors $\{u_k\}$;

c. Complete it also to a basis for $\mathscr{W}$ by adding the vectors $\{w_l\}$;

d. Check that $\{e_j\} \cup \{u_k\} \cup \{w_l\}$ is a basis for $\mathscr{U} + \mathscr{W}$;

e. Identify bases for the two quotient spaces invoved.

ex1.3.12 Show that if $\mathscr{W}_j$, $j = 1, \dots, k$, are finite-dimensional subspaces of a vector space $\mathscr{V}$, then $\sum \mathscr{W}_j$ is finite-dimensional and $\dim \sum \mathscr{W}_j \le \sum \dim \mathscr{W}_j$, with equality if and only if the subspaces $\mathscr{W}_j$ are independent.

ex1.3.13 Let $\mathscr{V}$ be an n-dimensional vector space, and let $\mathscr{V}_1 \subset \mathscr{V}$ be a subspace of dimension m.

a. Prove that the quotient space $\mathscr{V}/\mathscr{V}_1$ is finite-dimensional.

b. Let $\{v_1, \dots, v_m\}$ be a basis for $\mathscr{V}_1$ and let $\{\tilde{w}_1, \dots, \tilde{w}_k\}$ be a basis for $\mathscr{V}/\mathscr{V}_1$. For $j \in [1, k]$, let w_j be an element of the coset $\tilde{w}_j$.

Prove: $\{v_1, \dots, v_m\} \cup \{w_1, \dots, w_k\}$ is a basis for $\mathscr{V}$. Hence $k + m = n$.

ex1.3.14 Let $\mathscr{V}$ be a real vector space, and let $v_1, \dots, v_p \in \mathscr{V}$ be linearly independent. Let $r_l = [a_{l,1}, \dots, a_{l,p}]$, $1 \le l \le s$ be linearly independent vectors in $\mathbb{R}^p$. Prove that the vectors $u_l = \sum_1^p a_{l,j} v_j$, $l = 1, \dots, s$, are linearly independent in $\mathscr{V}$.

ex1.3.15 Let $\mathscr{V}$ and $\mathscr{U}$ be finite-dimensional spaces over $\mathbb{F}$. Prove that the tensor product $\mathscr{V} \otimes \mathscr{U}$ is finite-dimensional. Specifically, if $\{e_j\}_{j=1}^n$ is a basis for $\mathscr{V}$, and $\{f_k\}_{k=1}^m$ is a basis for $\mathscr{U}$, then $\{e_j \otimes f_k\}$, $1 \le j \le n$, $1 \le k \le m$, is a basis for $\mathscr{V} \otimes \mathscr{U}$, so that $\dim \mathscr{V} \otimes \mathscr{U} = \dim \mathscr{V} \dim \mathscr{U}$.

$*$**ex1.3.16** Assume that $\mathscr{V}$ is n-dimensional vector space over an infinite field $\mathbb{F}$, and let $\{\mathscr{W}_j\}$ be a finite collection of distinct m-dimensional subspaces.

a. Prove that no $\mathscr{W}_j$ is contained in the union of the others.

b. Prove that there is a subspace $\mathscr{U} \subset \mathscr{V}$ which is a complement of every $\mathscr{W}_j$.

Hint: See exercise **ex1.2.12**.

$*$**ex1.3.17** Assume that any three of the five $\mathbb{R}^3$-vectors $v_j = (x_j, y_j, z_j)$, $j = 1, \dots, 5$, are linearly independent. Prove that the vectors

$$w_j = (x_j^2, y_j^2, z_j^2, x_j y_j, x_j z_j, y_j z_j)$$

are linearly independent in $\mathbb{R}^6$.

Hint: Find nonzero (a,b,c) such that $ax_j + by_j + cz_j = 0$ for $j = 1,2$. Find nonzero (d,e,f) such that $dx_j + ey_j + fz_j = 0$ for $j = 3,4$. Observe (and use) the fact

$$(ax_5 + by_5 + cz_5)(dx_5 + ey_5 + fz_5) \neq 0.$$

1.4 Systems of linear equations

How do we find out if a set $\{v_j\}$, $j = 1, \dots, m$ of column vectors in $\mathbb{F}_{\mathbf{c}}^n$ is linearly dependent? How do we find out if a vector u belongs to $\mathsf{span}[v_1, \dots, v_m]$?

Given the vectors $v_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}$, $j = 1, \dots, m$, and $u = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, we express the conditions $\sum x_j v_j = 0$ for the first question, and $\sum x_j v_j = u$ for the second, in terms of the coordinates.

The first equation leads to the *system of homogeneous linear equations:*

$$
\begin{aligned}
a_{11}x_1 + \quad \dots \quad + a_{1m}x_m &= 0\\
a_{21}x_1 + \quad \dots \quad + a_{2m}x_m &= 0\\
\vdots \qquad\qquad\quad \vdots \quad &\\
a_{n1}x_1 + \quad \dots \quad + a_{nm}x_m &= 0
\end{aligned} \tag{1.4.1}
$$

or,

$$\sum_{j=1}^{m} a_{ij}x_j = 0, \quad i = 1, \dots, n. \tag{1.4.2}$$

The second equation gives the *nonhomogeneous system:*

$$\sum_{j=1}^{m} a_{ij}x_j = c_i, \quad i = 1, \dots, n. \tag{1.4.3}$$

DEFINITION: The *solution-set* of a system of linear equations, such as (1.4.2) or (1.4.3), is the set of all m-tuples $(x_1, \dots, x_m) \in \mathbb{F}^m$ for which all n equations hold.

To answer the question of the dependence of the v_j's, we need to determine if the solution-set of the system (1.4.2) is *trivial* or not, i.e., if there are solutions other than $(0, \dots, 0)$. To see if $u \in \mathsf{span}[v_1, \dots, v_m]$, we need to know if the solution-set of (1.4.3) is empty or not.

In both cases we would like to identify the solution-set as completely and as explicitly as possible.

1.4.1 Conversely, beginning with a homogeneous system like (1.4.2) we can rewrite it as

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \cdots + x_m \begin{bmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{bmatrix} = 0 \tag{1.4.4}$$

and use general properties of vector spaces to draw general conclusions. Our first result depends only on dimension.

Theorem. *A system of n homogeneous linear equations in $m > n$ unknowns has a nontrivial solution.*

PROOF: The m columns in (1.4.4) are elements of the n-dimensional space $\mathbb{F}_{\mathbf{c}}^n$. If $m > n$, then they are dependent, so (1.4.4) has a nontrivial solution. ◀

Remark: As noted in example ***a*** of subsection 1.2.3, the set of solutions of a homogeneous system of equations form a subspace of $\mathbb{F}^m$. The theorem above shows that if the number of variables is greater than the number of equations, then the dimension of the subspace of solutions is at least 1. See **ex1.4.1** for a refinement.

Similarly, a nonhomogeneous system like (1.4.3) can be rewritten in the form

$$x_1\begin{bmatrix}a_{11}\\ \vdots\\ a_{n1}\end{bmatrix}+\cdots+x_m\begin{bmatrix}a_{1m}\\ \vdots\\ a_{nm}\end{bmatrix}=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}. \tag{1.4.5}$$

It is then clear that the system given by (1.4.3) has a solution if and only if

the column $\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$ is in the span of the columns $\begin{bmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{bmatrix}$, $j = 1, \dots, m$.

1.4.2 The classical approach to solving systems of linear equations is by *Gaussian elimination*—an algorithm for replacing the given system by an *equivalent* system that can be solved easily. We need some terminology:

DEFINITION: The systems

$$\begin{aligned}(\mathfrak{A})\qquad & \sum_{j=1}^{m} a_{ij}x_j = c_i, \quad i = 1, \dots, k,\\ (\mathfrak{B})\qquad & \sum_{j=1}^{m} b_{ij}x_j = d_i, \quad i = 1, \dots, l,\end{aligned} \tag{1.4.6}$$

are *equivalent* if they have the same solution-set (in $\mathbb{F}^m$).

The matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & \dots & \vdots \\ a_{k1} & \dots & a_{km} \end{bmatrix} \quad \text{and} \quad A_{aug} = \begin{bmatrix} a_{11} & \dots & a_{1m} & c_1 \\ a_{21} & \dots & a_{2m} & c_2 \\ \vdots & \dots & \vdots & \vdots \\ a_{k1} & \dots & a_{km} & c_k \end{bmatrix}$$

are called the *coefficient matrix,* or simply the *matrix,* and the *augmented matrix* of the system $(\mathfrak{A})$. The augmented matrix is obtained from the matrix by appending the column of *values* (i.e., the right-hand sides of the equations in the system) as a (new) last column.

The augmented matrix contains all the information of the system $(\mathfrak{A})$. Any $k \times (m+1)$ matrix is the augmented matrix of a system of linear equations in m unknowns.

1.4.3 Row equivalence of matrices.

DEFINITION: The *row space* of a matrix $A \in \mathscr{M}(k,m)$ is the subspace of $\mathbb{F}_{\mathbf{r}}^m$ spanned by the rows of A. The dimension of this space is called the *row rank* of A.

The matrices

$$\begin{bmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & \dots & \vdots \\ a_{k1} & \dots & a_{km} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \vdots & \dots & \vdots \\ b_{l1} & \dots & b_{lm} \end{bmatrix} \tag{1.4.7}$$

are *row equivalent* if their rows span the same subspace of $\mathbb{F}_{\mathbf{r}}^m$; equivalently: if each row of either matrix is a linear combination of the rows of the other. Row equivalent matrices clearly have the same row rank.

Proposition. *Two systems of linear equations in m unknowns*

$$(\mathfrak{A}) \qquad \sum_{j=1}^{m} a_{ij}x_j = c_i, \quad i = 1,\dots,k,$$

$$(\mathfrak{B}) \qquad \sum_{j=1}^{m} b_{ij}x_j = d_i, \quad i = 1,\dots,l,$$

are equivalent if their respective augmented matrices are row equivalent.

PROOF: Assume that the augmented matrices are row equivalent.

If $(x_1, \dots, x_m)$ is a solution for system $(\mathfrak{A})$ and

$$(b_{i1}, \dots, b_{im}, d_i) = \sum \alpha_{i,k}(a_{k1}, \dots, a_{km}, c_k),$$

then

$$\sum_{j=1}^{m} b_{ij}x_j = \sum_{k,j} \alpha_{i,k}a_{kj}x_j = \sum_{k} \alpha_{i,k}c_k = d_i$$

and $(x_1, \dots, x_m)$ is a solution for system $(\mathfrak{B})$. ◀

1.4.4 Reduced-row-echelon form. We come now to *Gauss-Jordan elimination*. The equivalent system that is easier to solve is obtained by reducing the augmented matrix of the system to one in *reduced-row-echelon form*.

DEFINITION: A matrix $A \in \mathscr{M}(k,m)$ is in *reduced-row-echelon form* if the following conditions are satisfied:

rref–1 The first q rows of A are linearly independent in $\mathbb{F}^m$, and the remaining $k-q$ rows are zero.

rref–2 There are integers $1 \le l_1 < l_2 < \cdots < l_q \le m$ such that for $j \le q$, the first nonzero entry in the j'th row is 1, occuring in the l_j'th column.

rref–3 The entry 1 in row j is the only nonzero entry in the l_j column.

One can rephrase the last three conditions as: The l_j'th columns, called *pivot* columns, are the first q elements of the standard basis of $\mathbb{F}^k_{\mathbf{c}}$; every other column is a linear combination of the pivot columns that precede it.

Theorem. *Every matrix is row-equivalent to a matrix in reduced-row-echelon form. Furthermore, the reduced-row-echelon form of a matrix is unique.*

PROOF: We describe an algorithm that uses *elementary row operations* to reduce an arbitrary matrix A to a row-equivalent matrix in reduced-row-echelon form.

The *elementary row operations* are:

a. Reordering (i.e., permuting) the rows;

b. Multiplying a row by a nonzero constant;

c. Adding a multiple of one row to another.

These operations do not change the span of the rows so that the row equivalence class of the matrix is maintained. (We shall return later, in exercise **ex2.3.13**, to express these operations as matrix multiplications.)

If $A = 0$, there is nothing to prove, and we assume that $A \neq 0$. Denote the row rank of A by q. Let l_1 be the index of the first column that is not zero.

Reorder the rows so that $a_{1,l_1} \neq 0$, and multiply the first row by a_{1,l_1}^{-1}.

For every $j > 1$, subtract the first row multiplied by a_{j,l_1} from the j'th row.

Now all the columns before l_1 are zero and column l_1 has 1 in the first row, and zero elsewhere.

If the row rank q is 1, all the entries below the first row are now zero and we are done. Otherwise let l_2 be the index of the first column that has a nonzero entry in a row beyond the first. Notice that $l_2 > l_1$. Keep the first row in its place, reorder the remaining rows so that $a_{2,l_2} \neq 0$, and multiply the second row[6] by a_{2,l_2}^{-1}.

For every $j \neq 2$, subtract the second row multiplied by a_{j,l_2} from the j'th row.

Repeat the sequence of steps a total of q times. The first q rows, $\mathbf{r}_1, \dots, \mathbf{r}_q$, are (now) independent: a combination $\sum c_j \mathbf{r}_j$ has entry c_j in the l_j'th place, and can be zero only if $c_j = 0$ for all j.

If there is a nonzero entry beyond the current q'th row, necessarily beyond the l_q'th column, we could continue and get a row independent of the first q, contradicting the definition of q. Thus, after q steps, all the rows beyond the q'th are zero.

[6] We keep referring to the entries of the successively modified matrix as a_{ij}.

Uniqueness of the reduced-row-echelon-form of a matrix is left as exercise **ex1.4.4**. ◀

Observe that the scalars used in the process belong to the smallest field that contains all the coefficients of A.

1.4.5 If A and A_{aug} are the matrix and the augmented matrix of a system $(\mathfrak{A})$ and we apply the algorithm of the previous subsection to both, we observe that since the augmented matrix has the additional column on the right-hand side, the first q (the row rank of A) steps in the algorithm for either A or A_{aug} are identical. Having done q repetitions, A is in reduced-row-echelon form, while A_{aug} may or may not be. If the row rank of A_{aug} is q, then the algorithm for A_{aug} ends as well; otherwise we have $l_{q+1} = m+1$, and the reduced-row-echelon form for the augmented matrix is the same as that of A but with an added row and an added pivot column, both having 0 for all but the last entries, and 1 for the last entry. In the latter case, the system corresponding to the row-reduced augmented matrix has as its last equation $0 = 1$ and the system has no solutions.

On the other hand, if the row rank of the augmented matrix is the same as that of A, the reduced-row-echelon form of the augmented matrix is an augmentation of the reduced-row-echelon form of A. In this case we assign arbitrary values to the so-called *free variables*, i.e., the variables x_i, $i \neq l_j$, $j = 1, \dots, q$. We then move the corresponding terms to the right-hand side and, writing C_j for their sum, we obtain

$$x_{l_j} = C_j, \quad j = 1, \dots, q. \tag{1.4.8}$$

Theorem. *A necessary and sufficient condition for the system* $(\mathfrak{A})$ *to have solutions is that the row ranks of the matrix and of the augmented matrix of the system be equal.*

The discussion preceding the statement of the theorem not only proves the theorem but offers a concrete way to solve the system. The unknowns are now split into two groups, q pivot variables and $m-q$ free ones. We have "$m-q$ degrees of freedom": the $m-q$ free unknowns become free parameters that can be assigned arbitrary values, and these values determine the pivot unknowns uniquely.

Remark: Notice that the split into pivot and free unknowns depends on the specific definition of reduced-row-echelon form; counting the columns in a different order may result in a different split, though the number q of pivot variables would be the same, equal to the row rank of A. For example, for the "system" of one equation with two unknowns $x+y=1$, either x or y can be chosen freely, thereby determining the other.

Corollary. *A linear system of n equations in n unknowns with matrix A has solutions for all augmented matrices if and only if the only solution of the corresponding homogeneous system is the trivial solution.*

PROOF: The condition on the homogeneous system amounts to "the rows of A are independent", and no added columns can increase the row rank. ◀

1.4.6 DEFINITION: The *column space* of a matrix $A \in \mathscr{M}(k,m)$ is the subspace of $\mathbb{F}_{\mathbf{c}}^k$ spanned by the columns of A. The dimension of this space is called the *column rank* of A.

Theorem. *The column rank of a matrix A is equal to its row rank.*

PROOF: Linear relations between columns of A are solutions of the homogeneous system given by A. If B is row-equivalent to A, the columns of A and B have the same set of linear relations (see Proposition 1.4.3). In particular, if the row rank of A is q and B is in reduced-row-echelon form, then the q pivot columns in B are independent, and every other column is a linear combination of these. ◀

We refer to the common value of the row and column ranks of A simply as *the rank of* A and denote it by $\rho(A)$.

1.4.7 DEFINITION. A *submatrix* of a matrix A is a matrix B obtained by deleting from A some rows and some columns.[7]

B is a *principal submatrix* of a square matrix A if the set of indices of the deleted columns is the same as that of the deleted rows.

[7] We use the word *some* to mean *a nonnegative number of*, that is, *some or none.*

A simple corollary of Theorem 1.4.6 is the following proposition.

Proposition. *If A is a matrix and $\rho(A) = k$, then there is a $k \times k$ submatrix B of A such that $\rho(B) = k$.*

See **ex1.4.12** for a refinement.

EXERCISES FOR SECTION 1.4

ex1.4.1 Show that if $m > n$, then the dimension of the space of solutions of a homogeneous system of n linear equations in m variables is at least $m - n$.

ex1.4.2 Prove Proposition 1.4.7, and show an example of a 3×3 matrix of rank 2 without a principal submatrix of rank 2. (So that the second part of the theorem is false without some assumption of symmetry.)

ex1.4.3 Identify the matrix $A \in \mathscr{M}(n)$ of row rank n that is in reduced-row-echelon form.

ex1.4.4 Show that if $A, B \in \mathscr{M}(k,m)$ are both in reduced-row-echelon form, then either $A = B$, or A and B are not row equivalent. Conclude that the reduced-row-echelon form of a matrix is unique.
Hint: Consider the solution sets of the two homogeneous systems with coefficient matrices A and B, respectively. Show that these must be different if $A \neq B$.

ex1.4.5 A system of linear equations with rational coefficients that has a solution in $\mathbb{C}$, has a solution in $\mathbb{Q}$. Equivalently, vectors in $\mathbb{Q}^n$ that are linearly dependent over $\mathbb{C}$ are rationally dependent.
Hint: The last sentence of subsection 1.4.4.

ex1.4.6 A system of linear equations with rational coefficients, has the same number of degrees of freedom over $\mathbb{Q}$ as it does over $\mathbb{C}$.[8]

ex1.4.7 A subset $\mathscr{A}$ of a vector space $\mathscr{V}$ is called an *affine subspace* if $\mathscr{A}$ is the translate of a subspace $\mathscr{W}$ of $\mathscr{V}$, i.e., $\mathscr{A} = \{v_0 + w : w \in \mathscr{W}\}$. We call the subspace $\mathscr{W}$ the *corresponding subspace* to $\mathscr{A}$. (Thus a *line* in $\mathscr{V}$ is a translate of a one-dimensional subspace.)

Prove that a set $\mathscr{A} \subset \mathscr{V}$ is an affine subspace if and only if $\sum a_j u_j \in \mathscr{A}$ for all choices of $u_1, \dots, u_k \in \mathscr{A}$, and scalars a_j, $j = 1, \dots, k$ such that $\sum a_j = 1$.

[8]See 1.4.5.

ex1.4.8 Let $\mathscr{A} \subset \mathscr{V}$ be an affine subspace and $u_0 \in \mathscr{A}$. Prove that the set $\mathscr{A} - u_0 = \{u - u_0 : u \in \mathscr{A}\}$ is the corresponding subspace of $\mathscr{A}$ in $\mathscr{V}$. Show that the corresponding subspace $\mathscr{A} - u_0$ does not depend on the choice of u_0 in $\mathscr{A}$.

ex1.4.9 Show that the solution-set of a system of k linear equations in m unknowns is an affine subspace of $\mathbb{F}^m$. What is its corresponding subspace?

ex1.4.10 A column $v_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{kj} \end{bmatrix}$ of a matrix $A = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \dots & \vdots \\ a_{k1} & \dots & a_{km} \end{bmatrix}$ is called a pivot column if j is the index of a pivot column in the reduced-row-echelon form of A. Show that v_j is a pivot column of A if and only if it is linearly independent of the columns v_i, $i < j$.

ex1.4.11 Denote by $B = \begin{bmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \dots & \vdots \\ b_{k1} & \dots & b_{km} \end{bmatrix}$ the reduced-row-echelon form of the matrix A in the previous problem. Let $l_1 < l_2, \dots$ be the indices of the pivot columns in B and i the index of another column. Prove that

$$v_i = \sum_{l_j < i} b_{ji} v_{l_j}, \tag{1.4.9}$$

where, as in the previous problem, $v_1, \dots, v_m$ are the columns of A.

ex1.4.12 Show that if the matrix A in Proposition 1.4.7 is symmetric or skew-symmetric, then the submatrix B may be taken to be a *principal* submatrix.

ex1.4.13 What is the reduced-row-echelon form of the 7×6 matrix A, if its columns C_j, $j = 1, \dots, 6$, satisfy the following conditions:

a. $C_1 \neq 0$;

b. $C_2 = 3C_1$;

c. C_3 is not a (scalar) multiple of C_1;

d. $C_4 = C_1 + 2C_2 + 3C_3$;

e. $C_5 = 6C_3$;

f. C_6 is not in the span of C_2 and C_3.

∗**ex1.4.14** Given polynomials $P_1 = \sum_0^n a_j x^j$, $P_2 = \sum_0^m b_j x^j$, and $S = \sum_0^l s_j x^j$ of degrees n, m, and $l < n+m$ respectively, suppose that we want to find polynomials $q_1 = \sum_0^{m-1} c_j x^j$ and $q_2 = \sum_0^{n-1} d_j x^j$ such that

$$P_1 q_1 + P_2 q_2 = S. \tag{1.4.10}$$

The polynomial equation (1.4.10) is equivalent to the system of $m+n$ linear equations,

$$\sum_{j+k=l} a_j c_k + \sum_{r+t=l} b_r d_t = s_l, \quad l = 0, \dots, n+m-1, \tag{1.4.11}$$

where the unknowns are the coefficients $c_{m-1}, \dots, c_0$ of q_1, and $d_{n-1}, \dots, d_0$ of q_2. The coefficient matrix for this system is known as the *Sylvester matrix* of (P_1, P_2).

Write the matrix of the system for $n = 3$ and $m = 2$.

∗**ex1.4.15** The associated homogeneous system to the system (1.4.11) corresponds to the case $S = 0$. Show that it has a nontrivial solution if and only if P_1 and P_2 have a nontrivial common factor. (You may assume the unique factorization theorem; see A.6.3 in the appendix.) What is the rank of the Sylvester matrix if the degree of $\gcd(P_1, P_2)$ is r?

∗1.5 Normed finite-dimensional linear spaces

1.5.1 A norm on a real or complex vector space $\mathcal{V}$ is a nonnegative function $v \mapsto \|v\|$ that satisfies the conditions

a. Positivity: $\|0\| = 0$ and if $v \neq 0$ then $\|v\| > 0$.

b. Homogeneity: $\|av\| = |a|\|v\|$ for scalars a and vectors v.

c. The triangle inequality: $\|v+u\| \leq \|v\| + \|u\|$.

These properties guarantee that $\delta(v,u) = \|v-u\|$ is a metric on the space, the metric *defined* or *induced* by the norm; and with a metric one can use tools and notions from point-set topology such as limits, continuity, convergence, infinite series, etc.

A *normed vector space* is a vector space endowed with a norm. Since $\mathbb{R}$ and $\mathbb{C}$ are *complete* metric spaces, so is every normed finite-dimensional real or complex vector space.

1.5.2 If $\mathcal{V}$ and $\mathcal{W}$ are isomorphic real or complex finite-dimensional spaces and S is an isomorphism of $\mathcal{V}$ onto $\mathcal{W}$, then a norm $\|\cdot\|_{\mathcal{W}}$ on $\mathcal{W}$ can be "carried back" to $\mathcal{V}$ by defining $\|v\|_{\mathcal{V}} = \|Sv\|_{\mathcal{W}}$. This implies that all possible norms on a real n-dimensional space are copies of norms on $\mathbb{R}^n$, and all norms on a complex n-dimensional space are copies of norms on $\mathbb{C}^n$.

A finite-dimensional $\mathcal{V}$ can be endowed with many different norms; yet, all these norms are *equivalent* in the following sense (see **ex1.5.1**):

DEFINITION: The norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent* if there is a positive constant C such that for all $v \in \mathcal{V}$,

$$C^{-1}\|v\|_1 \le \|v\|_2 \le C\|v\|_1.$$

Metrics δ_1, δ_2, defined by equivalent norms $\|\cdot\|_1$ and $\|\cdot\|_2$, are equivalent: for $v, u \in \mathcal{V}$,

$$C^{-1}\delta_1(v,u) \le \delta_2(v,u) \le C\delta_1(v,u),$$

which means that they define the same topology—the familiar topology of $\mathbb{R}^n$ or $\mathbb{C}^n$.

EXERCISES FOR SECTION 1.5

ex1.5.1 Let $\mathcal{V}$ be an n-dimensional real or complex vector space, and let $\mathbf{v} = \{v_1, \dots, v_n\}$ be a basis for $\mathcal{V}$. Write

$$\left\|\sum a_j v_j\right\|_{\mathbf{v},1} = \sum |a_j| \qquad \text{and} \qquad \left\|\sum a_j v_j\right\|_{\mathbf{v},\infty} = \max|a_j|.$$

Prove:

a. $\|\cdot\|_{\mathbf{v},1}$ and $\|\cdot\|_{\mathbf{v},\infty}$ are norms on $\mathcal{V}$, and

(1.5.1)
$$\|\cdot\|_{\mathbf{v},\infty} \le \|\cdot\|_{\mathbf{v},1} \le n\|\cdot\|_{\mathbf{v},\infty}.$$

b. If $\|\cdot\|$ is *any* norm on $\mathcal{V}$ then, for all $v \in \mathcal{V}$,

(1.5.2)
$$\|v\|_{\mathbf{v},1} \max\|v_j\| \ge \|v\|.$$

c. Let $\|\cdot\|_j$, $j = 1,2$, be norms on $\mathcal{V}$, and δ_j the induced metrics. Let $\{v_n\}_{n=0}^{\infty}$ be a sequence in $\mathcal{V}$. Prove that if $\delta_1(v_n, v_0) \to 0$, then $\delta_2(v_n, v_0) \to 0$.

d. Show that if $\|\cdot\|$ is an arbitrary norm on $\mathscr{V}$, then the function $f : \mathbb{F}^n \to \mathbb{R}$ (where $\mathbb{F}$ is the field of scalars of $\mathscr{V}$, either $\mathbb{R}$ or $\mathbb{C}$), defined by

$$f(a_1, \dots, a_n) = \left\| \sum a_j v_j \right\|,$$

is continuous on $\mathbb{F}^n$. Conclude that f attains a strictly positive minimum on the set $B = \{(a_1, \dots, a_n) : \sum |a_j| = 1\} \subset \mathbb{F}^n$.

Hint: Show that the triangle inequality implies that $\big| \|v\| - \|u\| \big| \le \|v - u\|$, then use part **b** and the compactness of B.

e. Conclude that all norms on $\mathscr{V}$ are equivalent to $\|\cdot\|_{\mathbf{v},1}$.

ex1.5.2 Let $\{v_n\}_{n=0}^{\infty}$ be bounded in $\mathscr{V}$. Prove that $\sum_0^\infty z^n v_n$ converges for every z such that $|z| < 1$.

Hint: Prove that the partial sums form a Cauchy sequence in the metric defined by the norm.

ex1.5.3 Let $\mathscr{V}$ be n-dimensional real or complex normed vector space. The *unit ball* in $\mathscr{V}$ is the set

$$B_1 = \{v \in \mathscr{V} : \|v\| \le 1\}.$$

Prove that B_1 is

a. *Convex*: If $v, u \in B_1$, $0 \le a \le 1$, then $av + (1-a)u \in B_1$.

b. *Bounded:* For every $v \in \mathscr{V}$, there exists a (positive) constant λ such that $cv \notin B$ for $|c| > \lambda$.

c. *Circularly symmetric, centered at* 0*:* If $v \in B$ and $|a| = 1$ then $av \in B$.

Notice that convexity and circular symmetry together imply that if $v \in B$ and $|a| \le 1$ then $av \in B$.

ex1.5.4 Let $\mathscr{V}$ be n-dimensional real or complex vector space, and let B be a bounded circularly symmetric convex set centered at 0. Define

$$\|u\| = \inf\{a > 0 : a^{-1}u \in B\}.$$

Prove that this defines a norm on $\mathscr{V}$, and the unit ball for this norm is the given set B.

ex1.5.5 Describe a norm $\|\ \|_0$ on $\mathbb{R}^3$ such that the standard unit vectors have norm 1 while $\|(1,1,1)\|_0 < \frac{1}{100}$.