$\{\mathcal{L}\vec{i}\mathbb{N}\mathbf{e}^{*}\mathfrak{a}\|r\|\bigoplus_{i=0}^{\infty}\mathbf{Alg}|^{\perp}_{\langle\varepsilon,\beta\rangle(\mathrm{r}^{a})}\}\ \overline{\circ\tilde{\mathrm{v}}\mathrm{er}}\ \mathbf{R}\xrightarrow{\mathrm{and}'}\mathbf{C}$ $\cancel{\mathrm{without}}^{-1}\ [\mathcal{M}\mathrm{atrices}]^{\mathsf{T}}_{\cdot,j}$

Final Version (2021-9-5)

Qiū Cáiyóng

R should be understood by its only nontrivial algebraic extension: **C**.

R and **C** should be understood by their modules: real/complex vector spaces. Real/complex vector spaces should be understood by linear maps between them. The structure of linear operators on a finite-dimensional complex vector space (**Schur's Theorem**) is our final goal.

This part of mathematics is called **Linear Algebra**. We will cover some more advanced topics in the sequel **Advanced Linear Algebra**.

邱才敏.

Contents

Chapter 1

Vector Spaces

1.1 Complex numbers

Complex numbers were invented so that we can take square roots of negative numbers. The idea is to assume we have a square root of -1, denoted i, that obeys the usual rules of arithmetic.

1.1.1 Definition of complex numbers

Definition 1.1.1 (complex numbers)**.** A complex number is an ordered pair (a, b), where $a, b \in \mathbf{R}$, but we will write this as $a + bi$. The set of all complex numbers is denoted by $\mathbf{C}$:

$$\mathbf{C} = \{a + bi : a, b \in \mathbf{R}\}.$$

Addition and multiplication on $\mathbf{C}$ are defined by

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

here $a, b, c, d \in \mathbf{R}$.

If $a \in \mathbf{R}$, we identify $a + 0i$ with the real number a. Thus we can think of $\mathbf{R}$ as a subset of $\mathbf{C}$. We also usually write $0 + bi$ as just bi. and we usually write $0 + 1i$ as just i. Using multiplication as defined above, you should verify that $i^2 = -1$.

Example 1.1.1.

$$(2 + 3i)(4 + 5i) = -7 + 22i.$$

1.1.2 Properties of complex arithmetic

commutativity $\alpha+\beta=\beta+\alpha, \alpha\beta=\beta\alpha$ for all $\alpha,\beta\in\mathbf{C}$;

associativity $(\alpha+\beta)+\lambda=\alpha+(\beta+\lambda), (\alpha\beta)\lambda=\alpha(\beta\lambda)$ for all $\alpha,\beta,\lambda\in\mathbf{C}$;

identities $\lambda+0=\lambda, \lambda 1=\lambda$ for all $\lambda\in\mathbf{C}$;

additive inverse for every $\alpha\in\mathbf{C}$, there exists a unique $\beta\in\mathbf{C}$ such that $\alpha+\beta=0$;

multiplicative inverse for every $\alpha\in\mathbf{C}$ with $\alpha\neq 0$, there exists a unique $\beta\in\mathbf{C}$ such that $\alpha\beta=1$;

distributive property $\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$ for all $\lambda,\alpha,\beta\in\mathbf{C}$.

The properties above are proved using the familiar properties of real numbers and the definitions of complex addition and multiplication.

Example 1.1.2. Show that $\alpha\beta=\beta\alpha$ for all $\alpha,\beta\in\mathbf{C}$.

Proof. Suppose $\alpha=a+bi$ and $\beta=c+di$, where $a,b,c,d\in\mathbf{R}$. Then the definition of multiplication of complex numbers shows that

$$\alpha\beta=(ac-bd)+(ad+bc)i.$$

and

$$\beta\alpha=(ca-db)+(cb+da)i.$$

The equations above and the commutativity of multiplication and addition of real numbers show that $\alpha\beta=\beta\alpha$. ☐

Definition 1.1.2 (subtraction,division)**.** Let $\alpha,\beta\in\mathbf{C}$.

- Let $-\alpha$ denote the additive inverse of α. Thus $-\alpha$ is the unique complex number such that
$$\alpha+(-\alpha)=0.$$

- **Subtraction** on $\mathbf{C}$ is defined by
$$\beta-\alpha=\beta+(-\alpha).$$

- For $\alpha\neq 0$, let $1/\alpha$ denote the multiplicative inverse of α. Thus $1/\alpha$ is the unique complex number such that
$$\alpha(1/\alpha)=1.$$

- **Division** on $\mathbf{C}$ is defined by
$$\beta/\alpha=\beta(1/\alpha).$$

Throughout this book, $\mathbf{F}$ stands for either $\mathbf{R}$ or $\mathbf{C}$. Thus if we prove a theorem involving $\mathbf{F}$, we will know that it holds when $\mathbf{F}$ is replaced with $\mathbf{R}$ and when $\mathbf{F}$ is replaced with $\mathbf{C}$. Elements of $\mathbf{F}$ are called **scalars**.

For $\alpha \in \mathbf{F}$ and m a positive integer, we define α^m to denote the product of α with itself m times:

$$\alpha^m = \underbrace{\alpha \cdots \alpha}_{m \text{ times}}.$$

Clearly $(\alpha^m)^n = \alpha^{mn}$ and $(\alpha\beta)^m = \alpha^m\beta^m$ for all $\alpha, \beta \in \mathbf{F}$ and all positive integers m, n.

1.2 $\mathbf{F}^n$

1.2.1 Lists

Example 1.2.1.

- The set $\mathbf{R}^2$, which you can think of as a plane, is the set of all ordered pairs of real numbers:

$$\mathbf{R}^2 = \{(x, y) : x, y \in \mathbf{R}\}.$$

- The set $\mathbf{R}^3$, which you can think of as ordinary space, is the set of all ordered triples of real numbers:

$$\mathbf{R}^3 = \{(x, y, z) : x, y, z \in \mathbf{R}\}.$$

To generalize $\mathbf{R}^2$ and $\mathbf{R}^3$ to higher dimensions, we first need to discuss the concept of lists.

Definition 1.2.1 (list, length)**.** Suppose n is a nonnegative integer. A **list** of **length** n is an ordered collection of n elements (which might be numbers, other lists, or more abstract entities) separated by commas and surrounded by parentheses. A list of length n looks like this:

$$(x_1, \ldots, x_n).$$

Two lists are equal if and only if they have the same length and the same elements in the same order.

A list of length 2 is an ordered pair, and a list of length 3 is an ordered triple. We call a list of length n an n-tuple. Sometimes we will use the word **list** without specifying its length. Remember, however, that by definition each list has a finite length that is a nonnegative integer. Thus an object that looks like

$$(x_1, x_2, \ldots),$$

which might be said to have infinite length, is not a list.

A list of length 0 looks like this: (). We consider such an object to be a list so that some of our theorems will not have trivial exceptions.

Lists differ from sets in two ways: in lists, order matters and repetitions have meaning; in sets, order and repetitions are irrelevant:

Example 1.2.2 (lists versus sets)**.**

$$(3,5) \neq (5,3), \{3,5\} = \{5,3\}$$

$$(4,4) \neq (4,4,4), \{4,4\} = \{4,4,4\} = \{4\}$$

1.2.2 $\mathbf{F}^n$: the higher-dimensional analogues of $\mathbf{R}^2$ and $\mathbf{R}^3$

Fix a positive integer n for the rest of this section.

Definition 1.2.2 ($\mathbf{F}^n$)**.** $\mathbf{F}^n$ is the set of all lists of length n of elements of $\mathbf{F}$:

$$\mathbf{F}^n = \{(x_1, \ldots, x_n) : x_j \in \mathbf{F} \text{ for } j = 1, \ldots, n\}.$$

For $(x_1, \ldots, x_n) \in \mathbf{F}^n$ and $j \in \{1, \ldots, n\}$, we say that x_j is the j^{th} **coordinate** of $(x_1, \ldots, x_n)$.

If $n \geq 4$, we cannot visualize $\mathbf{R}^n$ as a physical object. Similarly, $\mathbf{C}^1$ can be thought of as a plane(the complex plane), but for $n \geq 2$, the human brain cannot provide a full image of $\mathbf{C}^n$. However, we can perform algebraic manipulations in $\mathbf{F}^n$.

Definition 1.2.3 (addition in $\mathbf{F}^n$)**. Addition** in $\mathbf{F}^n$ is defined by adding corresponding coordinates:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

Often the mathematics of $\mathbf{F}^n$ becomes cleaner if we use a single letter to denote a list of n numbers, without explicitly writing the coordinates. For example, the result below is stated with x and y in $\mathbf{F}^n$ even though the proof requires the more cumbersome notation of $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$.

Proposition 1.2.1 (Commutativity of addition in $\mathbf{F}^n$)**.** If $x, y \in \mathbf{F}^n$, then $x + y = y + x$.

Definition 1.2.4 (0)**.** Let 0 denote the list of length n whose coordinates are all 0:

$$0 = (0, \ldots, 0).$$

This potentially confusing practice actually causes no problems because the context always makes clear what is intended.

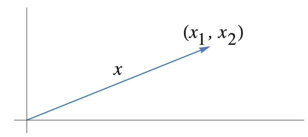
Example 1.2.3. Consider the statement that 0 is an additive identity for $\mathbf{F}^n$:

$$x + 0 = x \text{ for all } x \in \mathbf{F}^n.$$

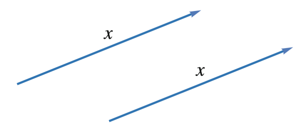
Is the 0 above the number 0 or the list 0?

Proof. Here 0 is a list, because we have not defined the sum of an element of $\mathbf{F}^n$ and the number 0. $\square$

1.2.3 Vectors



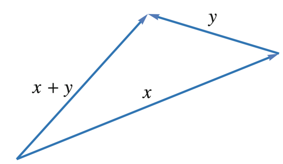
A picture can aid our intuition. We will draw pictures in $\mathbf{R}^2$ because we can sketch this space on 2-dimensional surfaces such as paper and blackboards. A typical element of $\mathbf{R}^2$ is a point $x = (x_1, x_2)$. Sometimes we think of x not as a point but as an arrow starting at the origin and ending at (x_1, x_2). When we think of x as an arrow, we refer to it as a **vector**.



When we think of vectors in $\mathbf{R}^2$ as arrows, we can move an arrow parallel to itself (not changing its length or direction) and still think of it as the same vector.

Whenever we use pictures in $\mathbf{R}^2$ or use the somewhat vague language of points and vectors, remember that these are just aids to our understanding, not substitutes for the actual mathematics that we will develop. Although we cannot draw good pictures in high-dimensional spaces, the elements of these spaces are as rigorously defined as elements of $\mathbf{R}^2$.

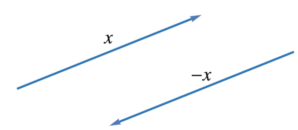
Recall that we defined the sum of two elements of $\mathbf{F}^n$ to be the element of $\mathbf{F}^n$ obtained by adding corresponding coordinates. As we will now see, addition has a simple geometric interpretation in the special case of $\mathbf{R}^2$. Suppose we have two vectors x and y in $\mathbf{R}^2$ that we want to add. Move the vector y parallel to itself so that its initial point coincides with the end point of the vector x, as shown here. The sum $x + y$ then equals the vector whose initial point equals the initial point of x and whose end point equals the end point of the vector y, as shown here.



Definition 1.2.5 (additive inverse in $\mathbf{F}^n$)**.** For $x \in \mathbf{F}^n$, the additive inverse of x, denoted $-x$, is the vector $-x \in \mathbf{F}^n$ such that

$$x + (-x) = 0.$$

In other words, if $x = x_1, \dots, x_n$, then $-x = (-x_1, \dots, -x_n)$.



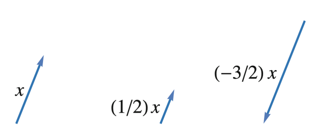
For a vector $x \in \mathbf{R}^2$, the additive inverse $-x$ is the vector parallel to x and with the same length as x but pointing in the opposite direction.

Definition 1.2.6 (scalar multiplication in $\mathbf{F}^n$)**.** The **product** of a number λ and a vector in $\mathbf{F}^n$ is computed by multiplying each coordinate of the vector by λ:

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n);$$

here $\lambda \in \mathbf{F}$ and $x_1, \dots, x_n \in \mathbf{F}^n$.

Scalar multiplication has a nice geometric interpretation in $\mathbf{R}^2$. If λ is a positive number and x is a vector in $\mathbf{R}^2$, then λx is the vector that points in the same direction as x and whose length is λ times the length of x. In other words, to get λx, we shrink or stretch x by a factor of λ, depending on whether $\lambda < 1$ or $\lambda > 1$. If λ is a negative number and x is a vector in $\mathbf{R}^2$, then λx is the vector that points in the direction opposite to that of x and whose length is $|\lambda|$ times the length of x.



1.3 Exercises

1. Suppose a and b are real numbers, not both 0. Find real numbers c and d such that
$$1/(a+bi) = c+di.$$
2. Show that $\frac{-1+\sqrt{3}i}{2}$ is a cube root of 1 (meaning that its cube equals 1).
3. Find two distinct square roots of i.
4. Show that $\alpha+\beta = \beta+\alpha$ for all $\alpha, \beta \in \mathbf{C}$.
5. Show that $(\alpha+\beta)+\lambda = \alpha+(\beta+\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.
6. Show that $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.
7. Show that for every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha+\beta = 0$.
8. Show that for every $\alpha \in \mathbf{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha\beta = 1$.
9. Show that $\lambda(\alpha+\beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.
10. Find $x \in \mathbf{R}^4$ such that
$$(4, -3, 1, 7) + 2x = (5, 9, -6, 8).$$
11. Explain why there does not exist $\lambda \in \mathbf{C}$ such that
$$\lambda(2-3i, 5+4i, -6+7i) = (12-5i, 7+22i, -32-9i).$$
12. Show that $(x+y)+z = x+(y+z)$ for all $x, y, z \in \mathbf{F}^n$.
13. Show that $(ab)x = a(bx)$ for all $x \in \mathbf{F}^n$ and all $a, b \in \mathbf{F}$.
14. Show that $1x = x$ for all $x \in \mathbf{F}^n$.
15. Show that $\lambda(x+y) = \lambda x + \lambda y$ for all $\lambda \in \mathbf{F}$ and all $x, y \in \mathbf{F}^n$.
16. Show that $(a+b)x = ax + bx$ for all $a, b \in \mathbf{F}$ and all $x \in \mathbf{F}^n$.

1.4 Vector spaces

The motivation for the definition of a vector space comes from properties of addition and scalar multiplication in $\mathbf{F}^n$: Addition is commutative, associative, and has an identity. Every element has an additive inverse. Scalar multiplication is associative. Scalar multiplication by 1 acts as expected. Addition and scalar multiplication are connected by distributive properties.

1.4.1 Definition of vector spaces

Definition 1.4.1 (addition, scalar multiplication)**.** We will define a vector space to be a set V with an addition and a scalar multiplication on V that satisfy the properties in the paragraph above.

- An **addition** on a set V is a function that assigns an element $u + v \in V$ to each pair of elements $u, v \in V$.
- A **scalar multiplication** on a set V is a function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbf{F}$ and each $v \in V$.

Definition 1.4.2 (vector space)**.** A **vector space** is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold:

commutativity $u + v = v + u$ for all $u, v \in V$;

associativity $(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and all $a, b \in \mathbf{F}$;

additive identity there exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$;

additive inverse for every $v \in V$, there exists $w \in V$ such that $v + w = 0$;

multiplicative identity $1v = v$ for all $v \in V$;

distributive properties $a(u+v) = au+av$ and $(a+b)v = av+bv$ for all $a, b \in \mathbf{F}$ and all $u, v \in V$.

Definition 1.4.3 (vector, point)**.** Elements of a vector space are called **vectors** or **points**.

Definition 1.4.4 (real vector space, complex vector space)**.** The scalar multiplication in a vector space depends on $\mathbf{F}$. Thus when we need to be precise, we will say that V is a vector space over $\mathbf{F}$ instead of saying simply that V is a vector space. For example, $\mathbf{R}^n$ is a vector space over $\mathbf{R}$, and $\mathbf{C}^n$ is a vector space over $\mathbf{C}$.

- A vector space over $\mathbf{R}$ is called a **real vector space**.
- A vector space over $\mathbf{C}$ is called a **complex vector space**.

Usually the choice of $\mathbf{F}$ is either obvious from the context or irrelevant. Thus we often assume that $\mathbf{F}$ is lurking in the background without specifically mentioning it.

Example 1.4.1. With the usual operations of addition and scalar multiplication, $\mathbf{F}^n$ is a vector space over $\mathbf{F}$.

Example 1.4.2. The simplest vector space contains only one point $\{0\}$.

Example 1.4.3 ($\mathbf{F}^\infty$)**.** $\mathbf{F}^\infty$ is defined to be the set of all sequences of elements of $\mathbf{F}$:

$$\mathbf{F}^\infty = \{(x_1, x_2, \dots) : x_j \in \mathbf{F} \text{ for } j = 1, 2, \dots\}.$$

Addition and scalar multiplication on $\mathbf{F}^\infty$ are defined as expected:

$$(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots)$$

$$\lambda(x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots)$$

With these definitions, $\mathbf{F}^\infty$ becomes a vector space over $\mathbf{F}$. The additive identity in this vector space is the sequence of all 0's.

1.4.2 $\mathbf{F}^S$

- If S is a set, then $\mathbf{F}^S$ denotes the set of functions from S to $\mathbf{F}$.
- For $f, g \in \mathbf{F}^S$, the **sum** $f + g \in \mathbf{F}^S$ is the function defined by

  $$(f + g)(x) = f(x) + g(x)$$

  for all $x \in S$.
- For $\lambda \in \mathbf{F}$ and $f \in \mathbf{F}^S$, the **product** $\lambda f \in \mathbf{F}^S$ is the function defined by

  $$(\lambda f)(x) = \lambda f(x)$$

  for all $x \in S$.
- If S is a nonempty set, then $\mathbf{F}^S$ (with the operations of addition and scalar multiplication as defined above) is a vector space over $\mathbf{F}$.
- The additive identity of $\mathbf{F}^S$ is the function $0 : S \to \mathbf{F}$ defined by

  $$0(x) = 0$$

  for all $x \in S$.
- For $f \in \mathbf{F}^S$, the additive inverse of f is the function $-f : S \to \mathbf{F}$ defined by

  $$(-f)(x) = -f(x)$$

  for all $x \in S$.
- Our previous examples of vector spaces, $\mathbf{F}^n$ and $\mathbf{F}^\infty$, are special cases of the vector space $\mathbf{F}^S$.

  Because a list of length n of numbers in $\mathbf{F}$ can be thought of as a function from $\{1, 2, \dots, n\}$ to $\mathbf{F}$ and a sequence of numbers in $\mathbf{F}$ can be thought of as a function from the set of positive integers to $\mathbf{F}$. In other words, we can think of $\mathbf{F}^n$ as $\mathbf{F}^{\{1,2,\dots,n\}}$ and we can think of $\mathbf{F}^\infty$ as $\mathbf{F}^{\{1,2,\dots\}}$.

1.4.3 Elementary properties of vector spaces

Proposition 1.4.1 (Unique additive identity)**.** A vector space has a unique additive identity.

Proof. Suppose 0 and $0'$ are both additive identities for some vector space V. Then

$$0' = 0' + 0 = 0 + 0' = 0.$$

$\square$

Proposition 1.4.2 (Unique additive inverse)**.** Every element in a vector space has a unique additive inverse.

Proof. Suppose V is a vector space. Let $v \in V$. Suppose w and w' are additive inverses of v. Then

$$w = w + 0 = w + (v + w') = (w + v) + w' = 0 + w' = w'.$$

$\square$

Because additive inverses are unique, the following notation now makes sense: Let $v, w \in V$. Then

- $-v$ denotes the additive inverse of v;
- $w - v$ is defined to be $w + (-v)$.

For the rest of the book, V denotes a vector space over $\mathbf{F}$.

Proposition 1.4.3 (The number 0 times a vector)**.** $0v = 0$ for every $v \in V$.

Proof. For $v \in V$, we have

$$0v = (0 + 0)v = 0v + 0v.$$

Adding the additive inverse of $0v$ to both sides of the equation above gives $0 = 0v$, as desired. $\square$

Proposition 1.4.4 (A number times the vector 0)**.** $a0 = 0$ for every $a \in \mathbf{F}$.

Proof. For $a \in \mathbf{F}$, we have

$$a0 = a(0 + 0) = a0 + a0.$$

Adding the additive inverse of $a0$ to both sides of the equation above gives $0 = a0$, as desired. $\square$

Proposition 1.4.5 (The number -1 times a vector)**.** $(-1)v = -v$ for every $v \in V$.

Proof. For $v \in V$, we have

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0.$$

Thus $(-1)v$ is the additive inverse of v, as desired. $\square$

1.5 Exercises

1. Prove that $-(-v) = v$ for every $v \in V$.
2. Suppose $a \in \mathbf{F}, v \in V$, and $av = 0$. Prove that $a = 0$ or $v = 0$.
3. Suppose $v, w \in V$. Explain why there exists a unique $x \in V$ such that $v + 3x = w$.
4. The empty set is not a vector space. Why?
5. Show that in the definition of a vector space, the additive inverse condition can be replaced with the condition that
$$0v = 0 \text{ for all } v \in V.$$
6. Let ∞ and $-\infty$ denote two distinct objects, neither of which is in $\mathbf{R}$. Define an addition and scalar multiplication on $\mathbf{R} \cup \{\infty\} \cup \{-\infty\}$ as you could guess from the notation. Is $\mathbf{R} \cup \{\infty\} \cup \{-\infty\}$ a vector space over $\mathbf{R}$? Explain.

1.6 Subspaces

1.6.1 Definition of subspaces

Definition 1.6.1 (subspace)**.** A subset U of V is called a **subspace** of V if U is also a vector space using the same addition and scalar multiplication as on V.

Example 1.6.1. $\{(x_1, x_2, 0) : x_1, x_2 \in \mathbf{F}\}$ is a subspace of $\mathbf{F}^3$.

Proposition 1.6.1 (Conditions for a subspace)**.** A subset U of V is a subspace of V if and only if U satisfies the following three conditions:

additive identity $0 \in U$

closed under addition $u, w \in U$ implies $u + w \in U$;

scalar multiplication $a \in \mathbf{F}$ and $u \in U$ implies $au \in U$.

Proof. If U is a subspace of V, then U satisfies the three conditions above by the definition of vector space.

Conversely, suppose U satisfies the three conditions above. The first condition above ensures that the additive identity of V is in U. The second condition above ensures that addition makes sense on U. The third condition ensures that scalar multiplication makes sense on U. If $u \in U$, then $-u = (-1)u$ is also in U by the third condition above. Hence every element of U has an additive inverse in U.

The other parts of the definition of a vector space, such as associativity and commutativity, are **automatically** satisfied for U because they hold on the larger space V. Thus U is a vector space and hence is a subspace of V. $\square$

Example 1.6.2. If $b \in \mathbf{F}$, then

$$\{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 = 5x_4 + b\}$$

is a subspace of $\mathbf{F}^4$ if and only if $b = 0$.

Clearly $\{0\}$ is the smallest subspace of V and V itself is the largest subspace of V. The subspaces of $\mathbf{R}^2$ are precisely $\{0\}, \mathbf{R}^2$, and all lines in $\mathbf{R}^2$ through the origin. The subspaces of $\mathbf{R}^3$ are precisely $\{0\}, \mathbf{R}^3$, all lines in $\mathbf{R}^3$ through the origin, and all planes in $\mathbf{R}^3$ through the origin.

The hard part is to show that they are the only subspaces of $\mathbf{R}^2$ and $\mathbf{R}^3$.

1.6.2 Sums of subspaces

Definition 1.6.2 (sum of subsets)**.** Suppose $U_1, \ldots, U_m$ are subsets of V. The **sum** of $U_1, \ldots, U_m$, denoted $U_1 + \cdots + U_m$, is the set of all possible sums of elements of $U_1, \ldots, U_m$. More precisely,

$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m : u_1 \in U_1, \ldots, u_m \in U_m\}.$$

Example 1.6.3.

$$U = \{(x, 0, 0) \in \mathbf{F}^3 : x \in \mathbf{F}\}, W = \{(0, y, 0) \in \mathbf{F}^3 : y \in \mathbf{F}\}.$$

Then

$$U + W = \{(x, y, 0) : x, y \in \mathbf{F}\},$$

as you should verify.

Example 1.6.4. Suppose that $U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}$ and $W = \{(x, x, x, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}$. Then

$$U + W = \{(x, x, y, z) \in \mathbf{F}^4 : x, y, z \in \mathbf{F}\},$$

as you should verify.

Proposition 1.6.2 (Sum of subspaces is the smallest containing subspace)**.** Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.

Proof. It is easy to see that $0 \in U_1 + \cdots + U_m$ and that $U_1 + \cdots + U_m$ is closed under addition and scalar multiplication. Thus $U_1 + \cdots + U_m$ **is** a subspace of V.

Clearly $U_1, \ldots, U_m$ are all contained in $U_1 + \cdots + U_m$. Conversely, every subspace of V containing $U_1, \ldots, U_m$ contains $U_1 + \cdots + U_m$ because subspaces must contain all finite sums of their elements. Thus $U_1 + \cdots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$. $\square$

1.6.3 Direct sums

Suppose $U_1, \ldots, U_m$ are subspaces of V. Every element of $U_1 + \cdots + U_m$ can be written in the form

$$u_1 + \cdots + u_m,$$

where each u_j is in U_j.

Definition 1.6.3 (direct sum)**.** Suppose $U_1, \ldots, U_m$ are subspaces of V.

- The sum $U_1 + \cdots + U_m$ is called a **direct sum** if each element of $U_1 + \cdots + U_m$ can be written in only one way as a sum $u_1 + \cdots + u_m$, where each u_j is in U_j.
- If $U_1 + \cdots + U_m$ is a direct sum, then $U_1 \oplus \cdots \oplus U_m$ denote $U_1 + \cdots + U_m$, with the $\oplus$ notation serving as an indication that this is a direct sum.

Example 1.6.5.

$$U = \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}, W = \{(0, 0, z) \in \mathbf{F}^3 : z \in \mathbf{F}\}.$$

Then $\mathbf{F}^3 = U \oplus W$, as you should verify.

Example 1.6.6. Suppose U_j is the subspace of $\mathbf{F}^n$ of those vectors whose coordinates are all 0, except possibly in the j^{th} slot (thus, for example, $U_2 = \{(0, x, 0, \ldots, 0) \in \mathbf{F}^n : x \in \mathbf{F}\}$). Then

$$\mathbf{F}^n = U_1 \oplus \cdots \oplus U_n,$$

as you should verify.

Example 1.6.7. Let

$$\begin{aligned} U_1 &= \{(x, y, 0) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}, \\ U_2 &= \{(0, 0, z) \in \mathbf{F}^3 : z \in \mathbf{F}\}, \\ U_3 &= \{(0, y, y) \in \mathbf{F}^3 : y \in \mathbf{F}\}. \end{aligned}$$

Show that $U_1 + U_2 + U_3$ is not a direct sum.

Proof. Clearly $(0, 0, 0) \in U_1 + U_2 + U_3$, but we have

$$(0, 0, 0) + (0, 0, 0) + (0, 0, 0) = (0, 1, 0) + (0, 0, 1) + (0, -1, -1).$$

$\square$

Proposition 1.6.3 (Condition for a direct sum)**.** Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is a direct sum if and only if the only way to write 0 as a sum $u_1 + \cdots + u_m$, where each u_j is in U_j, is by taking each u_j equal to 0.

Proof. To show that $U_1 + \cdots + U_m$ is a direct sum, let $v \in U_1 + \cdots + U_m$. We can write

$$v = u_1 + \cdots + u_m$$

for some $u_1 \in U_1, \ldots, u_m \in U_m$. To show that this representation is unique, suppose we also have

$$v = v_1 + \cdots + v_m,$$

where $v_1 \in U_1, \ldots, v_m \in U_m$. Subtracting these two equations, we have

$$0 = (u_1 - v_1) + \cdots + (u_m - v_m).$$

because $u_1 - v_1 \in U_1, \ldots, u_m - v_m \in U_m$, the equation above implies that each $u_j - v_j$ equals 0. Thus $u_1 = v_1, \ldots, u_m = v_m$, as desired. □

Proposition 1.6.4 (Direct sum of two subspaces)**.** Suppose U and W are subspaces of V. Then $U + V$ is a direct sum if and only if $U \cap W = \{0\}$.

Proof. First suppose that $U+W$ is a direct sum. If $v \in U \cap W$, then $0 = v + (-v)$, where $v \in U$ and $-v \in W$. By the unique representation of 0 as the sum of a vector in U and a vector in W, we have $v = 0$. Thus $U \cap W = \{0\}$, completing the proof in one direction.

To prove the other direction, now suppose $U \cap W = \{0\}$. To prove that $U + W$ is a direct sum, suppose $u \in U, w \in W$, and

$$0 = u + v.$$

To complete the proof, we need only show that $u = w = 0$. The equation above implies that $u = -v \in W$. Thus $u \in U \cap W$. Hence $u = 0$, which by the equation above implies that $w = 0$, completing the proof.

The result above deals only with the case of two subspaces. When asking about a possible direct sum with more than two subspaces, it is not enough to test that each pair of the subspaces intersect only at 0. □

1.7 Exercises

1. For each of the following subsets of $\mathbf{F}^3$, determine whether it is a subspace of $\mathbf{F}^3$:

   (a) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$;
   (b) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$;
   (c) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}$;
   (d) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$.

2. Is $\mathbf{R}^2$ a subspace of the complex vector space $\mathbf{C}^2$?

3. (a) Is $\{(a, b, c) \in \mathbf{R}^3 : a^3 = b^3\}$ a subspace of $\mathbf{R}^3$?
   (b) Is $\{(a, b, c) \in \mathbf{C}^3 : a^3 = b^3\}$ a subspace of $\mathbf{C}^3$?

4. Give an example of a nonempty subset U of $\mathbf{R}^2$ such that U is closed under addition and under taking additive inverses, but U is not a subspace of $\mathbf{R}^2$.
5. Give an example of a nonempty subset U of $\mathbf{R}^2$ such that U is closed under scalar multiplication, but U is not a subspace of $\mathbf{R}^2$.
6. A function $f : \mathbf{R} \to \mathbf{R}$ is called **periodic** if there exists a positive number p such that $f(x) = f(x+p)$ for all $x \in \mathbf{R}$. Is the set of periodic functions from $\mathbf{R}$ to $\mathbf{R}$ a subspace of $\mathbf{R}^{\mathbf{R}}$? Explain.
7. Suppose U_1 and U_2 are subspaces of V. Prove that the intersection $U_1 \cap U_2$ is a subspace of V.
8. Prove that the intersection of every collection of subspaces of V is a subspace of V.
9. Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.
10. Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.
11. Suppose U is a subspace of V. What is $U + U$?
12. Is the operation of addition on the subspaces of V commutative? In other words, if U and W are subspaces of V, is $U + W = W + U$?
13. Is the operation of addition on the subspaces of V associative? In other words, if U_1, U_2, U_3 are subspaces of V, is
$$(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3)?$$
14. Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?
15. Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that
$$U_1 + W = U_2 + W,$$
then $U_1 = U_2$.
16. Suppose
$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}.$$
Find a subspace W of $\mathbf{F}^4$ such that $\mathbf{F}^4 = U \oplus W$.
17. Suppose
$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$
Find a subspace W of $\mathbf{F}^5$ such that $\mathbf{F}^5 = U \oplus W$.
18. Suppose
$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$
Find three subspaces W_1, W_2, W_3 of $\mathbf{F}^5$, none of which equals $\{0\}$, such that $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

19. Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that

$$V = U_1 \oplus W \text{ and } V = U_2 \oplus W,$$

then $U_1 = U_2$.

20. A function $f : \mathbf{R} \to \mathbf{R}$ is called **even** if

$$f(-x) = f(x)$$

for all $x \in \mathbf{R}$. A function $f : \mathbf{R} \to \mathbf{R}$ is called **odd** if

$$f(-x) = -f(x)$$

for all $x \in \mathbf{R}$.

Let U_e denote the set of real-valued even functions on $\mathbf{R}$ and let U_o denote the set of real-valued odd functions on $\mathbf{R}$. Show that $\mathbf{R}^{\mathbf{R}} = U_e \oplus U_o$.

Chapter 2

Finite-Dimensional Vector Spaces

- **R** denotes **R** or **C**.
- V denotes a vector space over **F**.

2.1 Span and linear independence

2.1.1 Linear combinations and span

Definition 2.1.1 (linear combination)**.** Adding up scalar multiples of vectors in a list gives what is called a linear combination of the list.

A **linear combination** of a list $v_1, \ldots, v_m$ of vectors in V is a vector of the form

$$a_1 v_1 + \cdots + a_m v_m,$$

where $a_1, \ldots, a_m \in \mathbf{F}$.

Example 2.1.1. In $\mathbf{F}^3$,

- $(17, -4, 2)$ is a linear combination of $(2, 1, -3), (1, -2, 4)$ because

$$(17, -4, 2) = 6(2, 1, -3) + 5(1, -2, 4).$$

- $(17, -4, 5)$ is not a linear combination of $(2, 1, -3), (1, -2, 4)$ because there do not exists numbers $a_1, a_2 \in \mathbf{F}$ such that

$$(17, -4, 5) = a_1(2, 1, -3) + a_2(1, -2, 4).$$

In other words, the system of equations

$$\begin{aligned} 17 &= 2a_1 + a_2 \\ -4 &= a_1 - 2a_2 \\ 5 &= -3a_1 + 4a_2 \end{aligned}$$

has no solutions (as you should verify).

Definition 2.1.2 (span)**.** The set of all linear combinations of a list of vectors $v_1, \ldots, v_m$ in V is called the **span** of $v_1, \ldots, v_m$, denoted $\operatorname{span}(v_1, \ldots, v_m)$. In other words,

$$\operatorname{span}(v_1, \ldots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \ldots, a_m \in \mathbf{F}\}.$$

The span of the empty list () is defined to be $\{0\}$.

Example 2.1.2. The previous example shows that in $\mathbf{F}^3$,

- $(17, -4, 2) \in \operatorname{span}((2, 1, -3), (1, -2, 4))$
- $(17, -4, 5) \notin \operatorname{span}((2, 1, -3), (1, -2, 4))$

Proposition 2.1.1 (Span is the smallest containing subspace)**.** The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.

Proof. Omitted. □

Definition 2.1.3 (spans)**.** If $\operatorname{span}(v_1, \ldots, v_m)$ equals V, we say that $v_1, \ldots, v_m$ **spans** V.

Example 2.1.3. Suppose n is a positive integer. Show that

$$(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), (0, \ldots, 0, 1)$$

spans $\mathbf{F}^n$. Here the j^{th} vector in the list above is the n-tuple with 1 in the j^{th} slot and 0 in all other slots.

Proof.

$$(x_1, \ldots, x_n) = x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \cdots + x_n(0, \ldots, 0, 1).$$

□

Definition 2.1.4 (finite-dimensional vector space, infinite-dimensional vector space)**.** Recall that by definition every list has finite length.

A vector space is called **finite-dimensional** if some list of vectors in it spans the space. A vector space is called **infinite-dimensional** if it is not finite-dimensional.

Example 2.1.4. $\mathbf{F}^n$ is a finite-dimensional vector space for every positive integer n.

Proof. Example 2.1.3. □

2.1.2 Polynomials

Definition 2.1.5 (polynomial, $\mathcal{P}(\mathbf{F})$)**.** The definition of a polynomial is no doubt already familiar to you.

- A function $p : \mathbf{F} \to \mathbf{F}$ is called a **polynomial** with coefficients in $\mathbf{F}$ is there exists $a_0, \ldots, a_m \in \mathbf{F}$ such that

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

  for all $z \in \mathbf{F}$.
- $\mathcal{P}(\mathbf{F})$ is the set of all polynomials with coefficients in $\mathbf{F}$.
- With the usual operations of addition and scalar multiplication, $\mathcal{P}(\mathbf{F})$ is a vector space over $\mathbf{F}$, as you should verify.
- $\mathcal{P}(\mathbf{F})$ is a subspace of $\mathbf{F}^{\mathbf{F}}$.

If a polynomial (thought of as a function from $\mathbf{F}$ to $\mathbf{F}$) is represented by two sets of coefficients, then subtracting one representation of the polynomial from the other produces a polynomial that is identically zero as a function on $\mathbf{F}$ and hence has all zero coefficients. **Conclusion: the coefficients of a polynomial are uniquely determined by the polynomial.**

This is a fact. We will prove it later.

Definition 2.1.6 (degree of a polhnomial, $\deg p$)**.** The coefficients of a polynomial are uniquely determined by the polynomial!

- A polynomial $p \in \mathcal{P}(\mathbf{F})$ is said to have **degree** m if there exists scalars $a_0, a_1, \ldots, a_m \in \mathbf{F}$ with $a_m \neq 0$ such that

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

  for all $z \in \mathbf{F}$. If p has degree m, we write $\deg p = m$.
- The polynomial that is identically 0 is said to have degree $-\infty$. We use the convention that $-\infty < m$.

Definition 2.1.7 ($\mathcal{P}_m(\mathbf{F})$)**.** For m a nonnegative integer, $\mathcal{P}_m(\mathbf{F})$ denotes the set of all polynomials with coefficients in $\mathbf{F}$ and degree at most m.

Note that $\mathcal{P}_m(\mathbf{F}) = \operatorname{span}(1, z, \ldots, z^m)$, $\mathcal{P}_m(\mathbf{F})$ is a finite-dimensional vector space for each nonnegative integer m.

Example 2.1.5. Show that $\mathcal{P}(\mathbf{F})$ is infinite-dimensional.

Proof. Consider any list of elements of $\mathcal{P}(\mathbf{F})$. Let m denote the highest degree of the polynomials in this list. Then every polynomial in the span of this list has degree at most m. Thus z^{m+1} is not in the span of our list. Hence no list spans $\mathcal{P}(\mathbf{F})$. Thus $\mathcal{P}(\mathbf{F})$ is infinite-dimensional. □

2.1.3 Linear independence

Suppose $v_1, \ldots, v_m \in V$ and $v \in \operatorname{span}(v_1, \ldots, v_m)$. By the definition of span, there exists $a_1, \ldots, a_m \in \mathbf{F}$ such that

$$v = a_1 v_1 + \cdots + a_m v_m.$$

Consider the question of whether the choice of scalars in the equation above is unique. Suppose $c_1, \ldots, c_m$ is another set of scalars such that

$$v = c_1 v_1 + \cdots + c_m v_m.$$

Subtracting the last two equations, we have

$$0 = (a_1 - c_1)v_1 + \cdots + (a_m - c_m)v_m.$$

Thus we have written 0 as a linear combination of $(v_1, \ldots, v_m)$. If the only way to do this is the obvious way (using 0 for all scalars), then each $a_j - c_j$ equals 0, which means that each a_j equals c_j (and thus the choice of scalars was indeed unique).

Definition 2.1.8 (linearly independent)**.** This situation (the choice of scalars was indeed unique) is so important that we give it a special name:

- A list $v_1, \ldots, v_m$ of vectors in V is called **linearly independent** if the only choice of $a_1, \ldots, a_m \in \mathbf{F}$ that makes $a_1 v_1 + \cdots + a_m v_m$ equals 0 is $a_1 = \cdots = a_m = 0$.
- The empty list () is also declared to be linearly independent.
- $v_1, \ldots, v_m$ is linearly independent if and only if each vector in $\operatorname{span}(v_1, \ldots, v_m)$ has only one representation as a linear combination of $v_1, \ldots, v_m$.

Example 2.1.6. 1. A list v of one vector $v \in V$ is linearly independent if and only if $v \neq 0$.

2. A list of two vectors in V is linearly independent if and only if neither vector is a scalar multiple of the other.
3. $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$ is linearly independent in $\mathbf{F}^4$.
4. The list $1, z, \ldots, z^m$ is linearly independent in $\mathcal{P}(\mathbf{F})$ for each nonnegative integer m.
5. If some vectors are removed from a linearly independent list, the remaining list is also linearly independent, as you should verify.

Definition 2.1.9 (linearly dependent)**.** A list of vectors in V is called **linearly dependent** if it is not linearly independent. In other words, a list $v_1, \ldots, v_m$ of vectors in V is linearly dependent if there exists $a_1, \ldots, a_m \in \mathbf{R}$, not all 0, such that $a_1 v_1 + \cdots + a_m v_m = 0$.

Example 2.1.7. • $(2, 3, 1), (1, -1, 2), (7, 3, 8)$ is linearly dependent in $\mathbf{F}^3$ because

$$2(2, 3, 1) + 3(1, -1, 2) + (-1)(7, 3, 8) = (0, 0, 0).$$

- The list $(2, 3, 1), (1, -1, 2), (7, 3, c)$ is linearly dependent in $\mathbf{F}^3$ if and only if $c = 8$, as you should verify.
- If some vector in a list of vectors in V is a linear combination of the other vectors, then the list is linearly dependent.
- Every list of vectors in V containing the 0 vector is linearly dependent.

Lemma 2.1.1 (Linear Dependence Lemma)**.** Suppose $v_1, \ldots, v_m$ is a linearly dependent list in V. Then there exists $j \in \{1, 2, \ldots, m\}$ such that the following hold:

1. $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$;
2. if the j^{th} term is removed from $v_1, \ldots, v_m$, the span of the remaining list equals $\operatorname{span}(v_1, \ldots, v_m)$.

Proof. Because the list $v_1, \ldots, v_m$ is linearly dependent, there exist numbers $a_1, \ldots, a_m \in \mathbf{F}$, not all 0, such that

$$a_1 v_1 + \cdots + a_m v_m = 0$$

Let j be the largest element of $\{1, \ldots, m\}$ such that $a_j \neq 0$. Then

$$v_j = -\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1},$$

proving 1. To prove 2, suppose $u \in \operatorname{span}(v_1, \ldots, v_m)$. Then there exist numbers $c_1, \ldots, c_m \in \mathbf{F}$ such that

$$u = c_1 v_1 + \cdots + c_m v_m.$$

In the equation above, we can replace v_j with the right side, which shows that u is in the span of the list obtained by removing the j^{th} term from $v_1, \ldots, v_m$. Thus 2 holds. $\square$

Choosing $j = 1$ in the Linear Dependence Lemma above means that $v_1 = 0$, because if $j = 1$ then condition 1 above is interpreted to mean that $v_1 \in \operatorname{span}()$; recall that $\operatorname{span}() = \{0\}$. Note also that the proof of part 2 above needs to be modified in an obvious way if $v_1 = 0$ and $j = 1$.

Now we come to a key result. It says that no linearly independent list in V is longer than a spanning list in V.

Proposition 2.1.2 (Length of linearly independent list $\leq$ length of spanning list)**.** In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Proof. Suppose $u_1, \ldots, u_m$ is linearly independent in V. Suppose also that $w_1, \ldots, w_n$ spans V. We need to prove that $m \leq n$. We de so through the multi-step process described below; note that in each step we add one of the u's and remove one of the w's.

Step 1 Let B be the list $w_1, \dots, w_n$, which spans V. Thus adjoining any vector in V to this list produces a linearly dependent list (because the newly adjoined vector can be written as a linear combination of the other vectors). In particular, the list

$$u_1, w_1, \dots, w_n$$

is linearly dependent. Thus by the Linear Dependence Lemma, we can remove one of the w's so that the new list B (of length n) consisting of u_1 and the remaining w's spans V.

Step j The list B (of length n) from step $j-1$ spans V. Thus adjoining any vector to this list produces a linearly dependent list. In particular, the list of length $(n+1)$ obtained by adjoining u_j to B, placing it just after $u_1, \dots, u_{j-1}$, is linearly dependent. By the Linear Dependence Lemma, one of the vectors in this list is in the span of the previous ones, and because $u_1, \dots, u_j$ is linearly independent, this vector is one of the w's, not one of the u's. We can remove that w from B so that the new list B (of length n) consisting of $u_1, \dots, u_j$ and the remaining w's spans V.

After step m, we have added all the u's and the process stops. At each step as we add a u to B, the Linear Dependence Lemma implies that there is some w to remove. Thus there are at least as many w's as u's. □

Example 2.1.8. Show that the list $(1,2,3),(4,5,8),(9,6,7),(-3,2,8)$ is not linearly independent in $\mathbf{R}^3$.

Proof. The list $(1,0,0),(0,1,0),(0,0,1)$ spans $\mathbf{R}^3$. Thus no list of length larger than 3 is linearly independent in $\mathbf{R}^3$. □

Example 2.1.9. Show that the list $(1,2,3,-5),(4,5,8,3),(9,6,7,-1)$ does not span $\mathbf{R}^4$.

Proof. The list $(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)$ is linearly independent in $\mathbf{R}^4$. Thus no list of length less than 4 spans $\mathbf{R}^4$. □

Our intuition suggests that every subspace of a finite-dimensional vector space should also be finite-dimensional. We now prove that this intuition is correct.

Proposition 2.1.3 (Finite-dimensional subspaces)**.** Every subspace of a finite-dimensional vector space is finite-dimensional.

Proof. Suppose V is finite-dimensional and U is a subspace of V. We need to prove that U is finite-dimensional. We do this through the following multi-step construction.

Step 1 If $U = \{0\}$, then U is finite-dimensional and we are done. If $U \neq \{0\}$, then choose a nonzero vector $v_1 \in U$.

Step j If $U = \operatorname{span}(v_1, \ldots, v_{j-1})$, then U is finite-dimensional and we are done. If $U \neq \operatorname{span}(v_1, \ldots, v_{j-1})$, then choose a vector $v_j \in U$ such that

$$v_j \in \operatorname{span}(v_1, \ldots, v_{j-1}).$$

After each step, as long as the process continues, we have constructed a list of vectors such that no vector in this list is in the span of the previous vectors. Thus after each step we have constructed a linearly independent list, by the Linear Dependence Lemma. This linearly independent list cannot be longer than any spanning list of V. Thus the process eventually terminates, which means that U is finite-dimensional. $\square$

2.2 Exercises

1. Suppose v_1, v_2, v_3, v_4 spans V. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also spans V.

2. Find a number t such that

$$(3, 1, 4), (2, -3, 5), (5, 9, t)$$

is not linearly independent in $\mathbf{R}^3$.

3. (a) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{R}$, then the list $(1 + i, 1 - i)$ is linearly independent.
   (b) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{C}$, then the list $(1 + i, 1 - i)$ is linearly dependent.

4. Suppose v_1, v_2, v_3, v_4 is linearly independent in V. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

is also linearly independent.

5. Prove or give a counterexample: If $v_1, v_2, \ldots, v_m$ is a linearly independent list of vectors in V, then

$$5v_1 - 4v_2, v_2, v_3, \ldots, v_m$$

is linearly independent.

6. Prove or give a counterexample: If $v_1, v_2, \ldots, v_m$ is a linearly independent list of vectors in V and $\lambda \in \mathbf{F}$ with $\lambda \neq 0$, then $\lambda v_1, \lambda v_2, \ldots, \lambda v_m$ is linearly independent.

7. Prove or give a counterexample: If $v_1, \ldots, v_m$ and $w_1, \ldots, w_m$ are linearly independent lists of vectors in V, then $v_1 + w_1, \ldots, v_m + w_m$ is linearly independent.

8. Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that if $v_1 + w, \ldots, v_m + w$ is linearly dependent, then $w \in \operatorname{span}(v_1, \ldots, v_m)$.
9. Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Show that $v_1, \ldots, v_m, w$ is linearly independent if and only if
$$w \notin \operatorname{span}(v_1, \ldots, v_m).$$
10. Explain why there does not exist a list of six polynomials that is linearly independent in $\mathcal{P}_4(\mathbf{F})$.
11. Explain why no list of four polynomials spans $\mathcal{P}_4(\mathbf{F})$.
12. Prove that V is infinite-dimensional if and only if there is a sequence $v_1, v_2, \ldots$ of vectors in V such that $v_1, \ldots, v_m$ is linearly independent for every positive integer m.
13. Prove that $\mathbf{F}^\infty$ is infinite-dimensional.
14. Prove that the real vector space of all continuous real-valued functions on the interval $[0, 1]$ is infinite-dimensional.
15. Suppose $p_0, p_1, \ldots, p_m$ are polynomials in $\mathcal{P}_m(\mathbf{F})$ such that $p_j(2) = 0$ for each j. Prove that $p_0, p_1, \ldots, p_m$ is not linearly independent in $\mathcal{P}_m(\mathbf{F})$.

2.3 Bases

Definition 2.3.1 (basis)**.** A **basis** of V is a list of vectors in V that is linearly independent and spans V.

Example 2.3.1.

1. The list $(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$ is a basis of $\mathbf{F}^n$, called the **standard basis** of $\mathbf{F}^n$. In addition to the standard basis, $\mathbf{F}^n$ has many other bases.
2. The list $(1, 2), (3, 5)$ is a basis of $\mathbf{F}^2$.
3. The list $(1, 2, -4), (7, -5, 6)$ is linearly independent in $\mathbf{F}^3$ but is not a basis of $\mathbf{F}^3$ because it does not span $\mathbf{F}^3$.
4. The list $(1, 2), (3, 5), (4, 13)$ spans $\mathbf{F}^2$ but is not a basis of $\mathbf{F}^2$ because it is not linearly independent.
5. The list $(1, 1, 0), (0, 0, 1)$ is a basis of $\{(x, x, y) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}$.
6. The list $(1, -1, 0), (1, 0, -1)$ is a basis of
$$\{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}.$$
7. The list $1, z, \ldots, z^m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

Proposition 2.3.1 (Criterion for basis)**.** A list $v_1, \ldots, v_n$ of vectors in V is a basis of V if and only if every $v \in V$ **can** be written **uniquely** in the form
$$v = a_1 v_1 + \cdots a_n v_n,$$
where $a_1, \ldots, a_n \in \mathbf{F}$.

Proof. Omitted. □

A spanning list in a vector space may not be a basis because it is not linearly independent. Our next result says that given any spanning list, some (possibly none) of the vectors in it can be discarded so that the remaining list is linearly independent and still spans the vector space.

Proposition 2.3.2 (Spanning list contains a basis)**.** Every spanning list in a vector space can be reduced to a basis of the vector space.

Proof. Suppose $v_1, \ldots, v_n$ spans V. We want to remove some of the vectors from $v_1, \ldots, v_n$ so that the remaining vectors form a basis of V. We do this through the multi-step process described below.

Start with B equal to the list $v_1, \ldots, v_n$.

Step 1 If $v_1 = 0$, delete v_1 from B. If $v_1 \neq 0$, leave B unchanged.

Step j If v_j is in $\operatorname{span}(v_1, \ldots, v_{j-1})$, delete v_j from B. If v_j is not in $\operatorname{span}(v_1, \ldots, v_{j-1})$, leave B unchanged.

Stop the process after step n, getting a list B. This list B spans V because our original list spanned V and we have discarded only vectors that were already in the span of the previous vectors. The process ensures that no vector in B is in the span of the previous ones. Thus B is linearly independent, by the Linear Dependence Lemma. Hence B is a basis of V. □

Proposition 2.3.3 (Basis of finite-dimensional vector space)**.** Every finite-dimensional vector space has a basis.

Proof. By definition, a finite-dimensional vector space has a spanning list. The previous result tells us that each spanning list can be reduced to a basis. □

Now we show that given any linearly independent list, we can adjoin some additional vectors (this includes the possibility of adjoining no additional vectors) so that the extended list is still linearly independent but also spans the space.

Proposition 2.3.4 (Linearly independent list extends to a basis)**.** Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

Proof. Suppose $u_1, \ldots, u_m$ is linearly independent in a finite-dimensional vector space V. Let $w_1, \ldots, w_n$ be a basis of V (Because every finite-dimensional vector space has a basis). Thus the list

$$u_1, \ldots, u_m, w_1, \ldots, w_n$$

spans V. Since every spanning list in a vector space can be reduced to a basis of the vector space. We can reduce this list to a basis of V consisting of the vectors $u_1, \ldots, u_m$. (None of the u's get deleted in this procedure because $u_1, \ldots, u_m$ is linearly independent.) □

Proposition 2.3.5 (Every subspace of V is part of a direct sum equal to V)**.** Suppose V is finite-dimensional and U is a subspace of V. Then there is a subspace W of V such that $V = U \oplus W$.

Proof. Because V is finite-dimensional, so is U. Thus there is a basis $u_1, \ldots, u_m$ of U, which is a linearly independent list of vectors in V. Hence this list can be extended to a basis $u_1, \ldots, u_m, w_1, \ldots, w_n$ of V. Let $W = \operatorname{span}(w_1, \ldots, w_n)$. $\square$

2.4 Exercises

1. Find all vector spaces that have exactly one basis.
2. (a) Let U be the subspace of $\mathbf{R}^5$ defined by

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2, x_3 = 7x_4\}.$$

   Find a basis of U.

   (b) Extend the basis in part (a) to a basis of $\mathbf{R}^5$.

   (c) Find a subspace W of $\mathbf{R}^5$ such that $\mathbf{R}^5 = U \oplus W$.
3. (a) Let U be the subspace of $\mathbf{C}^5$ defined by

$$U = \{(z_1, z_2, z_3, z_4, z_5) \in \mathbf{C}^5 : 6z_1 = z_2, z_3 + 2z_4 + 3z_5 = 0\}.$$

   Find a basis of U.

   (b) Extend the basis in part (a) to a basis of $\mathbf{C}^5$.

   (c) Find a subspace W of $\mathbf{C}^5$ such that $\mathbf{C}^5 = U \oplus W$.
4. Prove or disprove: there exists a basis p_0, p_1, p_2, p_3 of $\mathcal{P}_3(\mathbf{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2.
5. Suppose v_1, v_2, v_3, v_4 is a basis of V. Prove that

$$v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$$

   is also a basis of V.
6. Prove or give a counterexample: If v_1, v_2, v_3, v_4 is a basis of V and U is a subspace of V such that $v_1, v_2 \in U$ and $v_3 \notin U$ and $v_4 \notin U$, then v_1, v_2 is a basis of U.
7. Suppose U and W are subspaces of V such that $V = U \oplus W$. Suppose also that $u_1, \ldots, u_m$ is a basis of U and $w_1, \ldots, w_n$ is a basis of W. Prove that

$$u_1, \ldots, u_m, w_1, \ldots, w_n$$

   is a basis of V.

2.5 Dimension

Although we have been discussing finite-dimensional vector spaces, we have not yet defined the dimension of such an object. How should dimension be defined? A reasonable definition should force the dimension of $\mathbf{F}^n$ to equal n. Notice that the standard basis

$$(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$$

of $\mathbf{F}^n$ has length n. Thus we are tempted to define the dimension as the length of a basis. However, a finite-dimensional vector space in general has many different bases, and our attempted definition makes sense only if all bases in a given vector space have the same length. Fortunately that turns out to be the case, as we now show.

Proposition 2.5.1 (Basis length does not depend on basis)**.** Any two bases of a finite-dimensional vector space have the same length.

Proof. Suppose V is finite-dimensional. Let B_1 and B_2 be two bases of V. Then B_1 is linearly independent in V and B_2 spans V, so the length of B_1 is at most the length of B_2. Interchanging the roles of B_1 and B_2, we also see that the length of B_2 is at most the length of B_1. Thus the length of B_1 equals the length of B_2, as desired. □

Definition 2.5.1 (dimension, $\dim V$)**.** The **dimension** of a finite-dimensional vector space is the length of any basis of the vector space. The dimension of V (if V is finite-dimensional) is denoted by $\dim V$.

Example 2.5.1.

- $\dim \mathbf{F}^n = n$ because the standard basis of $\mathbf{F}^n$ has length n.
- $\dim \mathcal{P}_m(\mathbf{F}) = m + 1$ because the basis $1, z, \ldots, z^m$ of $\mathcal{P}_m(\mathbf{F})$ has length $m + 1$.

Proposition 2.5.2 (Dimension of a subspace)**.** If V is finite-dimensional and U is a subspace of V, then $\dim U \leq \dim V$.

Proof. Hint: Suppose V is finite-dimensional and U is a subspace of V. Think of a basis of U as a linearly independent list in V, and think of a basis of V as a spanning list in V. □

The real vector space $\mathbf{R}^2$ has dimension 2; the complex vector space $\mathbf{C}$ has dimension 1. As sets, $\mathbf{R}^2$ can be identified with $\mathbf{C}$ (and addition is the same on both spaces, as is scalar multiplication by real numbers). Thus when we talk about the dimension of a vector space, the role played by the choice of $\mathbf{F}$ cannot be neglected.

Proposition 2.5.3 (Linearly independent list of the right length is a basis)**.** Suppose V is finite-dimensional. Then every linearly independent list of vectors in V with length $\dim V$ is a basis of V.

Proof. Suppose $\dim V = n$ and $v_1, \ldots, v_n$ is linearly independent in V. The list $v_1, \ldots, v_n$ can be extended to a basis of V. However, every basis of V has length n, so in this case the extensions is the trivial one, meaning that no elements are adjoined to $v_1, \ldots, v_n$. □

Example 2.5.2. Show that the list $(5, 7), (4, 3)$ is a basis of $\mathbf{F}^2$.

Proof. This list of two vectors in $\mathbf{F}^2$ is obviously linearly independent (because neither vector is a scalar multiple of the other). Note that $\mathbf{F}^2$ has dimension 2. We do not need to bother checking that it spans $\mathbf{F}^2$. □

Proposition 2.5.4 (Spanning list of the right length is a basis)**.** Suppose V is finite-dimensional. Then every spanning list of vectors in V with length $\dim V$ is a basis of V.

Proof. Suppose $\dim V = n$ and $v_1, \ldots, v_n$ spans V. The list $v_1, \ldots, v_n$ can be reduced to a basis of V. However, every basis of V has length n, so in this case the reduction is the trivial one, meaning that no elements are deleted from $v_1, \ldots, v_n$. □

Proposition 2.5.5 (Dimension of a sum)**.** If U_1 and U_2 are subspaces of a finite-dimensional vector space, then

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

Proof. Omitted. □

2.6 Exercises

1. Suppose V is finite-dimensional and U is a subspace of V such that $\dim U = \dim V$. Prove that $U = V$.
2. Show that the subspaces of $\mathbf{R}^2$ are precisely $\{0\}, \mathbf{R}^2$, and all lines in $\mathbf{R}^2$ through the origin.
3. Show that the subspaces of $\mathbf{R}^3$ are precisely $\{0\}, \mathbf{R}^3$, all lines in $\mathbf{R}^3$ through the origin, and all planes in $\mathbf{R}^3$ through the origin.
4. (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(6) = 0\}$. Find a basis of U.
   (b) Extend the basis in part(a) to a basis of $\mathcal{P}_4(\mathbf{F})$.
   (c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}(\mathbf{F}) = U \oplus W$.
5. (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5)\}$. Find a basis of U.
   (b) Extend the basis in part(a) to a basis of $\mathcal{P}_4(\mathbf{F})$.
   (c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}(\mathbf{F}) = U \oplus W$.
6. (a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5) = p(6)\}$. Find a basis of U.
   (b) Extend the basis in part(a) to a basis of $\mathcal{P}_4(\mathbf{F})$.
   (c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}(\mathbf{F}) = U \oplus W$.

7. Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that

$$\dim \operatorname{span}(v_1 + w, \ldots, w_m + w) \geq m - 1.$$

8. Suppose $p_0, p_1, \ldots, p_m \in \mathcal{P}(\mathbf{F})$ are such that each p_j has degree j. Prove that $p_0, p_1, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

9. Suppose that U and W are subspaces of $\mathbf{R}^8$ such that $\dim U = 3$, $\dim W = 5$, and $U + W = \mathbf{R}^8$. Prove that $\mathbf{R}^8 = U \oplus W$.

10. Suppose U and W are both five-dimensional subspaces of $\mathbf{R}^9$. Prove that $U \cap W \neq \{0\}$.

11. Suppose U and W are both 4-dimensional subspaces of $\mathbf{C}^6$. Prove that there exist two vectors in $U \cap W$ such that neither of these vectors is a scalar multiple of the other.

12. Suppose $U_1, \ldots, U_m$ are finte-dimensional subspaces of V. Prove that $U_1 + \cdots + U_m$ is finite-dimensional and

$$\dim(U_1 + \cdots + U_m) \leq \dim U_1 + \cdots + \dim U_m.$$

13. Suppose V is finite-dimensional, with $\dim V = n \geq 1$. Prove that there exists 1-dimensional subspaces $U_1, \ldots, U_n$ of V such that

$$V = U_1 \oplus \cdots \oplus U_n.$$

14. Suppose $U_1, \ldots, U_m$ are finite-dimensional subspaces of V such that $U_1 + \cdots + U_m$ is a direct sum. Prove that $U_1 \oplus \cdots \oplus U_m$ is a finite-dimensional and

$$\dim(U_1 \oplus \cdots w \oplus U_m) = \dim U_1 + \cdots + \dim U_m.$$

15. If U_1, U_2, U_3 are subspaces of a finite-dimensional vector space, then

$$\dim(U_1+U_2+U_3) = \dim U_1+\dim U_2+\dim U_3-\dim(U_1\cap U_2)-\dim(U_1\cap U_3)-\dim(U_2\cap U_3)+\dim(U_1\cap U_2\cap U_3)$$

Prove this or give a counterexample.

Chapter 3

Linear Maps

In this chapter we will frequently need another vector space, which we will call W, in addition to V.

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V and W denote vector spaces over $\mathbf{F}$.

3.1 The vector space of linear maps

3.1.1 Linear maps

Definition 3.1.1 (linear map, $\mathcal{L}(V, W)$)**.** A **linear map** from V to W is a function $T : V \to W$ with the following properties:

additivity $T(u + v) = Tu + Tv$ for all $u, v \in V$;

homogeneity $T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbf{F}$ and all $v \in V$.

The set of all linear maps from V to W is denoted $\mathcal{L}(V, W)$.

Example 3.1.1 (linear maps)**.** Make sure you verify that each of the functions defined below is indeed a linear map:

zero In addition to its other uses, we let the symbol 0 denote the function that takes each element of some vector space to the additive identity of another vector space. To be specific, $0 \in \mathcal{L}(V, W)$ is defined by

$$0v = 0.$$

identity The **identity map**, denoted I, is the function on some vector space that takes each element to itself. To be specific, $I \in \mathcal{L}(V, V)$ is defined by

$$Iv = v.$$

multiplication by x^2 Define $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$ by

$$(Tp)(x) = x^2 p(x)$$

for $x \in \mathbf{R}$.

backward shift Define $T \in \mathcal{L}(\mathbf{F}^\infty, \mathbf{F}^\infty)$ by

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

from $\mathbf{R}^3$ to $\mathbf{R}^2$ Define $T \in \mathcal{L}(\mathbf{R}^3, \mathbf{R}^2)$ by

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z).$$

from $\mathbf{F}^n$ to $\mathbf{F}^m$ Generalizing the previous example, let m and n be positive integers, let $A_{j,k} \in \mathbf{F}$ for $j = 1, \dots, m$ and $k = 1, \dots, n$, and define $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ by

$$T(x_1, \dots, x_n) = (\sum_{k=1}^{n} A_{1,k} x_k, \dots, \sum_{k=1}^{n} A_{m,k} x_k)$$

Actually every linear map from $\mathbf{F}^n$ to $\mathbf{F}^m$ is of this form.

The existence part of the next result means that we can find a linear map that takes on whatever values we wish on the vectors in a basis. The uniqueness part of the next result means that a linear map is completely determined by its values on a basis.

Proposition 3.1.1 (Linear maps and basis of domain)**.** Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_n \in W$. Then there exists a unique linear map $T : V \to W$ such that

$$Tv_j = w_j$$

for each $j = 1, \dots, n$.

Proof. First we show the existence of a linear map T with the desired property. Define $T : V \to W$ by

$$T(\sum_{j=1}^{n} c_i v_i) = \sum_{j=1}^{n} c_i w_i$$

where $c_1, \dots, c_n$ are arbitrary elements of $\mathbf{F}$. The list $v_1, \dots, v_n$ is a basis of V, and thus the equation above does indeed define a function T from V to W, and T is a linear map from V to W.

To prove the uniqueness, now suppose that $T \in \mathcal{L}(V, W)$ and that $Tv_j = w_j$ for $j = 1, \dots, n$. Let $c_1, \dots, c_n \in \mathbf{F}$. The homogeneity of T implies that $T(c_j v_j) = c_j w_j$ for $j = 1, \dots, n$. The additivity of T now implies that

$$T(\sum_{j=1}^{n} c_i v_i) = \sum_{j=1}^{n} c_i w_i$$

Thus T is uniquely determined on $\operatorname{span}(v_1, \dots, v_n)$ by the equation above. Because $v_1, \dots, v_n$ is a basis of V, this implies that T is uniquely determined on V. $\square$

3.1.2 Algebraic operations on $\mathcal{L}(V, W)$

Definition 3.1.2 (addition and scalar multiplication on $\mathcal{L}(V, W)$)**.** Suppose $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbf{F}$. The **sum** $S + T$ and the **product** λT are the linear maps from V to W defined by

$$(S + T)(v) = Sv + Tv \text{ and } (\lambda T)(v) = \lambda(Tv)$$

for all $v \in V$. You should verity that $S + T, \lambda T \in \mathcal{L}(V, W)$

Proposition 3.1.2 ($\mathcal{L}(V, W)$ is a vector space)**.** With the operations of addition and scalar multiplication as defined above, $\mathcal{L}(V, W)$ is a vector space.

Proof. Omitted. Note that the additive identity of $\mathcal{L}(V, W)$ is the zero linear map defined in Example 3.1.1. □

Definition 3.1.3 (Product of Linear Maps)**.** If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then the product $ST \in \mathcal{L}(U, W)$ is defined by

$$(ST)(u) = S(Tu)$$

for $u \in U$. In other words, ST is just the usual composition $S \circ T$ of two functions. You should verify that ST is indeed a linear map from U to W. Note that ST is defined **only when** T maps into the domain of S.

associativity

$$(T_1 T_2)T_3 = T_1(T_2 T_3)$$

whenever T_1, T_2, T_3 are linear maps such that the products make sense.

identity

$$TI = IT = T$$

whenever $T \in \mathcal{L}(V, W)$, the first I is the identity map on V, and the second I is the identity map on W.

distributive properties

$$(S_1 + S_2)T = S_1 T + S_2 T \text{ and } S(T_1 + T_2) = ST_1 + ST_2$$

whenever $T, T_1, T_2 \in \mathcal{L}(U, V)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$.

not commutative Multiplication of linear maps is not commutative. In other words, it is not necessarily true that $ST = TS$, even if both sides of the equation make sense.

Proposition 3.1.3 (Linear maps take 0 to 0)**.** Suppose T is a linear map from V to W. Then $T(0) = 0$.

Proof. By additivity, we have

$$T(0) = T(0 + 0) = T(0) + T(0).$$

Add the additive inverse of $T(0)$ to each side of the equation above to conclude that $T(0) = 0$. □

3.2 Exercises

1. Suppose $b, c \in \mathbf{R}$. Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by
$$T(x, y, z) = (2x - 4y + 3z + b, 6x + cxyz).$$
Show that T is linear if and only if $b = c = 0$.
2. Suppose $T \in \mathcal{L}(\mathbf{F}^n\mathbf{F}^m)$. Show that there exists scalars $A_{j,k} \in \mathbf{F}$ for $j = 1, \dots, m$ and $k = 1, \dots, n$ such that
$$T(x_1, \dots, x_n) = (\sum_{k=1}^{n} A_{1,k}x_k, \dots, \sum_{k=1}^{n} A_{m,k}x_k)$$
for every $(x_1, \dots, x_n) \in \mathbf{F}^n$.
3. Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_m$ is a list of vectors in V such that $Tv_1, \dots, Tv_m$ is a linearly independent list in W. Prove that $v_1, \dots, v_m$ is linearly independent.
4. Prove that if $\dim V = 1$ and $T \in \mathcal{L}(V, V)$, then there exists $\lambda \in \mathbf{F}$ such that $Tv = \lambda v$ for all $v \in V$.
5. Give an example of a function $\varphi : \mathbf{R}^2 \to \mathbf{R}$ such that
$$\varphi(av) = a\varphi(v)$$
for all $a \in \mathbf{R}$ and all $v \in \mathbf{R}^2$ but φ is not linear.
6. Give an example of a function $\varphi : \mathbf{C} \to \mathbf{C}$ such that
$$\varphi(w + z) = \varphi(w) + \varphi(z)$$
for all $w, z \in \mathbf{C}$ but φ is not linear.
7. Give an example of a function $\varphi : \mathbf{R} \to \mathbf{R}$ such that
$$\varphi(w + z) = \varphi(w) + \varphi(z)$$
for all $w, z \in \mathbf{R}$ but φ is not linear.
8. Suppose U is a subspace of V with $U \neq V$. Suppose $S \in \mathcal{L}(U, W)$ and $S \neq 0$. Define $T : V \to W$ by
$$Tv = \begin{cases} Sv & \text{if } v \in U, \\ 0 & \text{if } v \in V \text{ and } v \notin U. \end{cases}$$
Prove that T is **not** a linear map on V.
9. Suppose V is finite-dimensional. Show that if U is a subspace of V and $S \in \mathcal{L}(U, W)$, then there exists $T \in \mathcal{L}(V, W)$ such that $Tu = Su$ for all $u \in U$.

10. Suppose V is finite-dimensional with $\dim V > 0$, and suppose W is infinite-dimensional. Prove that $\mathcal{L}(V,W)$ is infinite-dimensional.
11. Suppose $v_1, \dots, v_m$ is a linearly dependent list of vectors in V. Suppose also that $W \neq \{0\}$. Prove that there exists $w_1, \dots, w_m \in W$ such that no $T \in \mathcal{L}(V,W)$ satisfies $Tv_k = w_k$ for each $k = 1, \dots, m$.
12. Suppose V is finite-dimensional with $\dim V \geq 2$. Prove that there exist $S, T \in \mathcal{L}(V,V)$ such that $ST \neq TS$.

3.3 Null spaces and ranges

3.3.1 Null space and injectivity

Definition 3.3.1 (null space, null T)**.** For $T \in \mathcal{L}(V,W)$, the **null space** of T, denoted null T, is the subset of V consisting of those vectors that T maps to 0:

$$\text{null } T = \{v \in V : Tv = 0\}.$$

Example 3.3.1.
- If T is the zero map from V to W, in other words if $Tv = 0$ for every $v \in V$, then null $T = V$.
- Suppose $\varphi \in \mathcal{L}(\mathbf{C}^3, \mathbf{C})$ is defined by $\varphi(z_1, z_2, z_3) = z_1 + 2z_2 + 3z_3$. Then null $T = \{(z_1, z_2, z_3) \in \mathbf{C}^3 : z_1 + 2z_2 + 3z_3 = 0\}$. A basis of null φ is $(-2,1,0), (-3,0,1)$.
- Suppose $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$ is the multiplication by x^2 map defined by $(Tp)(x) = x^2p(x)$. The only polynomial p such that $x^2p(x) = 0$ for all $x \in \mathbf{R}$ is the 0 polynomial. Thus null $T = \{0\}$.
- Suppose $T \in \mathcal{L}(\mathbf{F}^\infty, \mathbf{F}^\infty)$ is the backward shift defined by

  $$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

  Clearly $T(x_1, x_2, x_3, \dots) = 0$ if and only if $x_2, x_3, \dots$ are all 0. Thus in this case we have null $T = \{(a, 0, 0, \dots) : a \in \mathbf{F}\}$.

Proposition 3.3.1 (The null space is a subspace)**.** Suppose $T \in \mathcal{L}(V,W)$. Then null T is a subspace of V.

Proof. Omitted. □

Definition 3.3.2 (injective)**.** A function $T : V \to W$ is called **injective** is $Tu = Tv$ implies $u = v$. In other words, T is injective if it maps distinct inputs to distinct outputs.

Proposition 3.3.2 (Injectivity is equivalent to null space equals $\{0\}$)**.** Let $T \in \mathcal{L}(V,W)$. Then T is injective if and only if null $T = \{0\}$.

Proof. First suppose T is injective. We want to prove that null $T = \{0\}$. We already know that $\{0\} \subset$ null T. Th prove the inclusion in the other direction, suppose $v \in$ null T. Then

$$T(v) = 0 = T(0).$$

Because T is injective, the equation above implies that $v = 0$. Thus we can conclude that null $T = \{0\}$, as desired.

To prove the implication in the other direction, now suppose null $T = \{0\}$. We want to prove that T is injective. To do this, suppose $u, v \in V$ and $Tu = Tv$. Then

$$0 = Tu - Tv = T(u - v).$$

Thus $u - v$ is in null T, which equals $\{0\}$. Hence $u - v = 0$, which implies that $u = v$. Hence T is injective, as desired. □

3.3.2 Range and surjectivity

Definition 3.3.3 (range)**.** For T a function from V to W, the **range** of T is the subset of W consisting of those vectors that are of the form Tv for some $v \in V$:

$$\text{range } T = \{Tv : v \in V\}.$$

Example 3.3.2.

- If T is the zero map from V to W, in other words if $Tv = 0$ for every $v \in V$, then range $T = \{0\}$.
- Suppose $T \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^3)$ is defined by $T(x, y) = (2x, 5y, x + y)$, then range $T = \{(2x, 5y, x+y) : x, y \in \mathbf{R}\}$. A basis of range T is $(2, 0, 1), (0, 5, 1)$.

Proposition 3.3.3 (The range is a subspace)**.** If $T \in \mathcal{L}(V, W)$, then range T is a subspace of W.

Proof. Omitted. □

Definition 3.3.4 (surjective)**.** A function $T : V \to W$ is called **surjective** is its range equals W. **Whether a linear map is surjective depends on what we are thinking of as the vector space into which it maps.**

3.3.3 Fundamental theorem of linear maps

Theorem 3.3.1 (Fundamental Theorem of Linear Maps)**.** Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Then range T is finite-dimensional and

$$\dim V = \dim \text{null } T + \dim \text{range } T.$$

In particular, we have

$$\dim \text{null } T \geq \dim V - \dim W, \dim \text{range } T \leq \dim V.$$

Proof. Let $u_1, \ldots, u_m$ be a basis of null T; thus $\dim \operatorname{null} T = m$. The linearly independent list $u_1, \ldots, u_m$ can be extended to a basis

$$u_1, \ldots, u_m, v_1, \ldots, v_n$$

of V. Thus $\dim V = m + n$. To complete the proof, we need only show that the range T is finite-dimensional and $\dim \operatorname{range} T = n$. We will do this by proving that $Tv_1, \ldots, Tv_n$ is a basis of range T.

Let $v \in V$. Because $u_1, \ldots, u_m, v_1, \ldots, v_n$ spans V, we can write

$$v = \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j,$$

where the a's and b's are in $\mathbf{F}$. Applying T to both sides of this equation, we get

$$Tv = \sum_{j=1}^{n} b_j T v_j,$$

where the terms of the form Tu_i disappeared because each u_i is in null T. The last equation implies that $Tv_1, \ldots, Tv_n$ spans range T. In particular, range T is finite-dimensional. To show $Tv_1, \ldots, Tv_n$ is linearly independent, suppose $c_1, \ldots, c_n \in \mathbf{F}$ and

$$\sum_{j=1}^{n} c_j T v_j = T\left(\sum_{j=1}^{n} c_j v_j\right) = 0.$$

Hence $\sum_{j=1}^{n} c_j v_j \in \operatorname{null} T$. Because $u_1, \ldots, u_m$ spans null T, we can write

$$c_1 v_1 + \cdots + c_n v_n = d_1 u_1 + \cdots + d_m u_m,$$

where the d's are in $\mathbf{F}$. This equation implies that all the c's and d's are 0, because $u_1, \ldots, u_m, v_1, \ldots, v_n$ is linearly independent. Thus $Tv_1, \ldots, Tv_n$ is linearly independent and hence is a basis of range T, as desired. □

Proposition 3.3.4 (A map to a smaller dimensional space is not injective)**.** Suppose V and W are finite-dimensional vector spaces such that $\dim V > \dim W$. Then no linear map from V to W is injective.

Proof.

$$\dim \operatorname{null} T \geq \dim V - \dim W > 0.$$

□

Proposition 3.3.5 (A map to a larger dimensional space is not surjective)**.** Suppose V and W are finite-dimensional vector spaces such that $\dim V < \dim W$. Then no linear map from V to W is surjective.

Proof.

$$\dim \operatorname{range} T \leq \dim V < \dim W.$$

□

3.4 Linear equations

3.4.1 Homogeneous system of linear equations

Fix positive integers m and n, and let $A_{j,k} \in \mathbf{F}$ for $j = 1, \ldots, m$ and $k = 1, \ldots, n$. Consider the homogeneous system of linear equations

$$\sum_{k=1}^{n} A_{1,k} x_k = 0$$
$$\vdots$$
$$\sum_{k=1}^{n} A_{m,k} x_k = 0.$$

Obviously $x_1 = \cdots = x_n = 0$ is a solution of the system of equations above; the question here is whether any other solutions exist.

Define $T : \mathbf{F}^n \to \mathbf{F}^m$ by

$$T(x_1, \ldots, x_n) = \left(\sum_{k=1}^{n} A_{1,k} x_k, \ldots, \sum_{k=1}^{n} A_{m,k} x_k \right)$$

The equation $T(x_1, \ldots, x_n) = 0$ is the same as the homogeneous system of linear equations above.

Thus we want to know if null T is strictly bigger than $\{0\}$. In other words, we can rephrase our question about nonzero solutions as follows: What condition ensures that T is not injective?

Proposition 3.4.1 (Homogeneous system of linear equations)**.** A homogeneous system of linear equations with more variables than equations has nonzero solutions.

3.4.2 Inhomogeneous system of linear equations

Fix positive integers m and n, and let $A_{j,k} \in \mathbf{F}$ for $j = 1, \ldots, m$ and $k = 1, \ldots, n$. For $c_1, \ldots, c_m \in \mathbf{F}$, consider the system of linear equations

$$\sum_{k=1}^{n} A_{1,k} x_k = c_1$$
$$\vdots$$
$$\sum_{k=1}^{n} A_{m,k} x_k = c_m.$$

The question here is whether there is some choice of $c_1, \ldots, c_m \in \mathbf{F}$ such that no solution exists to the system above.

Define $T : \mathbf{F}^n \to \mathbf{F}^m$ by

$$T(x_1, \ldots, x_n) = \left(\sum_{k=1}^{n} A_{1,k} x_k, \ldots, \sum_{k=1}^{n} A_{m,k} x_k \right)$$

The equation $T(x_1, \ldots, x_n) = (c_1, \ldots, c_m)$ is the same as the system of equations above. Thus we want to know if range $T \neq \mathbf{F}^m$. Hence we can rephrase our question about not having a solution for some choice of $c_1, \ldots, c_m \in \mathbf{F}$ as follows: What condition ensures that T is not surjective?

Proposition 3.4.2 (Inhomogeneous system of linear equations)**.** An inhomogeneous system of linear equations with more equations than variables has no solution for some choice of the constant terms.

3.5 Exercises

1. Give an example of a linear map T such that

$$\dim \operatorname{null} T = 3, \dim \operatorname{range} T = 2$$

2. Suppose V is a vector space and $S, T \in \mathcal{L}(V, V)$ are such that

$$\operatorname{range} S \subset \operatorname{null} T$$

   Prove that $(ST)^2 = 0$.

3. Suppose $v_1, \ldots, v_m$ is a list of vectors in V. Define $T \in \mathcal{L}(\mathbf{F}^m, V)$ by

$$T(z_1, \ldots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$

   (a) What property of T corresponds to $v_1, \ldots, v_m$ spanning V?

   (b) What property of T corresponds to $v_1, \ldots, v_m$ being linearly independent?

4. Show that

$$\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$$

   is not a subspace of $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$.

5. Give an example of a linear map $T : \mathbf{R}^4 \to \mathbf{R}^4$ such that

$$\operatorname{range} T = \operatorname{null} T.$$

6. Prove that there does not exist a linear map $T : \mathbf{R}^5 \to \mathbf{R}^5$ such that

$$\operatorname{range} T = \operatorname{null} T.$$

7. Suppose V and W are finite-dimensional with $2 \leq \dim V \leq \dim W$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

8. Suppose V and W are finite-dimensional with $\dim V \geq \dim W \geq 2$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not surjective}\}$ is not a subspace of $\mathcal{L}(V, W)$
9. Suppose $T \in \mathcal{L}(V, W)$ is injective and $v_1, \ldots, v_n$ is linearly independent in V. Prove that $Tv_1, \ldots, Tv_n$ is linearly independent in W.
10. Suppose $v_1, \ldots, v_n$ spans V and $T \in \mathcal{L}(V, W)$. Prove that the list $Tv_1, \ldots, Tv_n$ spans range T.
11. Suppose $S_1, \ldots, S_n$ are injective linear maps such that $S_1 S_2 \cdots S_n$ makes sense. Prove that $S_1 S_2 \cdots S_n$ is injective.
12. Suppose that V is finite-dimensional and that $T \in \mathcal{L}(V, W)$. Prove that there exists a subspace U of V such that $U \cap \operatorname{null} T = \{0\}$ and range $T = \{Tu : u \in U\}$.
13. Suppose T is a linear map from $\mathbf{F}^4$ to $\mathbf{F}^2$ such that

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2, x_3 = 7x_4\}.$$

Prove that T is surjective.
14. Suppose U is a 3-dimensional subspace of $\mathbf{R}^8$ and that T is a linear map from $\mathbf{R}^8$ to $\mathbf{R}^5$ such that null $T = U$. Prove that T is surjective.
15. Prove that there does not exist a linear map from $\mathbf{F}^5$ to $\mathbf{F}^2$ whose null space equals

$$\{(x_1, x_3, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2, x_3 = x_4 = x_5\}.$$

16. Suppose there exists a linear map on V whose null space and range are both finite-dimensional. Prove that V is finite-dimensional.
17. Suppose V and W are both finite-dimensional. Prove that there exists an injective linear map from V to W if and only if $\dim V \leq \dim W$.
18. Suppose V and W are both finite-dimensional. Prove that there exists an surjective linear map from V to W if and only if $\dim V \geq \dim W$.
19. Suppose V and W are finite-dimensional and that U is a subspace of V. Prove that there exists $T \in \mathcal{L}(V, W)$ such that null $T = U$ if and only if $\dim U \geq \dim V - \dim W$.
20. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is injective if and only if there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity map on V.
21. Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is surjective if and only if there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity map on W.
22. Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

23. Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

24. Suppose W is finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that null $T_1 \subset$ null T_2 if and only if there exists $S \in \mathcal{L}(W, W)$ such that $T_2 = ST_1$.
25. Suppose V is finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that range $T_1 \subset$ range T_2 if and only if there exists $S \in \mathcal{L}(V, V)$ such that $T_1 = T_2S$.
26. Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$ is such that $\deg(Dp) = (\deg p) - 1$ for every nonconstant polynomial $p \in \mathcal{P}(\mathbf{R})$. Prove that D is surjective.
27. Suppose $T \in \mathcal{L}(V, W)$, and $w_1, \ldots, w_m$ is a basis of range T. Prove that there exists $\varphi_1, \ldots, \varphi_m \in \mathcal{L}(V, \mathbf{F})$ such that

$$Tv = \sum_{i=1}^{m} \varphi_i(v) w_i$$

for every $v \in V$.
28. Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$. Suppose $u \in V$ is not in null φ. Prove that

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbf{F}\}.$$

29. Suppose φ_1 and φ_2 are linear maps from V to $\mathbf{F}$ that have the same null space. Show that there exists a constant $c \in \mathbf{F}$ such that $\varphi_1 = c\varphi_2$.
30. Give an example of two linear maps T_1 and T_2 from $\mathbf{R}^5$ to $\mathbf{R}^2$ that have the same null space but are such that T_1 is not a scalar multiple of T_2.

3.6 Invertibility and isomorphic vector spaces

3.6.1 Invertible linear maps

Definition 3.6.1 (invertible, inverse)**.** We begin this section by defining the notions of invertible and inverse in the context of linear maps.

- A linear map $T \in \mathcal{L}(V, W)$ is called **invertible** if there exists a linear map $S \in \mathcal{L}(W, V)$ such that ST equals the identity map on V and TS equals the identity map on W.
- A linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I_V$ and $TS = I_W$ is called an inverse of T.

Proposition 3.6.1 (Inverse is unique)**.** An invertible linear map has a unique inverse.

Now that we know that the inverse is unique, we can give it a notation. If T is invertible, then its inverse is denoted by T^{-1}. In other words, if $T \in \mathcal{L}(V, W)$ is invertible, then T^{-1} is the unique element of $\mathcal{L}(W, V)$ such that $T^{-1}T = I$ and $TT^{-1} = I$.

Proposition 3.6.2 (Invertibility is equivalent to injectivity and surjectivity). A linear map is invertible if and only if it is injective and surjective.

Proof. Suppose $T \in \mathcal{L}(V, W)$. We need to show that T is invertible if and only if it is injective and surjective.

First suppose T is invertible. To show that T is injective, suppose $u, v \in V$ and $Tu = Tv$. Then

$$u = T^{-1}(Tu) = T^{-1}(Tv) = v,$$

so $u = v$. Hence T is injective.

We are still assuming that T is invertible. Now we want to prove that T is surjective. To do this, let $w \in W$. Then $w = T(T^{-1}w)$, which shows that w is in the range of T. Thus range $T = W$. Hence T is surjective, completing this direction of the proof.

Now suppose T is injective and surjective. We want to prove that T is invertible. For each $w \in W$, define Sw to be the unique element of V such that $T(Sw) = w$ (the existence and uniqueness of such an element follow from the surjectivity and injectivity of T). Clearly $T \circ S$ equals the identity map on W.

To prove that $S \circ T$ equals the identity map on V, let $v \in V$. Then

$$T((S \circ T)v) = (T \circ S)(Tv) = I(Tv) = Tv.$$

This equation implies that $(S \circ T)v = v$ because T is injective. Thus $S \circ T$ equals the identity map on V.

To complete the proof, we need to show that S is linear. To do this, suppose $w_1, w_2 \in W$. Then

$$T(Sw_1 + Sw_2) = T(Sw_1) + T(Sw_2) = w_1 + w_2.$$

Thus $Sw_1 + Sw_2$ is the unique element of V that T maps to $w_1 + w_2$. By the definition of S, this implies that $S(w_1 + w_2) = Sw_1 + Sw_2$. Hence S satisfies the additive property required for linearity.

The proof of homogeneity is similar. Specifically, if $w \in W$ and $\lambda \in \mathbf{F}$, then

$$T(\lambda Sw) = \lambda T(Sw) = \lambda w.$$

Thus λSw is the unique element of V that T maps to λw. By the definition of S, this implies that $S(\lambda w) = \lambda Sw$. Hence S is linear, as desired. $\square$

Example 3.6.1. linear maps that are not invertible:

- The multiplication by x^2 linear map from $\mathcal{P}(\mathbf{R})$ to $\mathcal{P}(\mathbf{R})$ is not invertible because it is not surjective.
- The backward shift linear map from $\mathbf{F}^\infty$ to $\mathbf{F}^\infty$ is not invertible because it is not injective.

3.6.2 Isomorphic vector spaces

Definition 3.6.2 (isomorphism, isomorphic)**.** This definition captures the idea of two vector spaces that are essentially the same, except for the names of the elements of the vector spaces.

- An **isomorphism** is an invertible linear map.
- Two vector spaces are called **isomorphic** if there is an isomorphism from one vector space onto the other one.

The Greek word isos means equal; the Greek word morph means shape. Thus isomorphic literally means equal shape. Think of an isomorphism $T : V \to W$ as relabeling $v \in V$ as $Tv \in W$. This viewpoint explains why two isomorphic vector spaces have the same vector space properties.

Proposition 3.6.3 (Dimension shows whether vector spaces are isomorphic)**.** Two finite-dimensional vector spaces over $\mathbf{F}$ are isomorphic if and only if they have the same dimension.

Proof. Omitted. This result implies that each finite-dimensional vector space V is isomorphic to $\mathbf{F}^{\dim V}$. $\square$

Proposition 3.6.4 ($\mathcal{L}(V,W)$ and $\mathbf{F}^{m\times n}$ are isomorphic)**.** Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ is a basis of W. Then we can define $A_{j,k} \in \mathbf{F}$ by

$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j$$

Then we define $\mathcal{M}(T) = (A_{j,k}) \in \mathbf{F}^{m\times n}$. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V,W)$ and $\mathbf{F}^{m\times n}$.

Proof. Omitted. Notice that we have $\dim \mathcal{L}(V,W) = (\dim V)(\dim W)$. $\square$

3.6.3 Operators

Definition 3.6.3 (operator, $\mathcal{L}(V)$)**.** Linear maps from a vector space to itself are so important that they get a special name and special notation.

- A linear map from a vector space to itself is called an **operator**.
- The notation $\mathcal{L}(V)$ denotes the set of all operators on V. In other words, $\mathcal{L}(V) = \mathcal{L}(V,V)$.

Example 3.6.2. On infinite-dimensional vector spaces, neither condition alone implies invertibility:

- The multiplication by x^2 operator on $\mathcal{P}(\mathbf{R})$ is injective but not surjective.
- The backward shift operator on $\mathbf{F}^\infty$ is surjective but not injective.

Proposition 3.6.5 (Injectivity is equivalent to surjectivity in finite dimensions)**.** Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivalent:

1. T is invertible;
2. T is injective;
3. T is surjective.

Proof. This is nothing but the Fundamental Theorem of Linear Maps. □

3.7 Exercises

1. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that with respect to each choice of bases of V and W, $\mathcal{M}(T)$ has at least $\dim \operatorname{range} T$ nonzero entries.
2. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ are 0 except that the entries $A_{j,j}$ equal 1 for $1 \leq j \leq \dim \operatorname{range} T$.
3. Suppose $v_1, \ldots, v_m$ is a basis of V and W is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $w_1, \ldots, w_n$ of W such that $\mathcal{M}(T)_{j,1}$ are 0 except for possibly $\mathcal{M}(T)_{1,1} = 1$.
4. Suppose $w_1, \ldots, w_n$ is a basis of W and V is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $v_1, \ldots, v_m$ of V such that $\mathcal{M}(T)_{1,k}$ are 0 excepr for possibly $\mathcal{M}(T)_{1,1} = 1$.
5. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that $\dim \operatorname{range} T = 1$ if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ equal 1.
6. In the following result, the assumption is that the same bases are used for all three linear maps.
   (a) $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$
   (b) $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$
   (c) $\mathcal{M}(ST)_{j,k} = \sum_r \mathcal{M}(S)_{j,r} \mathcal{M}(T)_{r,k}$
   (d) $\mathcal{M}(T^3)_{j,k} = \sum_{p,r} \mathcal{M}(T)_{j,p} \mathcal{M}(T)_{p,r} \mathcal{M}(T)_{r,k}$
7. Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$ are both invertible linear maps. Prove that $ST \in \mathcal{L}(U, W)$ is invertible and that $(ST)^{-1} = T^{-1} S^{-1}$.
8. Suppose V is finite-dimensional and $\dim V > 1$. Prove that the set of noninvertible operators on V is not a subspace of $\mathcal{L}(V)$.
9. Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove there exists an invertible operator $T \in \mathcal{L}(V)$. such that $Tu = Su$ for every $u \in U$ if and only if S is injective.

10. Suppose W is finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} T_1 = \operatorname{null} T_2$ if and only if there exists an invertible operator $S \in \mathcal{L}(W)$ such that $T_1 = ST_2$.

11. Suppose V is finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} T_1 = \operatorname{range} T_2$ if and only if there exists an invertible operator $S \in \mathcal{L}(V)$ such that $T_1 = T_2 S$.

12. Suppose V and W are finite-dimensional and $T_1, T_2 \in \mathcal{L}(V, W)$. Prove that there exist invertible operators $R \in \mathcal{L}(V)$ and $S \in \mathcal{L}(W)$ such that $T_1 = ST_2R$ if and only if $\dim \operatorname{null} T_1 = \dim \operatorname{null} T_2$.

13. Suppose V and W are finite-dimensional. Let $v \in V$. Let

$$E = \{T \in \mathcal{L}(V, W) : Tv = 0\}.$$

   (a) Show that E is a subspace of $\mathcal{L}(V, W)$.
   (b) Suppose $v \neq 0$. What is $\dim E$?

14. Suppose V is finite-dimensional and $T : V \to W$ is a surjective linear map of V onto W. Prove that there is a subspace U of V such that $T|_U$ is an isomorphism of U onto W.

15. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST is invertible if and only if both S and T are invertible.

16. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that $ST = I$ if and only if $TS = I$.

17. Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

18. Show that the result in the previous exercise can fail without the hypothesis that V is finite-dimensional.

19. Suppose V is a finite-dimensional vector space and $R, S, T \in \mathcal{L}(V)$ are such that RST is surjective. Prove that S is injective.

20. Suppose $v_1, \ldots, v_n$ is a basis of V. Prove that the map $T : V \to \mathbf{F}^n$ defined by

$$Tv = (c_1, \ldots, c_n), v = \sum_{i=1}^{n} c_i v_i$$

    is an isomorphism of V onto $\mathbf{F}^n$.

21. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

22. Suppose V is finite-dimensional and $\mathcal{E}$ is a subspace of $\mathcal{L}(V)$ such that $ST \in \mathcal{E}$ and $TS \in \mathcal{E}$ for all $S \in \mathcal{L}(V)$ and all $T \in \mathcal{E}$. Prove that $\mathcal{E} = \{0\}$ or $\mathcal{E} = \mathcal{L}(V)$.

23. Show that V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

24. Suppose $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is such that T is injective and $\deg Tp \leq \deg p$ for every nonzero polynomial $p \in \mathcal{P}(\mathbf{R})$.

(a) Prove that T is surjective.

(b) Prove that $\deg Tp = \deg p$ for every nonzero $p \in \mathcal{P}(\mathbf{R})$.

25. Suppose n is a positive integer and $A_{j,k} \in \mathbf{F}$ for $i, j = 1, \ldots, n$. Prove that the following are equivalent:

    (a) The trivial solution $x_1 = \cdots = x_n = 0$ is the only solution to the homogeneous system of equations

    $$\sum_{k=1}^{n} A_{1,k} x_k = 0$$
    $$\vdots$$
    $$\sum_{k=1}^{n} A_{n,k} x_k = 0.$$

    (b) For every $c_1, \ldots, c_n \in \mathbf{F}$, there exists a solution to the system of equations

    $$\sum_{k=1}^{n} A_{1,k} x_k = c_1$$
    $$\vdots$$
    $$\sum_{k=1}^{n} A_{n,k} x_k = c_n.$$

Chapter 4

Products, Quotients, and Duality

4.1 Products and quotients

4.1.1 Product of vector spaces

Definition 4.1.1 (product of vector spaces)**.** Suppose $V_1, \ldots, V_m$ are vector spaces over $\mathbf{F}$.

- The **product** $V_1 \times \cdots \times V_m$ is defined by

$$V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) : v_1 \in V_1, \ldots, v_m \in V_m\}.$$

- Addition on $V_1 \times \cdots \times V_m$ is defined by

$$(u_1, \ldots, u_m) + (v_1, \ldots, v_m) = (u_1 + v_1, \ldots, u_m + v_m).$$

- Scalar multiplication on $V_1 \times \cdots \times V_m$ is defined by

$$\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m).$$

Example 4.1.1.

$$(5 - 6x + 4x^2, (3, 8, 7)) \in \mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^3$$

Proposition 4.1.1 (Product of vector spaces is a vector space)**.** Suppose $V_1, \ldots, V_m$ are vector spaces over $\mathbf{F}$. Then $V_1 \times \cdots \times V_m$ is a vector space over $\mathbf{F}$.

Proof. Omitted. Note that the additive identity of $V_1 \times \cdots \times V_m$ is $(0, \ldots, 0)$, where the 0 in the j^{th} slot is the additive identity of V_j. The additive inverse of $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$ is $(-v_1, \ldots, -v_m)$. $\square$

Example 4.1.2. $\mathbf{R}^2 \times \mathbf{R}^3$ is not equal to $\mathbf{R}^5$, but it is isomorphic to $\mathbf{R}^5$. The isomorphism is:

$$((x_1, x_2, x_3), (x_4, x_5)) \mapsto (x_1, x_2, x_3, x_4, x_5).$$

Example 4.1.3 (Find a basis of $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$)**.**

$$(1, (0, 0)), (x, (0, 0)), (x^2(0, 0)), (0, (1, 0)), (0, (0, 1))$$

Proposition 4.1.2 (Dimension of a product is the sum of dimensions)**.** Suppose $V_1, \ldots, V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

Proof. Choose a basis of each V_j. For each basis vector of each V_j, consider the element of $V_1 \times \cdots \times V_m$ that equals the basis vector in the j^{th} slot and 0 in the other slots. The list of all such vectors is linearly independent and spans $V_1 \times \cdots \times V_m$. Thus it is a basis of $V_1 \times \cdots \times V_m$. The length of this basis is $\dim V_1 + \cdots + \dim V_m$, as desired. □

Proposition 4.1.3 (Products and direct sums)**.** Suppose that $U_1, \ldots, U_m$ are subspaces of V. Define a linear map $\Gamma : U_1 \times \cdots \times U_m \to U_1 + \cdots + U_m$ by

$$\Gamma(u_1, \ldots, u_m) = u_1 + \cdots + u_m.$$

Then $U_1 + \cdots + U_m$ is a direct sum if and only if Γ is injective.

Proof. Omitted. □

Proposition 4.1.4 (A sum is a direct sum if and only if dimensions add up)**.** Suppose V is finite-dimensional and $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is a direct sum if and only if

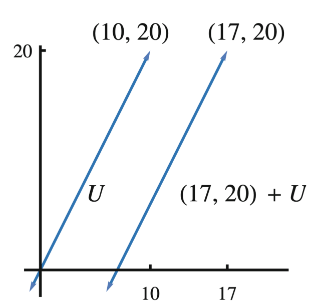
$$\dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m.$$

Proof. Omitted. □

4.1.2 Quotients of vector vpaces

Definition 4.1.2 ($v + U$)**.** Suppose $v \in V$ and U is a subspace of V. Then $v + U$ is the subset of V defined by

$$v + U = \{v + u : u \in U\}.$$



Example 4.1.4. Suppose $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$. Then U is the line in $\mathbf{R}^2$ through the origin with slope 2. Thus

$$(17, 20) + U$$

is the line in $\mathbf{R}^2$ that contains the point $(17, 20)$ and has slope 2.

Definition 4.1.3 (affine subset, parallel, quotient space, V/U)**.** Let V be a vector space.

- An **affine subset** of V is a subset of V of the form $v + U$ for some $v \in V$ and some subspace U of V.
- For $v \in V$ and U a subspace of V, the affine subset $v + U$ is said to be **parallel** to U.
- Suppose U is a subspace of V. Then the **quotient space** V/U is the set of all affine subsets of V parallel to U. In other words,

$$V/U = \{v + U : v \in V\}.$$

Example 4.1.5. With the definition of parallel given, no line in $\mathbf{R}^3$ is considered to be an affne subset that is parallel to the plane U.

- In the example 4.1.4 above, all the lines in $\mathbf{R}^2$ with slope 2 are parallel to U.
- If $U = \{(x, y, 0) \in \mathbf{R}^3 : x, y \in \mathbf{R}\}$, then the affine subsets of $\mathbf{R}^3$ parallel to U are the planes in $\mathbf{R}^3$ that are parallel to the xy-plane U in the usual sense.
- If $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$, then $\mathbf{R}^2/U$ is the set of all lines in $\mathbf{R}^2$ that have slope 2.
- If U is a line in $\mathbf{R}^3$ containing the origin, then $\mathbf{R}^3/U$ is the set of all lines in $\mathbf{R}^3$ parallel to U.
- If U is a plane in $\mathbf{R}^3$ containing the origin, then $\mathbf{R}^3/U$ is the set of all planes in $\mathbf{R}^3$ parallel to U.

Proposition 4.1.5 (Two affine subsets parallel to U are equal or disjoint)**.** Suppose U is a subspace of V and $v, w \in V$. Then the following are equivalent:

1. $v - u \in U$
2. $v + U = w + U$
3. $(v + U) \cap (w + U) \neq \emptyset$

Proof. First suppose (1) holds, so $v - u \in U$. If $u \in U$, then

$$v + u = w + ((v - u) + u) \in w + U$$

Thus $v+U \subset w+U$. Similarly, $w+U \subset v+U$. Thus $v+U = w+U$. Completing the proof that (1) implies (2). Obviously (2) implies (3).

Now suppose (3) holds, so $(v+U)\cap(w+U) \neq \emptyset$. Thus there exist $u_1, u_2 \in U$ such that

$$v + u_1 = w + u_2.$$

Thus $v - w = u_2 - u_1$. Hence $v - w \in U$, showing that (3) implies (1) and completing the proof. $\square$

Definition 4.1.4 (addition and scalar multiplication on V/U)**.** Suppose U is a subspace of V. Then **addition** and **scalar multiplication** are defined on V/U by

$$(v+U)+(w+U)=(v+w)+U$$
$$\lambda(v+U)=(\lambda v)+U$$

for all $v, w \in V$ and $\lambda \in \mathbf{F}$.

Proposition 4.1.6 (Quotient space is a vector space)**.** Suppose U is a subspace of V. Then V/U, with the operations of addition and scalar multiplication as defined above, is a vector space.

Proof. The potential problem with the definitions above of addition and scalar multiplication on V/U is that the representation of an affine subset parallel to U is not unique. Specifically, suppose $v, w \in V$. Suppose also that $\hat{v}, \hat{w} \in V$ are such that $v+U=\hat{v}+U$ and $w+U=\hat{w}+U$. To show that the definition of addition on V/U given above makes sense, we must show that $(v+w)+U=(\hat{v}+\hat{w})+U$.

By Proposition 4.1.5, we have

$$v-\hat{v} \in U \text{ and } w-\hat{w} \in U$$

Because U is a subspace of V and thus is closed under addition, this implies that $(v-\hat{v})+(w-\hat{w}) \in U$. Thus $(v+w)-(\hat{v}+\hat{w}) \in U$. Using Proposition 4.1.5 again, we see that

$$(v+w)+U=(\hat{v}+\hat{w})+U$$

as desired. Thus the definition of addition on V/U makes sense.

Similarly, suppose $\lambda \in \mathbf{F}$. Because U is a subspace of V and thus is closed under scalar multiplication, we have $\lambda(v-\hat{v}) \in U$. Thus $\lambda v-\lambda\hat{v} \in U$. Hence Proposition 4.1.5 implies that $(\lambda v)+U=(\lambda\hat{v})+U$. Thus the definition of scalar multiplication on V/U makes sense.

Now that addition and scalar multiplication have been defined on V/U, the verification that these operations make V/U into a vector space is straightforward and is left to the reader. Note that the additive identity of V/U is $0+U$ (which equals U) and that the additive inverse of $v+U$ is $(-v)+U$. □

4.1.3 Quotient map and $\tilde{T}$

Definition 4.1.5 (quotient map, π)**.** Suppose U is a subspace of V. The **quotient map** π is the linear map $\pi : V \to V/U$ defined by

$$\pi(v)=v+U$$

for all $v \in V$. The reader should verify that π is indeed a linear map.

Proposition 4.1.7 (Dimension of a quotient space)**.** Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim V/U = \dim V - \dim U.$$

Proof. Let π be the quotient map from V to V/U. We see that null $\pi = U$. Clearly range $\pi = V/U$. The Fundamental Theorem of Linear Maps thus tells us that

$$\dim V = \dim \operatorname{null} \pi + \dim \operatorname{range} \pi,$$

which gives the desired result. □

Definition 4.1.6 ($\tilde{T}$)**.** Suppose $T \in \mathcal{L}(V, W)$. Define $\tilde{T} : V/(\text{null } T) \to W$ by

$$\tilde{T}(v + \text{null } T) = Tv.$$

To show that the definition of $\tilde{T}$ makes sense, suppose $u, v \in V$ are such that $u + \text{null } T = v + \text{null } T$. By Proposition 4.1.5, we have $u - v \in \text{null } T$. Thus $T(u - v) = 0$. Hence $Tu = Tv$. Thus the definition of $\tilde{T}$ indeed makes sense.

Theorem 4.1.1 (Structure of linear maps)**.** Suppose $T \in \mathcal{L}(V, W)$. Then

1. $\tilde{T}$ is a linear map from $V/(\text{null } T)$ to W;
2. $\tilde{T}$ is injective;
3. range $\tilde{T}$ = range T;
4. $V/(\text{null } T)$ is isomorphic to range T.
5. We can always factorize T into:

$$V \xrightarrow{\text{surjection } \pi} V/(\text{null } T) \xrightarrow{\text{isomorphism } \tilde{T}} \text{range } T \xrightarrow{\text{injection}} W$$

Proof. Omitted. □

4.2 Exercises

1. Suppose T is a **function** from V to W. The **graph** of T is the subset of $V \times W$ defined by

$$\text{graph } T = \{(v, Tv) \in V \times W : v \in V\}.$$

   Prove that T is a linear map if and only if the graph of T is a subspace of $V \times W$.

2. Suppose $V_1, \ldots, V_m$ are vector spaces such that $V_1 \times \cdots \times V_m$ is finite-dimensional. Prove that V_j is finite-dimensional for each $j = 1, \ldots, m$.
3. Give an example of a vector space V and subspaces U_1, U_2 of V such that $U_1 \times U_2$ is isomorphic to $U_1 + U_2$ but $U_1 + U_2$ is not a direct sum. (Hint: Use infinite-dimensional spaces)

4. Suppose $V_1, \dots, V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.
5. Suppose $W_1, \dots, W_m$ are vector spaces. Prove that $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.
6. For n a positive integer, define V^n by

$$V^n = \underbrace{V \times \cdots \times V}_{n \text{ times}}.$$

   Prove that V^n and $\mathcal{L}(\mathbf{F}^n, V)$ are isomorphic vector spaces.
7. Suppose v, x are vectors in V. U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.
8. Prove that a nonempty subset A of V is an affine subset of V if and only if $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$.
9. Suppose A_1 and A_2 are affine subsets of V. Prove that the intersection $A_1 \cap A_2$ is either an affine subset of V or the empty set.
10. Prove that the intersection of every collection of affine subsets of V is either an affine subset of V or the empty set.
11. Suppose $v_1, \dots, v_m \in V$. Let

$$A = \{\sum_{j=1}^{m} \lambda_j v_j : \lambda_1, \dots, \lambda_m \in \mathbf{F} \text{ and } \sum_{j=1}^{m} \lambda_j = 1\}.$$

    (a) Prove that A is an affine subset of V.
    (b) Prove that every affine subset of V that contains $v_1, \dots, v_m$ also contains A.
    (c) Prove that $A = v + U$ for some $v \in V$ and some subspace U of V with $\dim U \leq m - 1$.
12. Suppose U is a subspace of V such that V/U is finite-dimensional. Prove that V is isomorphic to $U \times (V/U)$.
13. Suppose U is a subspace of V and $v_1 + U, \dots, v_m + U$ is a basis of V/U and $u_1, \dots, u_n$ is a basis of U. Prove that $v_1, \dots, v_m, u_1, \dots, u_n$ is a basis of V.
14. Suppose $\mathbf{F}^{\oplus\infty} = \{(x_1, x_2, \dots) \in \mathbf{F}^\infty : x_j \neq 0 \text{ for only finitely many } j\}$.
    (a) Show that $\mathbf{F}^{\oplus\infty}$ is a subspace of $\mathbf{F}^\infty$.
    (b) Prove that $\mathbf{F}^\infty / \mathbf{F}^{\oplus\infty}$ is infinite-dimensional.
15. Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Prove that $\dim V/(\text{null } \varphi) = 1$.
16. Suppose U is a subspace of V such that $\dim V/U = 1$. Prove that there exists $\varphi \in \mathcal{L}(V, \mathbf{F})$ such that null $\varphi = U$.

17. Suppose U is a subspace of V such that V/U is finite-dimensional. Prove that there exists a subspace W of V such that $\dim W = \dim V/U$ and $V = U \oplus W$.
18. Suppose $T \in \mathcal{L}(V, W)$ and U is a subspace of V. Let π denote the quotient map from V onto V/U. Prove that there exists $S \in \mathcal{L}(V/U, W)$ such that $T = S \circ \pi$ if and only if $U \subset \operatorname{null} T$.
19. Suppose U is a subspace of V. Define $\Gamma : \mathcal{L}(V/U, W) \to \mathcal{L}(V, W)$ by

$$\Gamma(S) = S \circ \pi.$$

    (a) Show that Γ is a linear map.
    (b) Show that Γ is injective.
    (c) Show that $\operatorname{range} \Gamma = \{T \in \mathcal{L}(V, W) : Tu = 0 \text{ for every } u \in U\}$.

4.3 Duality

4.3.1 The dual space

Definition 4.3.1 (linear functional)**.** A **linear functional** on V is a linear map from V to $\mathbf{F}$. In other words, a linear functional is an element of $\mathcal{L}(V, \mathbf{F})$.

Example 4.3.1 (linear functionals)**.** Fix $(c_1, \ldots, c_n) \in \mathbf{F}^n$. Define $\varphi : \mathbf{F}^n \to \mathbf{F}$ by

$$\varphi(x_1, \ldots, x_n) = \sum_{j=1}^{n} c_j x_j.$$

Then φ is a linear functional on $\mathbf{F}^n$.

Definition 4.3.2 (dual space, V', $\dim V' = \dim V$)**.** The **dual space** of V, denoted V', is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V, \mathbf{F})$. Suppose V is finite-dimensional. Then V' is also finite-dimensional and $\dim V' = (\dim V)(\dim \mathbf{F}) = \dim V$.

Definition 4.3.3 (dual basis)**.** If $v_1, \ldots, v_n$ is a basis of V, then the **dual basis** of $v_1, \ldots, v_n$ is the list $\varphi_1, \ldots, \varphi_n$ of elements of V', where each φ_j is the linear functional on V such that

$$\varphi_j(v_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

Example 4.3.2. What is the dual basis of the standard basis $e_1, \ldots, e_n$ of $\mathbf{F}^n$?

Proof. For $1 \leq j \leq n$, define φ_j to be the linear functional on $\mathbf{F}^n$ that selects the j^{th} coordinate of a vector in $\mathbf{F}^n$. In other words,

$$\varphi_j(x_1, \ldots, x_n) = x_j$$

for $(x_1, \ldots, x_n) \in \mathbf{F}^n$. Clearly

$$\varphi_j(e_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

Thus $\varphi_1, \ldots, \varphi_n$ is the dual basis of the standard basis $e_1, \ldots, e_n$ of $\mathbf{F}^n$. $\square$

Proposition 4.3.1 (Dual basis is a basis of the dual space)**.** Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.

Proof. Suppose $v_1, \ldots, v_n$ is a basis of V. Let $\varphi_1, \ldots, \varphi_n$ denote the dual basis.To show that $\varphi_1, \ldots, \varphi_n$ is a linearly independent list of elements of V', suppose $a_1, \ldots, a_n \in \mathbf{F}$ are such that

$$\sum_{j=1}^{n} a_j \varphi_j = 0$$

Now $\left(\sum_{j=1}^{n} a_j \varphi_j\right)(v_j) = a_j$ for $j = 1, \ldots, n$. The equation above thus shows that $a_1 = \cdots = a_n = 0$. Hence $\varphi_1, \ldots, \varphi_n$ is linearly independent. Since $\dim V' = \dim V$, $\varphi_1, \ldots, \varphi_n$ is a basis of V'. $\square$

4.3.2 The dual map

We know that $\dim \mathcal{L}(V, W) = \dim \mathcal{L}(W, V)$. Given $T \in \mathcal{L}(V, W)$, is there some canonical $T' \in \mathcal{L}(W, V)$ which corresponds to T? Unfortunately, there is no such a correspondence. But we have $T' \in \mathcal{L}(W', V')$, which is the dual of T:

Definition 4.3.4 (dual map, T')**.** If $T \in \mathcal{L}(V, W)$, then the **dual map** of T is the linear map $T' \in \mathcal{L}(W', V')$ defined by $T'(\varphi) = \varphi \circ T$ for $\varphi \in W'$. In other words,

$$\varphi = \left(W \xrightarrow{\varphi} \mathbf{F}\right) \mapsto \left(V \xrightarrow{T} W \xrightarrow{\varphi} \mathbf{F}\right) = T'(\varphi)$$

$T'(\varphi)$ is defined above to be the composition of the linear maps φ and T. Thus $T'(\varphi)$ is indeed a linear map from V to $\mathbf{F}$. In other words, $T'(\varphi) \in V'$.

The verification that T' is a linear map from W' to V' is easy:

- If $\varphi, \psi \in W'$, then

$$T'(\varphi + \psi) = (\varphi + \psi) \circ T = \varphi \circ T + \psi \circ T = T'(\varphi) + T'(\psi).$$

- If $\lambda \in \mathbf{F}$ and $\varphi \in W'$, then

$$T'(\lambda\varphi) = (\lambda\varphi) \circ T = \lambda(\varphi \circ T) = \lambda T'(\varphi).$$

Proposition 4.3.2 (Algebraic properties of dual maps)**.** The first two bullet points in the result below imply that the function $\cdot' : \mathcal{L}(V, W) \xrightarrow{T \mapsto T'} \mathcal{L}(W'V')$ that takes T to T' is a linear map from $\mathcal{L}(V, W)$ to $\mathcal{L}(W'V')$

- $(S+T)' = S' + T'$ for all $S, T \in \mathcal{L}(V, W)$.
- $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$ and all $T \in \mathcal{L}(V, W)$.
- $(ST)' = T'S'$ for all $T \in \mathcal{L}(U, V)$ and all $\mathcal{L}(V, W)$.

Proof. We only prove the third bullet point, suppose $\varphi \in W'$. Then

$$(ST)'\varphi = \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = (T'S')(\varphi).$$

where the first, third, and fourth equalities above hold because of the definition of the dual map, the second equality holds because composition of functions is associative, and the last equality follows from the definition of composition. The equality of the first and last terms above for all $\varphi \in W'$ means that $(ST)' = T'S'$. $\square$

4.3.3 The null space and range of the dual of a linear map

Definition 4.3.5 (annihilator, U_V^0)**.** For $U \subset V$, the **annihilator** of U, denoted U_V^0, is defined by

$$U_V^0 = \{\varphi \in V' : \varphi(u) = 0 \text{ for all } u \in U\}.$$

Example 4.3.3. Let e_1, e_2, e_3, e_4, e_5 denote the standard basis of $\mathbf{R}^5$, and let $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5$ denote the dual basis of $(\mathbf{R}^5)'$. Suppose

$$U = \operatorname{span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbf{R}^5 : x_1, x_2 \in \mathbf{R}\}.$$

Show that $U_{\mathbf{R}^5}^0 = \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$.

Proof. Recall that φ_j is the linear functional on $\mathbf{R}^5$ that selects that j^{th} coordinate: $\varphi_j(x_1, x_2, x_3, x_4, x_5) = x_j$.

First suppose $\varphi \in \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$. Then there exists $c_3, c_4, c_5 \in \mathbf{R}$ such that $\varphi = c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5$. If $(x_1, x_2, 0, 0, 0) \in U$, then

$$\varphi(x_1, x_2, 0, 0, 0) = (c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5)(x_1, x_2, 0, 0, 0) = 0.$$

Thus $\varphi \in U_{\mathbf{R}^5}^0$. In other words, we have shown that $\operatorname{span}(\varphi_3, \varphi_4, \varphi_5) \subset U_{\mathbf{R}^5}^0$.

To show the inclusion in the other direction, suppose $\varphi \in U_{\mathbf{R}^5}^0$. Because the dual basis is a basis of $(\mathbf{R}^5)'$, there exists $c_1, c_2, c_3, c_4, c_5 \in \mathbf{R}$ such that $\varphi = \sum_{j=1}^5 c_j\varphi_j$. Because $e_1 \in U$ and $\varphi \in U_{\mathbf{R}^5}^0$, we have

$$0 = \varphi(e_1) = \left(\sum_{j=1}^5 c_j\varphi_j\right)(e_1) = c_1$$

Similarly, $e_2 \in U$ and thus $c_2 = 0$. Hence $\varphi = c_3\varphi_3 + c_4\varphi_4 + c_5\varphi_5 \in \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$. Which shows that $U_{\mathbf{R}^5}^0 \subset \operatorname{span}(\varphi_3, \varphi_4, \varphi_5)$. $\square$

Proposition 4.3.3 (The annihilator is a subspace)**.** Suppose $U \subset V$, then U_V^0 is a subspace of V'.

Proof. Omitted. $\square$

Theorem 4.3.1 (Dual of inclusion)**.** Suppose U is a subspace of V, let $i \in \mathcal{L}(U, V)$ be the inclusion map defined by $i(u) = u$ for $u \in U$. Then i' is a linear map from V' to U'. Then we have

$$\text{null } i' = U_V^0$$

If V is finite-dimensional, we have

$$\text{range } i' = U' \text{ (So } i' \text{ is surjective)}$$

Proof.

$$\text{null } i' = \{\varphi \in V' : i'\varphi = 0\} = \{\varphi \in V' : \varphi \circ i = 0\}$$

But $\varphi \circ i \in \mathcal{L}(U, \mathbf{F})$, so $\varphi \circ i = 0$ if and only if $\varphi(i(u)) = \varphi(u) = 0$ for all $u \in U$. And we have

$$\{\varphi \in V' : \varphi \circ i = 0\} = \{\varphi \in V' : \varphi(u) = 0 \text{ for all } u \in U\} = U_V^0$$

By Exercise 3.2.9, if $\varphi \in U'$, then φ can be extended to a linear functional ψ on V. The definition of i' show that $i'(\psi) = \varphi$. Thus $\varphi \in \text{range } i'$, which implies that range $i' = U'$. $\square$

Proposition 4.3.4 (Dimension of the annihilator)**.** Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U + \dim U_V^0 = \dim V.$$

Proof. The Fundamental Theorem of Linear Maps applied to i' shows that

$$\dim \text{range } i' + \dim \text{null } i' = \dim V',$$

which is just

$$\dim U' + \dim U_V^0 = \dim V',$$

which is just

$$\dim U + \dim U_V^0 = \dim V.$$

$\square$

Theorem 4.3.2 (The null space and range of T')**.** Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

1. null $T' = (\text{range } T)_W^0$
2. $\dim \text{null } T' = \dim \text{null } T + \dim W - \dim V$
3. $\dim \text{range } T' = \dim \text{range } T$
4. range $T' = (\text{null } T)_V^0$

Proof. The proof of part (1) of the result below does not use the hypothesis that V and W are finite-dimensional.

1. First suppose $\varphi \in \text{null } T'$. Thus $0 = T'(\varphi) = \varphi \circ T$. Hence
$$0 = (\varphi \circ T)(v) = \varphi(Tv) \text{ for every } v \in V.$$
Thus $\varphi \in (\text{range } T)^0_W$. This implies that $\text{null } T' \subset (\text{range } T)^0_W$ To prove the inclusion in the opposite direction, now suppose that $\varphi \in (\text{range } T)^0_W$. Thus $\varphi(Tv) = 0$ for every vector $v \in V$. Hence $0 = \varphi \circ T = T'(\varphi)$. In other words, $\varphi \in \text{null } T'$, which show that $(\text{range } T)^0_W \subset \text{null } T'$, completing the proof of (1).
2. We have
$$\begin{aligned}\dim \text{null } T' &= \dim(\text{range } T)^0_W \\ &= \dim W - \dim \text{range } T \\ &= \dim W - (\dim V - \dim \text{null } T).\end{aligned}$$
3. We have
$$\begin{aligned}\dim \text{range } T' &= \dim W' - \dim \text{null } T' \\ &= \dim W - \dim(\text{range } T)^0_W \\ &= \dim \text{range } T.\end{aligned}$$
4. First suppose $\varphi \in \text{range } T'$. Thus there exists $\psi \in W'$ such that $\varphi = T'(\psi)$. If $v \in \text{null } T$, then
$$\varphi(v) = (T'(\psi))v = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0.$$
Hence $\varphi \in (\text{null } T)^0_V$. This implies that $\text{range } T' \subseteq (\text{null } T)^0_V$.

   We will complete the proof by showing that $\text{range } T'$ and $(\text{null } T)^0_V$ have the same dimension. To do this, note that
$$\begin{aligned}\dim \text{range } T' &= \dim \text{range } T \\ &= \dim V - \dim \text{null } T \\ &= \dim(\text{null } T)^0_V.\end{aligned}$$

□

Proposition 4.3.5 (T surjective is equivalent to T' injective)**.** Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is surjective if and only if T' is injective.

Proof.
$$\text{range } T = W \iff (\text{range } T)^0_W = \{0\} \iff \text{null } T' = \{0\}$$

□

Proposition 4.3.6 (T injective is equivalent to T' surjective)**.** Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is injective if and only if T' is surjective.

Proof.

$$\text{null } T = \{0\} \iff (\text{null } T)^0_V = W' \iff \text{range } T' = V'$$

$\square$

The setting for the next result is the assumption that we have a basis $v_1, \ldots, v_n$ of V, along with its dual basis $\varphi_1, \ldots, \varphi_n$ of V'. We also have a basis $w_1, \ldots, w_m$ of W, along with its dual basis $\psi_1, \ldots, \psi_m$ of W'.

Proposition 4.3.7. Suppose $T \in \mathcal{L}(V, W)$. Let

$$Tv_k = \sum_{s=1}^{m} A_{s,k} w_s, T'(\psi_j) = \sum_{r=1}^{n} C_{r,j} \varphi_r$$

Then $C_{k,j} = A_{j,k}$ for every $j = 1, \ldots, m$ and every $k = 1, \ldots, n$.

Proof. We have $T'(\psi_j) = \psi_j \circ T$, thus

$$(T'(\psi_j))(v_k) = \sum_{r=1}^{n} C_{r,j} \varphi_r(v_k) = C_{k,j}$$

but we also have

$$(\psi_j \circ T)(v_k) = \psi_j(Tv_k) = \psi_j \left(\sum_{s=1}^{m} A_{s,k} w_s \right) = A_{j,k}$$

Hence we have

$$A_{j,k} = (\psi_j \circ T)(v_k) = (T'(\psi_j))(v_k) = C_{k,j}$$

$\square$

4.4 Exercises

1. Explain why every linear functional is either surjective or the zero map.
2. Give three distinct examples of linear functionals on $\mathbf{R}^{[0,1]}$
3. Suppose V is finite-dimensional and $v \in V$ with $v \neq 0$. Prove that there exists $\varphi \in V'$ such that $\varphi(v) = 1$.
4. Suppose V is finite-dimensional and U is a subspace of V such that $U \neq V$. Prove that there exists $\varphi \in V'$ such that $\varphi(u) = 0$ for every $u \in U$ but $\varphi \neq 0$.

5. Suppose $V_1, \ldots, V_m$ are vector spaces. Prove that $(V_1 \times \cdots \times V_m)'$ and $V_1' \times \cdots \times V_m'$ are isomorphic vector spaces
6. Suppose V is finite-dimensional and $v_1, \ldots, v_m \in V$. Define a linear map $\Gamma : V' \to \mathbf{F}^m$ by
$$\Gamma(\varphi) = (\varphi(v_1), \ldots, \varphi(v_m)).$$
   (a) Prove that $v_1, \ldots, v_m$ spans V if and only if Γ is injective.
   (b) Prove that $v_1, \ldots, v_m$ is linearly independent if and only if Γ is surjective.
7. Suppose m is a positive integer.
   (a) Show that $1, x-5, \ldots, (x-5)^m$ is a basis of $\mathcal{P}_m(\mathbf{R})$.
   (b) What is the dual basis of the basis in part (a)?
8. Suppose $v_1, \ldots, v_n$ is a basis of V and $\varphi_1, \ldots, \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that
$$\psi = \sum_{j=1}^{n} \psi(v_j)\varphi_j.$$
9. Suppose $T \in \mathcal{L}(V, W)$. Choose a basis $v_1, \ldots, v_n$ for V and a basis $w_1, \ldots, w_m$ for W. Let
$$Tv_k = \sum_{j=1}^{m} A_{j,k} w_j$$
Prove that $\dim \operatorname{range} T = 1$ if and only if there exists $(c_1, \ldots, c_m) \in \mathbf{F}^m$ and $(d_1, \ldots, d_n) \in \mathbf{F}^n$ such that $A_{j,k} = c_j d_k$ for every $j = 1, \ldots, m$ and every $k = 1, \ldots, n$.
10. Show that the dual map of the identity map on V is the identity map on V'.
11. Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x, y, z) = (4x + 5y + 6z, 7x + 8y + 9z)$. Suppose φ_1, φ_2 denotes the dual basis of the standard basis of $\mathbf{R}^2$ and ψ_1, ψ_2, ψ_3 denote the dual basis of the standard basis of $\mathbf{R}^3$.
    (a) Describe the linear functionals $T'(\varphi_1)$ and $T'(\varphi_2)$.
    (b) Write $T'(\varphi_1)$ and $T'(\varphi_2)$ as linear combinations of ψ_1, ψ_2, ψ_3.
12. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that $T' = 0$ if and only if $T = 0$.
13. Suppose V and W are finite-dimensional. Prove that the map $\cdot'$ that takes $T \in \mathcal{L}(V, W)$ to $T' \in \mathcal{L}(W', V')$ is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathcal{L}(W', V')$.
14. Suppose $U \subset V$. Explain why $U_V^0 = \{\varphi \in V' : U \subset \operatorname{null} \varphi\}$.

15. Suppose V is finite-dimensional and $U \subset V$ is nonempty. Show that $U = \{0\}$ if and only if $U^0_V = V'$.
16. Suppose V is finite-dimensional and U is a subspace of V. Show that $U = V$ if and only if $U^0_V = \{0\}$.
17. Suppose U and W are subsets of V with $U \subset W$. Prove that $W^0_V \subset U^0_V$.
18. Suppose V is finite-dimensional and U and W are subspaces of V with $W^0_V \subset U^0_V$. Prove that $U \subset W$.
19. Suppose U, W are subspaces of V. Show that $(U + W)^0_V = W^0_V \cap U^0_V$
20. Suppose V is finite-dimensional and U and W are subspaces of V. Prove that $(U \cap W)^0_V = W^0_V + U^0_V$.
21. Suppose V is finite-dimensional and U is a subspace of V. Show that

$$U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0_V\}$$

22. Suppose V is finite-dimensional and Γ is a subspace of V'. Show that

$$\Gamma = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Gamma\}^0_V$$

23. Suppose $T \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_5(\mathbf{R}))$ and null $T' = \operatorname{span}(\varphi)$, where φ is the linear functional on $\mathcal{P}_5(\mathbf{R})$ defined by $\varphi(p) = p(8)$. Prove that range $T = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}$.
24. Suppose V and W are finite-dimensional, $T \in \mathcal{L}(V, W)$, and there exists $\varphi \in W'$ such that null $T' = \operatorname{span}(\varphi)$. Prove that range T = null φ.
25. Suppose V and W are finite-dimensional, $T \in \mathcal{L}(V, W)$, and there exists $\varphi \in V'$ such that range $T' = \operatorname{span}(\varphi)$. Prove that null T = null φ.
26. Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m$ is a linearly independent list in V'. Prove that

$$\dim \bigcap_{j=1}^{m} \operatorname{null} \varphi_j = \dim V - m$$

27. The **double dual space** of V, denoted V'', is defined to be the dual space of V'. In other words, $V'' = (V')'$. Define $\Lambda : V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for $v \in V$ and $\varphi \in V'$.

(a) Show that Λ is a linear map from V to V''.
(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.
(c) Show that if V if finite-dimensional, then Λ is an isomorphism from V onto V''.

Remark: Suppose V is finite-dimensional. Then V and V' are isomorphic, but finding an isomorphism from V onto V' generally requires choosing a basis of V. In contrast, the isomorphism Λ from V onto V'' does not require a choice of basis and thus is considered more natural.

28. Show that $(\mathcal{P}(\mathbf{R}))'$ and $\mathbf{R}^\infty$ are isomorphic.
29. Suppose U is a subspace of V. Let $i : U \to V$ be the inclusion map. Describe $\tilde{i'} : V'/U^0_V \to U'$.
30. Suppose U is a subspace of V. Let $\pi : V \to V/U$ be the usual quotient map. Prove that $\pi' : (V/U)' \to U^0_V$ is an isomorphism.

Chapter 5

Eigenvalues, Eigenvectors, and Invariant Subspaces

5.1 Polynomials

5.1.1 Complex conjugate and absolute value

Definition 5.1.1 ($\Re z, \Im z$)**.** Suppose $z = a+bi$, where a and b are real numbers.

- The **real part** of z, denoted $\Re z$, is defined by $\Re z = a$.
- The **imaginary part** of z, denoted $\Im z$, is defined by $\Im z = b$.

Thus for every complex number z, we have

$$z = \Re z + (\Im z)i.$$

Definition 5.1.2 (complex conjugate, $\bar{z}$, absolute value, $|z|$)**.** Suppose $z \in \mathbf{C}$.

- The **complex conjugate** of $z \in \mathbf{C}$, denoted $\bar{z}$, is defined by

$$\bar{z} = \Re z - (\Im z)i.$$

- The **absolute value** of a complex number z, denoted $|z|$, is defined by

$$|z| = \sqrt{(\Re z)^2 + (\Im z)^2}.$$

Note that $|z|$ is a nonnegative number for every $z \in \mathbf{C}$.

Proposition 5.1.1 (Properties of complex numbers)**.** Suppose $w, z \in \mathbf{C}$. Then

sum of z and $\bar{z}$: $z + \bar{z} = 2\Re z$

difference of z and $\bar{z}$: $z - \bar{z} = 2(\Im z)i$

product of z and $\bar{z}$: $z\bar{z} = |z|^2$

additivity and multiplicativity: $\overline{w+z} = \bar{w} + \bar{z}, \overline{wz} = \bar{w}\bar{z}$

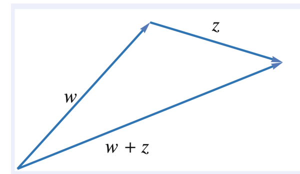
conjugate of conjugate: $\bar{\bar{z}} = z$

bounded by $|z|$: $|\Re z| \leq |z|, |\Im z| \leq |z|$

absolute value: $|\bar{z}| = |z|$

multiplicativity of absolute value: $|wz| = |w||z|$

Triangle Inequality: $|w+z| \leq |w| + |z|$



5.1.2 Uniqueness of coefficients for polynomials

Recall that a function $p : \mathbf{R} \to \mathbf{R}$ is called a polynomial with coefficients in $\mathbf{F}$ if there exist $a_0, \dots, a_m \in \mathbf{F}$ such that

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$.

Proposition 5.1.2 (If a polynomial is the zero function, then all coefficients are 0)**.** Suppose $a_0, \dots, a_m \in \mathbf{F}$. If

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m = 0$$

for every $z \in \mathbf{F}$, then $a_0 = \cdots = a_m = 0$.

Proof. We will prove the contrapositive. If not all the coefficients are 0, then by changing m we can assume $a_m \neq 0$. Let

$$z = \frac{|a_0| + |a_1| + \cdots + |a_{m-1}|}{|a_m|} + 1$$

Note that $z \geq 1$, and thus $z^j \leq z^{m-1}$ for $j = 0, 1, \dots, m-1$. Using the triangle inequality, we have

$$|a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1}| \leq (|a_0| + |a_1| + \cdots + |a_{m-1}|) z^{m-1} < |a_m z^m|.$$

Thus $a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} \neq -a_m z^m$. Hence we conclude that $a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m \neq 0$. □

The result above implies that the coefficients of a polynomial are uniquely determined (because if a polynomial had two different sets of coefficients, then subtracting the two representations of the polynomial would give a contradiction to the result above).

Recall that if a polynomial p can be written in the form

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

with $a_m \neq 0$, then we say that p has degree m and we write $\deg p = m$. The degree of the 0 polynomial is defined to be $-\infty$. When necessary, use the obvious arithmetic with ∞. For example, $-\infty < m$ and $-\infty + m = -\infty$ for every integer m.

5.1.3 The division algorithm for polynomials

Recall that $\mathcal{P}(\mathbf{F})$ denotes the vector space of all polynomials with coefficients in $\mathbf{F}$ and that $\mathcal{P}_m(\mathbf{F})$ is the subspace of $\mathcal{P}(\mathbf{F})$ consisting of the polynomials with coefficients in $\mathbf{F}$ and degree at most m.

Theorem 5.1.1 (Division Algorithm for Polynomials)**.** Suppose that $p, s \in \mathcal{P}(\mathbf{F})$, with $s \neq 0$. Then there exist unique polynomials $q, r \in \mathcal{P}(\mathbf{F})$ such that

$$p = sq + r$$

and $\deg r < \deg s$.

Proof. Let $n = \deg p$ and $m = \deg s$. If $n < m$, then take $q = 0$ and $r = p$ to get the desired result. Thus we can assume that $n \geq m$.

Define $T : \mathcal{P}_{n-m}(\mathbf{F}) \times \mathcal{P}_{m-1}(\mathbf{F}) \to \mathcal{P}_n(\mathbf{F})$ by

$$T(q, r) = sq + r.$$

The reader can easily verify that T is a linear map. If $(q, r) \in \operatorname{null} T$, then $sq + r = 0$, which implies that $q = 0$ and $r = 0$. Because otherwise $\deg sq \geq m$ and thus sq cannot equal $-r$. Thus $\dim \operatorname{null} T = 0$, proving the unique part of the result. Also, we know that T is injective (hence surjective).

We have

$$\dim(\mathcal{P}_{n-m}(\mathbf{F}) \times \mathcal{P}_{m-1}(\mathbf{F})) = \dim \mathcal{P}_{n-m}(\mathbf{F}) + \dim \mathcal{P}_{m-1}(\mathbf{F}) = (n-m+1) + (m-1+1) = n+1$$

The Fundamental Theorem of Linear Maps and the equation displayed above now imply that $\dim \operatorname{range} T = n + 1$, which equals $\dim \mathcal{P}_n(\mathbf{F})$. Thus $\operatorname{range} T = \mathcal{P}_n(\mathbf{F})$, and hence there exist $q \in \mathcal{P}_{n-m}(\mathbf{F})$ and $r \in \mathcal{P}_{m-1}(\mathbf{F})$ such that $p = T(q, r) = sq + r$. $\square$

5.1.4 Zeros of polynomials

Definition 5.1.3 (zero of a polynomial)**.** A number $\lambda \in \mathbf{F}$ is called a **zero** (or **root**) of a polynomial $p \in \mathcal{P}(\mathbf{F})$ if

$$p(\lambda) = 0.$$

Definition 5.1.4 (factor)**.** A polynomial $s \in \mathcal{P}(\mathbf{F})$ is called a **factor** of $p \in \mathcal{P}(\mathbf{F})$ if there exists a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that $p = sq$.

Proposition 5.1.3 (Each zero of a polynomial corresponds to a degree-1 factor)**.** Suppose $p \in \mathcal{P}(\mathbf{F})$ and $\lambda \in \mathbf{F}$. Then $p(\lambda) = 0$ if and only if there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that

$$p(z) = (z - \lambda)q(z)$$

for every $z \in \mathbf{F}$.

Proof. Suppose $p(\lambda) = 0$. The polynomial $z - \lambda$ has degree 1. Because a polynomial with degree less than 1 is a constant function, the Division Algorithm for Polynomials implies that there exist a polynomial $q \in \mathcal{P}(\mathbf{F})$ and a number $r \in \mathbf{F}$ such that

$$p(z) = (z - \lambda)q(z) + r$$

for every $z \in \mathbf{F}$. The equation above and the equation $p(\lambda) = 0$ imply that $r = 0$. □

Proposition 5.1.4 (A polynomial has at most as many zeros as its degree)**.** Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial with degree $m \geq 0$. Then p has at most m distinct zeros in $\mathbf{F}$.

Proof. We use induction on m, assuming that every polynomial with degree $m - 1$ has at most $m - 1$ distinct zeros. □

5.2 Fundamental Theorem of Algebra

The next result, although called the Fundamental Theorem of **Algebra**, uses analysis its proof. The Fundamental Theorem of Algebra is an existence theorem. Its proof does not lead to a method for finding zeros. The quadratic formula gives the zeros explicitly for polynomials of degree 2. Similar but more complicated formulas exist for polynomials of degree 3 and 4. **No such formulas exist for polynomials of degree 5 and above.**

Theorem 5.2.1 (Fundamental Theorem of Algebra)**.** Every nonconstant polynomial with complex coefficients has a zero.

Proof. Let p be a nonconstant polynomial with complex coefficients. Suppose p has no zeros. Then $1/p$ is a bounded analytic function on $\mathbf{C}$. By Liouville's theorem, $1/p$ is constant, a contradiction. □

Remarkably, mathematicians (Galois, Abel, Ruffini,...) have proved that no formula (using only $+, -, \times, \div, \sqrt[n]{}$) exists for the zeros of polynomials of degree 5 or higher. The real reason is that group $\mathfrak{S}_5$ is not solvable, so equations of degree 5 is not solvable by radicals.

Proposition 5.2.1 (cubic formula)**.** Suppose $p(x) = ax^3 + bx^2 + cx + d$ where $a \neq 0$. Set

$$u = \frac{9abc - 2b^3 - 27a^2d}{54a^3}, v = u^2 + \left(\frac{3ac - b^2}{9a^2}\right)^3$$

When $v \geq 0$, we have a root of p:

$$x = -\frac{b}{3a} + \sqrt[3]{u + \sqrt{v}} + \sqrt[3]{u - \sqrt{v}}$$

5.2.1 Factorization of polynomials

Proposition 5.2.2 (Factorization of a polynomial over $\mathbf{C}$)**.** If $p \in \mathcal{P}(\mathbf{C})$ is a nonconstant polynomial, then p has a unique factorization (except for the order of the factors) of the form

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where $c, \lambda_1, \ldots, \lambda_m \in \mathbf{C}$.

Proof. Let $p \in \mathcal{P}(\mathbf{C})$ and let $m = \deg p$. We will use induction on m. If $m = 1$, then clearly the desired factorization exists and is unique. So assume that $m > 1$ and that the desired factorization exists and is unique for all polynomials of degree $m - 1$.

First we will show that the desired factorization of p exists. By the Fundamental Theorem of Algebra, p has a zero λ. There is a polynomial q such that

$$p(z) = (z - \lambda)q(z)$$

for all $z \in \mathbf{C}$. Because $\deg q = m - 1$, our induction hypothesis implies that q has the desired factorization, which when plugged into the equation above gives the desired factorization of p.

Now we turn to the question of uniqueness. Clearly c is uniquely determined as the coefficient of z^m in p. So we need only show that except for the order, there is only one way to choose $\lambda_1, \ldots, \lambda_m$. If

$$(z - \lambda_1) \cdots (z - \lambda) = (z - \tau_1) \cdots (z - \tau_m)$$

for all $z \in \mathbf{C}$, then because the left side of the equation above equals 0 when $z = \lambda_1$, one of the τ's onn the right side equals λ_1. Relabeling, we can assume that $\tau_1 = \lambda_1$. Now for $z \neq \lambda_1$, we can divide both sides of the equation above by $z - \lambda_1$, getting

$$(z - \lambda_2) \cdots (z - \lambda) = (z - \tau_2) \cdots (z - \tau_m)$$

for all $z \in \mathbf{C}$ except possibly $z = \lambda_1$. Actually the equation above holds for all $z \in \mathbf{C}$, because otherwise by subtracting the right side from the left side we would get a nonzero polynomial that has infinitely many zeros. The equation above and our induction hypothesis imply that except for the order, the λ's are the same as the τ's, completing the proof of uniqueness. $\square$

Proposition 5.2.3 (Polynomials with real coefficients have zeros in pairs)**.** Suppose $p \in \mathcal{P}(\mathbf{C})$ is a polynomial with real coefficients. If $\lambda \in \mathbf{C}$ is a zero of p, then so is $\bar{\lambda}$.

Proof. Let

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$

where $a_0, \ldots, a_m$ are real numbers. Suppose $\lambda \in \mathbf{C}$ is a zero of p. Then

$$a_0 + a_1\lambda + \cdots + a_m\lambda^m = 0.$$

Take the complex conjugate of both sides of this equation, obtaining

$$a_0 + a_1\bar{\lambda} + \cdots + a_m\bar{\lambda}^m = 0.$$

The equation above shows that $\bar{\lambda}$ is a zero of p. $\square$

Proposition 5.2.4 (Factorization of a quadratic polynomial)**.** Suppose $b, c \in \mathbf{R}$. Then there is a polynomial factorization of the form

$$x^2 + bx + c = (x - \lambda_1)(x - \lambda_2)$$

with $\lambda_1, \lambda_2 \in \mathbf{R}$ if and only if $b^2 \geq 4c$.

Proof. Omitted. $\square$

Proposition 5.2.5 (Factorization of a polynomial over $\mathbf{R}$)**.** Suppose $p \in \mathcal{P}(\mathbf{R})$ is a nonconstant polynomial. Then p has a unique factorization (except for the order of the factors) of the form

$$p(x) = c(x - \lambda_1) \cdot (x - \lambda_m)(x^2 + b_1x + c_1) \cdots (x^2 + b_Mx + c_M),$$

where $c, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M \in \mathbf{R}$, with $b_j^2 < 4c_j$ for each j.

Proof. Think of p as an element of $\mathcal{P}(\mathbf{C})$. If all the (complex) zeros of p are real, then we are done. Thus suppose p has a zero $\lambda \in \mathbf{C}$ with $\lambda \notin \mathbf{R}$. By Proposition 5.2.3, $\bar{\lambda}$ is a zero of p. Thus we can write

$$p(x) = (x - \lambda)(x - \bar{\lambda})q(x) = (x^2 - 2(\Re\lambda)x + |\lambda|^2)q(x)$$

for some polynomial $q \in \mathcal{P}(\mathbf{C})$ with degree two less than the degree of p. If we can prove that q has real coefficients, then by using induction on the degree of p, we can conclude that $(x - \lambda)$ appears in the factorization of p exactly as many times as $(x - \bar{\lambda})$.

To prove that q has real coefficients, we solve the equation above for q, getting

$$q(x) = \frac{p(x)}{x^2 - 2(\Re\lambda)x + |\lambda|^2}$$

for all $x \in \mathbf{R}$. The equation above implies that $q(x) \in \mathbf{R}$ for all $x \in \mathbf{R}$. Writing

$$q(x) = a_0 + a_1x + \cdots + a_{n-2}x^{n-2},$$

where $n = \deg p$ and $a_0, \ldots, a_{n-2} \in \mathbf{C}$, we thus have

$$0 = \Im q(x) = (\Im a_0) + (\Im a_1)x + \cdots + (\Im a_{n-2})x^{n-2}$$

for all $x \in \mathbf{R}$. This implies that $\Im a_0, \ldots, \Im a_{n-2}$ all equal 0. Thus all the coefficients of q are real, as desired. Hence the desired factorization exists.

Now we turn to the question of uniqueness of our factorization. A factor of p of the form $x^2 + b_jx + c_j$ with $b_j^2 < 4c_j$ can be uniquely written as $(x - \lambda_j)(x - \bar{\lambda}_j)$ with $\lambda_j \in \mathbf{C}$. A moment's thought shows that two different factorizations of p as an element of $\mathcal{P}(\mathbf{R})$ would lead to two different factorizations of p as an element of $\mathcal{P}(\mathbf{C})$, contradicting Proposition 5.2.2. $\square$

5.3 Exercises

1. Suppose m is a positive integer. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p = m\}$$

a subspace of $\mathcal{P}(\mathbf{F})$?

2. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p \text{ is even}\}$$

a subspace of $\mathcal{P}(\mathbf{F})$?

3. Suppose m and n are positive integers with $m \leq n$, and suppose $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{F})$ with $\deg p = n$ such that $0 = p(\lambda_1) = \cdots = p(\lambda_m)$ and such that p has no other zeros.

4. Suppose m is a nonnegative integer, $z_1, \ldots, z_{m+1}$ are distinct elements of $\mathbf{F}$, and $w_1, \ldots, w_{m+1} \in \mathbf{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_m(\mathbf{F})$ such that

$$p(z_j) = w_j$$

for $j = 1, \ldots, m+1$.

5. Suppose $p \in \mathcal{P}(\mathbf{C})$ has degree m. Prove that p has m distinct zeros if and only if p and its derivative p' have no zeros in common.

6. Prove that every polynomial of odd degree with real coefficients has a real zero.

7. Define $T : \mathcal{P}(\mathbf{R}) \to \mathbf{R}^{\mathbf{R}}$ by

$$Tp = \begin{cases} \frac{p - p(3)}{x - 3} & \text{if } x \neq 3, \\ p'(3) & \text{if } x = 3. \end{cases}$$

Show that $Tp \in \mathcal{P}(\mathbf{R})$ for every polynomial $p \in \mathcal{P}(\mathbf{R})$ and that T is a linear map.

8. Suppose $p \in \mathcal{P}(\mathbf{C})$. Define $q : \mathbf{C} \to \mathbf{C}$ by

$$q(z) = p(z)\overline{p(\bar{z})}$$

Prove that q is a polynomial with real coefficients.

9. Suppose m is a nonnegative integer and $p \in \mathcal{P}_m(\mathbf{C})$ is such that there exist distinct real numbers $x_0, x_1, \ldots, x_m$ such that $p(x_j) \in \mathbf{R}$ for $j = 0, 1, \ldots, m$. Prove that all the coefficients of p are real.

10. Suppose $p \in \mathcal{P}(\mathbf{F})$ with $p \neq 0$. Let $U = \{pq : q \in \mathcal{P}(\mathbf{F})\}$.

    (a) Show that $\dim \mathcal{P}(\mathbf{F})/U = \deg p$.

    (b) Find a basis of $\mathcal{P}(\mathbf{F})/U$.

5.4 Invariant subspaces

Recall that an operator is a linear map from a vector space to itself. Recall also that we denote the set of operators on V by $\mathcal{L}(V)$; in other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Suppose $T \in \mathcal{L}(V)$. If we have a direct sum decomposition

$$V = \bigoplus_{j=1}^{m} U_j$$

where each U_j is a proper subspace of V, then to understand the behavior of T, we need only understand the behavior of each $T|_{U_j}$; here $T|_{U_j}$ denotes the restriction of T to the smaller domain U_j. Dealing with $T|_{U_j}$ should be easier than dealing with T because U_j is a smaller vector space than V.

However, if we intend to apply tools useful in the study of operators (such as taking powers), then we have a problem: $T|_{U_j}$ may not map U_j into itself; in other words, $T|_{U_j}$ may not be an operator on U_j. Thus we are led to consider only decompositions of V of the form above where T maps each U_j into itself.

The notion of a subspace that gets mapped into itself is sufficiently important to deserve a name.

Definition 5.4.1 (invariant subspace)**.** Suppose $T \in \mathcal{L}(V)$. A subspace U of V is called **invariant** under T if $u \in U$ implies $Tu \in U$. In other words, U is invariant under T if $T|_U$ is an operator on U.

Example 5.4.1. Suppose $T \in \mathcal{L}(V)$. Show that each of the following subspaces of V is invariant under T:

1. $\{0\}$;
2. V;
3. null T;
4. range T.

5.4.1 Eigenvalues and eigenvectors

Now we turn to an investigation of the simplest possible nontrivial invariant subspaces: invariant subspaces with dimension 1.

Take any $v \in V$ with $v \neq 0$ and let U equal the set of all scalar multiples of v:

$$U = \{\lambda v : \lambda \in \mathbf{F}\} = \operatorname{span}(v).$$

Then U is a 1-dimensional subspace of V, and every 1-dimensional subspace of V is of this form for an appropriate choice of v. If U is invariant under an operator $T \in \mathcal{L}(V)$, then $Tv \in U$, and hence there is a scalar $\lambda \in \mathbf{F}$ such that

$$Tv = \lambda v.$$

Conversely, if $Tv = \lambda v$ for some $\lambda \in \mathbf{F}$, then $\text{span}(v)$ is a 1-dimensional subspace of V invariant under T.

The equation

$$Tv = \lambda v,$$

which we have just seen is intimately connected with 1-dimensional invariant subspaces.

Definition 5.4.2 (eigenvalue)**.** Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \mathbf{F}$ is called an **eigenvalue** of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$.

The comments above show that T has a 1-dimensional invariant subspace if and only if T has an eigenvalue. In the definition above, we require that $v \neq 0$ because every scalar $\lambda \in \mathbf{F}$ satisfies $T0 = \lambda 0$.

Proposition 5.4.1 (Equivalent conditions to be an eigenvalue)**.** Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Then the following are equivalent:

1. λ is an eigenvalue of T;
2. $T - \lambda I$ is not injective;
3. $T - \lambda I$ is not surjective;
4. $T - \lambda I$ is not invertible.

Proof. Omitted. □

Definition 5.4.3 (eigenvector)**.** Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$ is an eigenvalue of T. A vector $v \in V$ is called an **eigenvector** of T corresponding to λ if $v \neq 0$ and $Tv = \lambda v$.

Because $Tv = \lambda v$ if and only if $(T - \lambda I)v = 0$, a vector $v \in V$ with $v \neq 0$ is an eigenvector of T corresponding to λ if and only if $v \in \text{null}\ (T - \lambda I)$.

Example 5.4.2. Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ is defined by

$$T(w, z) = (-z, w).$$

1. Find the eigenvalues and eigenvectors of T if $\mathbf{F} = \mathbf{R}$.
2. Find the eigenvalues and eigenvectors of T if $\mathbf{F} = \mathbf{C}$.

Proof.

1. No eigenvalues and thus no eigenvectors.
2. To find eigenvalues of T, we must find the scalars λ such that

   $$T(w, z) = \lambda(w, z)$$

   has **some** solution other than $w = z = 0$. The equation above is equivalent to the simultaneous equations

   $$-z = \lambda w, w = \lambda z.$$

Substituting the value for w given by the second equation into the first equation gives

$$-z = \lambda^2 z.$$

Now z cannot equal 0, otherwise $w = 0$. We are looking for solutions where (w, z) is not the 0 vector, so the equation above leads to the equation

$$-1 = \lambda^2.$$

The solutions to this equation are $\lambda = i, \lambda = -i$. The eigenvectors corresponding to the eigenvalue i are the vectors of the form $(w, -wi)$, with $w \in \mathbf{C}$ and $w \neq 0$, and the eigenvectors corresponding to the eigenvalue $-i$ are the vectors of the form (w, wi), with $w \in \mathbf{C}$ and $w \neq 0$.

□

Proposition 5.4.2 (Linearly independent eigenvectors)**.** Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and $v_1, \dots, v_m$ are corresponding eigenvectors. Then $v_1, \dots, v_m$ is linearly independent.

Proof. Suppose $v_1, \dots, v_m$ is linearly dependent. Let k be the smallest positive integer such that

$$v_k \in \operatorname{span}(v_1, \dots, v_{k-1});$$

and $v_1, \dots, v_{k-1}$ is linearly independent.

The existence of k with this property follows from the Linear Dependence Lemma. Thus there exists $a_1, \dots, a_{k-1} \in \mathbf{F}$ such that

$$v_k = \sum_{j=1}^{k-1} a_j v_j.$$

Hence

$$\lambda_k v_k = \sum_{j=1}^{k-1} a_j \lambda_k v_j.$$

Apply T to both sides of this equation, getting

$$\lambda_k v_k = \sum_{j=1}^{k-1} a_j \lambda_j v_j.$$

Subtract the equation above, getting

$$0 = \sum_{j=1}^{k-1} a_j (\lambda_k - \lambda_j) v_j$$

Since $v_1, \dots, v_{k-1}$ is linearly independent, the equation above implies that all the a's are 0 (recall that λ_k is not equal to any of $\lambda_1, \dots, \lambda_{k-1}$). However, this means that v_k equals 0, contradicting our hypothesis that v_k is an eigenvector. Therefore our assumption that $v_1, \dots, v_m$ is linearly dependent was false. □

Proposition 5.4.3 (Number of eigenvalues)**.** Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

5.4.2 Restriction and quotient operators

Definition 5.4.4 ($T|_U$ and T/U)**.** Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T.

1. The **restriction operator** $T|_U \in \mathcal{L}(U)$ is defined by
$$T|_U(u) = Tu$$
for $u \in U$.
2. The **quotient operator** $T/U \in \mathcal{L}(V/U)$ is defined by
$$(T/U)(v + U) = Tv + U$$
for $v \in V$.

For both the operators defined above, it is worthwhile to pay attention to their domains and to spend a moment thinking about why they are well defined as operators on their domains.

To show that the definition above of the quotient operator makes sense, we need to verify that if $v + U = w + U$, then $Tv + U = Tw + U$. Hence suppose $v + U = w + U$. Thus $v - w \in U$. Because U is invariant under T, we also have $T(v - w) \in U$, which implies that $Tv - Tw \in U$, which implies that $Tv + U = Tw + U$, as desired.

In some sense, we can learn about T by studying the operators $T|_U$ and T/U, each of which is an operator on a vector space with smaller dimension than V. However, sometimes $T|_U$ and T/U do not provide enough information about T.

Example 5.4.3. Defined an operator $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(x, y) = (y, 0)$. Let $U = \{(x, 0) : x \in \mathbf{F}\}$. Show that

1. U is invariant under T and $T|_U$ is the 0 operator on U;
2. there does not exist a subspace W of $\mathbf{F}^2$ that is invariant under T and such that $\mathbf{F}^2 = U \oplus W$;
3. T/U is the 0 operator on $\mathbf{F}^2/U$.

Proof.

1. For $(x, 0) \in U$, we have $T(x, 0) = (0, 0) \in U$. Thus U is invariant under T and $T|_U$ is the 0 operator on U.
2. Suppose W is subspace of $\mathbf{F}^2$ such that $\mathbf{F}^2 = U \oplus W$. Because $\dim \mathbf{F}^2 = 2$ and $\dim U = 1$, we have $\dim W = 1$. If W were invariant under T, then each nonzero vector in W would be an eigenvector of T. However, it is easy to see that 0 is the only eigenvalue of T and that all eigenvectors of T are in U. Thus W is not invariant under T.
3. For $(x, y) \in \mathbf{F}^2$, we have
$$(T/U)((x, y) + U) = T(x, y) + U = (y, 0) + U = 0 + U,$$
where the last equality holds because $(y, 0) \in U$. The equation above shows that T/U is the 0 operator.

$\square$

5.5 Exercises

1. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.
   (a) Prove that if $U \subset \operatorname{null} T$, then U is invariant under T.
   (b) Prove that if $\operatorname{range} T \subset U$, then U is invariant under T.
2. (a) Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Prove that null S is invariant under T.
   (b) Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Prove that range S is invariant under T.
3. Suppose that $T \in \mathcal{L}(V)$ and $U_1, \ldots, U_m$ are subspaces of V invariant under T. Prove that $U_1 + \cdots + U_m$ is invariant under T.
4. Suppose that $T \in \mathcal{L}(V)$. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.
5. Prove or give a counterexample: if V is finite-dimensional and U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$.
6. Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.
7. Define $T \in \mathcal{L}(\mathbf{F}^2)$ by
$$T(w, z) = (z, w).$$
   Find all eigenvalues and eigenvectors of T.
8. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by
$$T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$$
   Find all eigenvalues and eigenvectors of T.
9. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by
$$T(x_1, x_2, x_3, \ldots, x_n) = (x_1, 2x_2, 3x_3, \ldots, nx_n).$$
   (a) Find all eigenvalues and eigenvectors of T.
   (b) Find all invariant subspaces of T.
10. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Prove that there exists $\alpha \in \mathbf{F}$ such that $|\alpha - \lambda| < \frac{1}{1000}$ and $T - \alpha I$ is invertible.
11. Suppose $V = U \oplus W$, where U and W are nonzero subspaces of V. Define $P \in \mathcal{L}(V)$ by $P(u + w) = u$ for $u \in U$ and $w \in W$. Find all eigenvalues and eigenvectors of P.
12. Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible.
   (a) Prove that T and $S^{-1}TS$ have the same eigenvalues.
   (b) What is the relationship between the eigenvectors of T and the eigenvectors of $S^{-1}TS$.

13. Suppose V is a complex vector space, $T \in \mathcal{L}(V)$, and the matrix of T with respect to some basis of V contains only real entries. Show that if λ is an eigenvalue of T, then so is $\bar{\lambda}$.
14. Given an example of an operator $T \in \mathcal{L}(\mathbf{R}^4)$ such that T has no real eigenvalues.
15. Show that the operator $T \in \mathcal{L}(\mathbf{C}^\infty)$ defined by

$$T(z_1, z_2, \dots) = (0, z_1, z_2, \dots)$$

has no eigenvalues.

16. Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \dots, x_n) = (x_1 + \cdots + x_n, \dots, x_1 + \cdots + x_n).$$

Find all eigenvalues and eigenvectors of T.

17. Find all eigenvalues and eigenvectors of the backward shift operator $T \in \mathcal{L}(\mathbf{F}^\infty)$ defined by

$$T(z_1, z_2, z_3, \dots) = (z_2, z_3, \dots).$$

18. Suppose $T \in \mathcal{L}(V)$ is invertible.
    (a) Suppose $\lambda \in \mathbf{F}$ with $\lambda \neq 0$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.
    (b) Prove that T and T^{-1} have the same eigenvectors.
19. Suppose $T \in \mathcal{L}(V)$ and there exist nonzero vectors v and w in V such that

$$Tv = 3w, Tw = 3v.$$

Prove that 3 or -3 is an eigenvalue of T.

20. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.
21. Suppose $T \in \mathcal{L}(\mathbf{F}^n)$ is given by

$$T(x_1, \dots, x_n) = \left(\sum_{k=1}^{n} A_{1,k} x_k, \dots, \sum_{k=1}^{n} A_{n,k} x_k \right)$$

    (a) Suppose $\sum_{j=1}^{n} A_{j,k} = 1$ for all k. Prove that 1 is an eigenvalue of T.
    (b) Suppose $\sum_{k=1}^{n} A_{j,k} = 1$ for all j. Prove that 1 is an eigenvalue of T.
22. Suppose $T \in \mathcal{L}(V)$ and u, v are eigenvectors of T such that $u+v$ is also an eigenvector of T. Prove that u and v are eigenvectors of T corresponding to the same eigenvalue.
23. Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

24. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ is such that every subspace of V with dimension $\dim V - 1$ is invariant under T. Prove that T is a scalar multiple of the identity operator.
25. Suppose V is finite-dimensional with $\dim V \geq 3$ and $T \in \mathcal{L}(V)$ is such that every 2-dimensional subspace of V is invariant under T. Prove that T is a scalar multiple of the identity operator.
26. Suppose $T \in \mathcal{L}(V)$ and $\dim \operatorname{range} T = k$. Prove that T has at most $k + 1$ distinct eigenvalues.
27. Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ and $-4, 5$ and $\sqrt{7}$ are eigenvalues of T. Prove that there exists $x \in \mathbf{R}^3$ such that $Tx - 9x = (-4, 5, \sqrt{7})$.
28. Suppose V is finite-dimensional and $v_1, \ldots, v_m$ is a list of vectors in V. Prove that $v_1, \ldots, v_m$ is linearly independent if and only if there exists $T \in \mathcal{L}(V)$ such that $v_1, \ldots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.
29. Suppose $\lambda_1, \ldots, \lambda_n$ is a list of distinct real numbers. Prove that the list $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ is linearly independent in the vector space of real-valued functions on $\mathbf{R}$.
30. Suppose $T \in \mathcal{L}(V)$. Prove that $T/(\operatorname{range} T) = 0$.
31. Suppose $T \in \mathcal{L}(V)$. Prove that $T/(\operatorname{null} T)$ is injective if and only if $(\operatorname{null} T) \cap (\operatorname{range} T) = \{0\}$.
32. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is invariant under T. Prove that each eigenvalue of T/U is an eigenvalue of T.
33. Give an example of a vector space V, an operator $T \in \mathcal{L}(V)$, and a subspace U of V that is invariant under T such that T/U has an eigenvalue that is not an eigenvalue of T.

5.6 Upper-triangular operators

5.6.1 Polynomials applied to operators

Definition 5.6.1 (T^m)**.** Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

1. T^m is defined by
$$T^m = \underbrace{T \cdots T}_{m \text{ times}}.$$
2. T^0 is defined to be the identity operator I on V.
3. If T is invertible with inverse T^{-1}, then T^{-m} is defined by
$$T^{-m} = (T^{-1})^m.$$

You should verify that if T is an operator, then

$$T^m T^n = T^{m+n}, (T^m)^n = T^{mn},$$

where m and n are allowed to be arbitrary integers if T is invertible and non-negative integers if T is not invertible.

Definition 5.6.2 ($p(T)$)**.** Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$ si a polynomial given by

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for $z \in \mathbf{F}$. Then $p(T)$ is the operator defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$$

If we fix an operator $T \in \mathcal{L}(V)$, then the function from $\mathcal{P}(\mathbf{F})$ to $\mathcal{L}(V)$ given by $p \mapsto p(T)$ is linear, as you should verify.

Definition 5.6.3 (product of polynomials)**.** If $p, q \in \mathcal{P}(\mathbf{F})$, then $pq \in \mathcal{P}(\mathbf{F})$ is the polynomial defined by

$$(pq)(z) = p(z)q(z)$$

for $z \in \mathbf{F}$.

Proposition 5.6.1 (Multiplicative properties)**.** Suppose $p, q \in \mathcal{P}(\mathbf{F})$ and $T \in \mathcal{L}(V)$. Then

$$(pq)(T) = p(T)q(T) = q(T)p(T)$$

Proof. Omitted. □

5.6.2 Existence of eigenvalues

Proposition 5.6.2 (Operators on complex vector spaces have an eigenvalue)**.** Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

Proof. Suppose V is a complex vector space with dimension $n > 0$ and $T \in \mathcal{L}(V)$. Choose $v \in V$ with $v \neq 0$. Then

$$v, Tv, T^2 v, \ldots, T^n v$$

is not linearly independent, because V has dimension n and we have $n + 1$ vectors. Thus there exist complex numbers $a_0, \ldots, a_n$, not all 0, such that

$$0 = a_0 v + a_1 T v + \cdots + a_n T^n v.$$

Note that $a_1, \ldots, a_n$ cannot all be 0, because otherwise the equation above would become $0 = a_0 v$, which would force a_0 also to be 0.

Make the a's the coefficients of a polynomial, which by the Fundamental Theorem of Algebra has a factorization

$$a_0 + a_1 z + \cdots + a_n z^n = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where c is a nonzero complex number, each λ_j is in $\mathbf{C}$, and the equation holds for all $z \in \mathbf{C}$.

We then have

$$0 = a_0 v + a_1 T v + \cdots + a_n T^n v = c(T - \lambda_1 I) \cdots (T - \lambda_m I)v$$

Thus $T - \lambda_j I$ is not injective for at least one j. In other words, T has an eigenvalue. $\square$

5.6.3 Upper-triangular operators

Proposition 5.6.3 (Conditions for upper-triangular operator)**.** Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then the following are equivalent:

1. $Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$ for each $j = 1, \ldots, n$;
2. $\operatorname{span}(v_1, \ldots, v_j)$ is invariant under T for each $j = 1, \ldots, n$.

Proof. We need only prove that (1) implies (2). $\square$

Theorem 5.6.1 (Schur's theorem)**.** Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T can be upper-triangular with respect to some basis of V.

Proof. We will use induction on the dimension of V. Clearly the desired result holds if $\dim V = 1$.

Suppose now that $\dim V > 1$ and the desired result holds for all complex vector spaces whose dimension is less than the dimension of V. Let λ be any eigenvalue of T. Let

$$U = \text{range } (T - \lambda I).$$

Because $T - \lambda I$ is not surjective, $\dim U < \dim V$. Furthermore, U is invariant under T. To prove this, suppose $u \in U$. Then

$$Tu = (T - \lambda I)u + \lambda u.$$

Obviously $(T - \lambda I)u \in U$ (because U equals the range of $T - \lambda I$) and $\lambda u \in U$. Thus the equation above shows that $Tu \in U$. Hence U is invariant under T, as claimed.

Thus $T|_U$ is an operator on U. By our induction hypothesis, there is a basis $u_1, \ldots, u_m$ of U with respect to which $T|_U$ has an upper-triangular matrix. Thus for each j we have

$$Tu_j = (T|_U)(u_j) \in \operatorname{span}(u_1, \ldots, u_j).$$

Extend $u_1, \ldots, u_m$ to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V. For each k, we have

$$Tv_k = (T - \lambda I)v_k + \lambda v_k.$$

The definition of U show that $(T - \lambda I)v_k \in U = \operatorname{span}(u_1, \ldots, u_m)$. Thus the equation above shows that

$$Tv_k \in \operatorname{span}(u_1, \ldots, u_m, v_1, \ldots, v_n).$$

We conclude that T has an upper-triangular matrix with respect to the basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V, as desired. □

Proof. We will use induction on the dimension of V. Clearly the desired result holds if $\dim V = 1$.

Suppose now that $\dim V = n > 1$ and the desired result holds for all complex vector spaces whose dimension is $n - 1$. Let v_1 be any eigenvector of T. Let $U = \operatorname{span}(v_1)$. Then U is an invariant subspace of T and $\dim U = 1$.

Because $\dim V/U = n - 1$, we can apply our induction hypothesis to $T/U \in \mathcal{L}(V/U)$. Thus there is a basis $v_2 + U, \ldots, v_n + U$ of V/U such that T/U has an upper-triangular matrix with respect to this basis. Hence

$$(T/U)(v_j + U) \in \operatorname{span}(v_2 + U, \ldots, v_j + U)$$

for each $j = 2, \ldots, n$. Unraveling the meaning of the inclusion above, we see that

$$Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$$

for each $j = 1, \ldots, n$. Thus T has an upper-triangular basis $v_1, \ldots, v_n$. □

Proposition 5.6.4 (Determination of invertibility from upper-triangular basis)**.** Suppose $T \in \mathcal{L}(V)$ has an upper-triangular basis $v_1, \ldots, v_n$ of V. Then T is invertible if and only if $\prod_{j=1}^{n} v_j'(Tv_j) \neq 0$.

Proof. Omitted. □

Proposition 5.6.5 (Determination of eigenvalues from upper-triangular basis)**.** Suppose $T \in \mathcal{L}(V)$ has an upper-triangular basis $v_1, \ldots, v_n$ of V. Then the eigenvalues of T are precisely $v_j'(Tv_j)$.

Proof. Omitted. □

5.7 Exercises

1. Suppose $T \in \mathcal{L}(V)$ and there exists a positive integer n such that $T^n = 0$.

   (a) Prove that $I - T$ is invertible and that

   $$(I - T)^{-1} = I + T + \cdots + T^{n-1}.$$

   (b) Explain how you would guess the formula above.

2. Suppose $T \in \mathcal{L}(V)$ and $(T - 2I)(T - 3I)(T - 4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

3. Suppose $T \in \mathcal{L}(V)$ and $T^2 = 1$ and -1 is not an eigenvalue of T. Prove that $T = I$.
4. Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \text{null } P \oplus \text{range } P$.
5. Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial. Prove that
$$p(STS^{-1}) = Sp(T)S^{-1}.$$
6. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. Prove that U is invariant under $p(T)$ for every polynomial $p \in \mathcal{P}(\mathbf{F})$.
7. Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or -3 is an eigenvalue of T.
8. Give an example of $T \in \mathcal{L}(\mathbf{R}^2)$ such that $T^4 = -I$.
9. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$ with $v \neq 0$. Let p be a nonzero polynomial of smallest degree such that $p(T)v = 0$. Prove that every zero of p is an eigenvalue of T.
10. Suppose $T \in \mathcal{L}(V)$ and v is an eigenvector of T with eigenvalue λ. Suppose $p \in \mathcal{P}(\mathbf{F})$. Prove that $p(T)v = p(\lambda)v$.
11. Suppose $\mathbf{F} = \mathbf{C}, T \in \mathcal{L}(V), p \in \mathcal{P}(\mathbf{C})$ is a nonconstant polynomial, and $\alpha \in \mathbf{C}$. Prove that α is an eigenvalue of $p(T)$ if and only if $\alpha = p(\lambda)$ for some eigenvalue λ of T.
12. Show that the result in the previous exercise does not hold if $\mathbf{C}$ is replaced with $\mathbf{R}$.
13. Suppose W is a complex vector space and $T \in \mathcal{L}(W)$ has no eigenvalues. Prove that every subspace of W invariant under T is either $\{0\}$ or infinite-dimensional.
14. Give an example of an operator with respect to some basis satisfies $v_j'(Tv_j) = 0$ for all j, but the operator is invertible.
15. Rewrite the proof of Proposition 5.6.2 using the linear map that sends $p \in \mathcal{P}_n(\mathbf{C})$ to $(p(T))v \in V$.
16. Rewrite the proof of Proposition 5.6.2 using the linear map that sends $p \in \mathcal{P}_{n^2}(\mathbf{C})$ to $p(T) \in \mathcal{L}(V)$.
17. Suppose V is a finite-dimensional with $\dim V > 1$ and $T \in \mathcal{L}(V)$. Prove that
$$\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V).$$
18. Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T has an invariant subspace of dimension k for each $k = 1, \ldots, \dim V$.

Epilogue

Our next course, **Advanced Linear Algebra** will cover the following topics:

- Diagonal operators
- Generalized eigenvectors
- Nilpotent operators
- Multiplicity of an eigenvalue
- Block diagonal operators
- The Cayley–Hamilton theorem
- The minimal polynomial of an operator
- Jordan canonical form
- Complexification of a vector space
- Complexification of an operator
- Trace
- Determinant
- ......
- ......