Honors Linear Algebra

MATH0540

Professor Melody Chan

Brown University

Edited by

Richard Tang

Contents

Chapter 1

Fundamentals of Linear Algebra

§1.1 Sets

Sets serve as a fundamental construct in higher-level mathematics. We start with a brief introduction to set theory.

Definition 1.1.1: Sets

A **set** is a collection of elements.

1. $x \in X$ means x is an element of X.
2. $x \notin$ means x is not an element of X.
3. $X \subset Y$ means X is a subset of Y (i.e. $\forall x \in X, x \in Y$.)
4. $X = Y \iff X \subset Y \wedge Y \subset X$.
5. $A \cap B := \{x \mid x \in A \wedge x \in B\}$ means set intersection.
6. $A \cup B := \{x \mid x \in A \vee x \in B\}$ means set union.
7. $A \setminus B := \{x \mid x \in A \wedge x \notin B\}$ means set difference.

Example 1. *Let*

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, \dots\}.$$

denote the set of integers, and let

$$\mathbb{Z}^+ = \{0, 1, \dots\}.$$

denote the set of positive integers.

§1.1.1 Set Builder notation

Sets may be defined formally with set-builder notation:

$$X = \{\ expression \mid rule\}.$$

Example 2.

1. *Let E represent the set of all even numbers. This set is expressed*

$$E = \{n \in \mathbb{Q} \mid \exists k \in \mathbb{Z} \ s.t.\ n = 2k\}.$$

2. *Let A represent the set of real numbers whose squares are rational numbers:*

$$A = \{a \in \mathbb{R} \mid a^2 \in \mathbb{Q}\}.$$

§1.1.2 Cartesian Products

Definition 1.1.2: Ordered Tuples

An **ordered pair** is defined (x, y). An n**-ordered tuple** is an ordered list of n items

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

Definition 1.1.3: Cartesian Products

Let A, B be sets. The **cartesian product** $A \times B$ is defined

$$A \times B := \{(a, b) \mid a \in A, b \in B\}.$$

Similarly, define the n-fold cartesian product

$$A^n := A \times A \times \cdots \times A.$$

Example 3. $\mathbb{R}^2$ *and* $\mathbb{R}^3$ *are examples of commonly known Cartesian products, which represent the 2D- and 3D-plane respectively.*

Example 4. R^n *is a first example of a* ***vector space.*** *Let* $n \in \mathbb{Z}^+ \cup \{0\}$*:*

1. *(Addition in R^n) We define an* ***addition operation*** *on $\mathbb{R}^n$ by adding coordinate-wise*

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

2. *(Scaling) Given $(x_1, \ldots, x_n) \in \mathbb{R}^n, \lambda \in \mathbb{R}$, we define*

$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Remark 1. $\mathbb{R}_0 = \{0\}$.

§1.1.3 Functions

Let A, B be sets. Informally, a function $f : A \to B$ deterministically returns an element $b \in B$ for each $a \in A$. We write $f(a) = b$.

Example 5. *The function* $f : \mathbb{R} \to \mathbb{R}$ *given by* $f(x) = x^2$ *maps* $\mathbb{R}$ *to the subset*

$$S \subset \mathbb{R} = \{(x, x^2) \mid x \in \mathbb{R}\}.$$

Definition 1.1.4: Functions

Let A, B be sets. A function $f : A \to B$ is a subset $G_f \subset A \times B$ such that for all $a \in A$, there exists at most one $b \in B$ s.t. $(a, b) \in G_f$. We write $f(a) = b$ when $(a, b) \in G_f$.

Definition 1.1.5: Codomain

Given a function $f : A \to B$, A is the **domain** of f, and B is the **codomain** or **target** of f. Let the **range** of f be defined as

$$\{b \in B \mid f(a) = b, a \in A\}.$$

The range is the subset of B. Importantly, the number of elements in the range of f cannot be larger than the number of elements in A, as each $f(a)$ maps to at most one $b \in B$.

Definition 1.1.6: Bijectivity

Let $f : A \to B$ be a function.

1. f is **injective**, or an **injection**, if $a_1, a_2 \in A$ and $f(a_1) = f(a_2)$ implies $a_1 = a_2$.
2. f is **surjective**, or a **surjection**, if for any $b \in B$, there exists an $a \in A$ such that $f(a) = b$. Equivalently, the range is the whole codomain.
3. f is **bijective**, or a **bijection**, if it is both injective and surjective. Equivalently, for every $b \in B$, there is a unique $a \in A$ such that $f(a) = b$.

§1.2 Fields

Roughly speaking, a **field** is a set, together with operations addition and multiplication. Vector spaces may be defined *over* fields.

Definition 1.2.1: Fields

A **field** is a set $\mathbb{F}$ containing elements named 0 and 1, together with binary operations $+$ and $\cdot$ satisfying, for all $a, b, c \in \mathbb{F}$:

- **commutativity**: $a+b=b+a, a\cdot b=b\cdot a$
- **associativity**: $a+(b+c)=(a+b)+c, a\cdot(b\cdot c)=(a\cdot b)\cdot c$
- **identities**: $0+a=a, 1\cdot a=a$
- **additive inverse**: For any $a\in\mathbb{F}$, there exists a $b\in\mathbb{F}$ such that $a+b=0$. We denote this $b=-a$
- **multiplicative inverse**: For any $a\in\mathbb{F}$, $a\neq 0$, there exists a $b\in\mathbb{F}$ such that $ab=1$.
- **distributivity**: $a\cdot(b+c)=a\cdot b+a\cdot c$.

Example 6. $\mathbb{R}^+\setminus\{0\}$ *is **not** a field under* $+,\cdot$.

Example 7. *(Finite Fields) Let* p *prime (e.g.* $p=5$*). Define*

$$\mathbb{F}_p=\{0,\ldots,p-1\},$$

with binary operations $+_p,\cdot_p$ *given by addition and multiplication modulo* p*. We claim (without proof) that* $\mathbb{F}_p$ *is a field.*

Example 8. *Let* $\mathbb{C}=\{a+bi\mid a,b\in\mathbb{R}\}$*. Elements of* $\mathbb{C}$ *are called **complex numbers**. Formally, a complex number is an ordered pair* $(a,b), a,b\in\mathbb{R}$*. We define addition as*

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

and multiplication as

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

Showing $\mathbb{C}$ *is a field is left as an exercise for the reader.*

Proposition 1.2.1: $\mathbb{C}$ Multiplicative Inverse

For every $\alpha\in\mathbb{C}\setminus\{0\}$, there exists $\beta\in\mathbb{C}$ with $\alpha\cdot\beta=1$.

Proof. Given $\alpha\in\mathbb{C}\setminus\{0\}$, let us write $\alpha=a+bi$. Then not both $a,b=0$. Let $\beta=\frac{a}{a^2+b^2}+-\frac{b}{a^2+b^2}i$. Then $\alpha\beta=(a+bi)\left(\frac{a}{a^2+b^2}+-\frac{b}{a^2+b^2}\right)=1$.
Thus $\forall\alpha\in\mathbb{C}\setminus\{0\},\exists\beta\in\mathbb{C}$ s.t. $\alpha\cdot\beta=1$. $\square$

$\mathbb{R}^n$ and $\mathbb{C}^n$ are specific examples of fields, but by no means the only ones (for instance, $\mathbb{F}^2$ with addition and multiplication modulo 2 is a field). Fields serve as the underlying set of numbers and operations that vector spaces are built on. In this course, we focus primarily on $\mathbb{R}$ and $\mathbb{C}$; but many of the definitions, theorems, and proofs work interchangeably with abstract fields.

§1.3 Vector Spaces

Vector spaces serve as the fundamental abstract structure of linear algebra. All future topics will build on vector spaces. Roughly, a vector space V is a set of **vectors** with an addition operation and scalar multiplication, where scalars are drawn from a field $\mathbb{F}$. We now formalize this definition.

Definition 1.3.1: Vector Spaces

Given a field $\mathbb{F}$, A **vector space** over $\mathbb{F}$, denoted $V_{\mathbb{F}}$, is a set V, together with vector addition on V

$$+ : V \times V \longrightarrow V$$

and scalar multiplication on V

$$\cdot : \mathbb{F} \times V \longrightarrow V$$

satisfying the following properties:

- (additive associativity) For all $u, v, w \in V$, $u + (v + w) = (u + v) + w$.
- (additive identity) There exists an element $0 \in V$ such that $v + 0 = 0 + v = 0$.
- (additive inverse) For all $v \in V$, there exists $w \in V$ such that $v + w = w + v = 0$. We denote $w = -v$.
- (additive commutativity) For all $v, w \in V$, $v + w = w + v$.
- (scalar multiplicative associativity) For all $\alpha, \beta \in \mathbb{F}, v \in V$, $\alpha(\beta v) = (\alpha\beta)v$.
- (scalar multiplicative identity) There exists an element $1 \in \mathbb{F}$ such that $1v = v$ for all $v \in V$.
- (Distributive Law I) For every $\alpha \in \mathbb{F}$, $v, w \in V$, $a \cdot (v + w) = a \cdot v + a \cdot w$.
- (Distributive Law II) For every $\alpha, \beta \in \mathbb{F}, v \in V$, $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.

We call elements of $\mathbb{F}$ **scalars**, and elements of V **vectors**, or **points**.

Example 9. *We say V is a vector space over $\mathbb{F}$. A vector space over $\mathbb{R}$ is called a* ***real vector space****, and a vector space over $\mathbb{C}$ is called a* ***complex vector space****.*

Example 10. *Let $\mathbb{F}$ be a field.*

1. *For some integers $n \geq 0$, $\mathbb{F}^n = \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{F}\}$ with vector addition defined*

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

and scalar multiplication defined

$$\lambda \cdot (v_1, v_2, \ldots, v_n) = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n).$$

Note that $F^0 = \{0\}$.

2. $\mathbb{F}^\infty = P\{(a_1, a_2, a_3, \dots) \mid a_j \in \mathbb{F}, j \in \mathbb{N}\}$ *with vector addition and scalar multiplication defined similarly.*

3. *Let* S *be any set; consider* $\{g : S \to \mathbb{F}\}$ *be the set of functions from* S *to* $\mathbb{F}$. *Given* $f, g : S \to \mathbb{F}$, $\lambda \in \mathbb{F}$, *define vector addition* $(f + g) : S \to \mathbb{F}$ *as*

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

and scalar multiplication $\lambda f : S \to \mathbb{F}$ *as*

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

Perhaps counterintuitively, example 3 subsumes example 1! For example, let $S = \{1, 2, \dots, n\}$, and let $\mathbb{R}^{\{1,\dots,n\}}$ be the set of all functions from $\{1, \dots, n\} \to \mathbb{R}$. One such f may be

$$\begin{aligned} f : \{1, \dots, n\} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = x^2 - 3. \end{aligned}$$

But f can also be thought of as an n-tuple. For instance, with $n = 3$, we can define a function

$$f = (-2, 1, 6) \in \mathbb{R}^3.$$

This is equivalent to $f(1) = -2, f(2) = 1, f(3) = 6$. Similarly, if $f(x) = e^x$, then $f \in \mathbb{R}^{\{1,2,3\}} = \left(e, e^2, e^3\right) \in \mathbb{R}^3$, since $f(1) = e, f(2) = e^2, f(3) = e^3$.

In other words, every n-tuple $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ could be represented as a function $f : \{1, 2, \dots, n\} \to \mathbb{R}$, where $f(1) = x_1, f(2) = x_2, \dots, f(n) = x_n$. The key insight here is that **the function f is the n-tuple**; the one function $f(x) = e^x$ is equivalent to the n-tuple $\left(e, e^2, \dots, e^n\right)$.

From this, we get that the set of functions $\mathbb{R}^{\{1,\dots,n\}} = \mathbb{R}^n$, the set of n-tuples.

Remark 2. *Reinterpret* $\mathbb{F}^0 = \{$*functions* $f : \varnothing \longrightarrow \mathbb{F}\}$. *How many functions are there from* $\varnothing \longrightarrow \mathbb{F}$*?*
One function $\varnothing = \varnothing \times \mathbb{F}$.

Example 11. *The set of continuous functions* $f : \mathbb{R} \to \mathbb{R}$ *forms a vector space over* $\mathbb{R}$. *In particular, the sum of two continuous functions is continuous; and* $a \cdot f$ *is continuous for any* $a \in \mathbb{R}$, *and* f *continuous.*

But what about fields over fields? Are these vector spaces?

Example 12. *Let $\mathbb{K}$ be a field, and say $\mathbb{F} \subseteq \mathbb{K}$ ($\mathbb{F}$ is a subfield of $\mathbb{K}$). Then $\mathbb{K}$ is a vector space over $\mathbb{F}$, with addition defined as in $\mathbb{K}$, and with scalar multiplication defined*

$$\lambda \cdot x = \lambda x, \text{ where } \lambda \in \mathbb{F}, x \in \mathbb{K}.$$

Thus $\mathbb{C}$ is a real vector space (this is why we draw the complex plane like $\mathbb{R}^2$!).

§1.3.1 Properties of Vector Spaces

We now observe some fascinating properties of vector spaces. Let V be a vector space over a field $\mathbb{F}$.

Proposition 1.3.1: Unique Additive Identity

V has a unique additive identity.

Proof. Suppose $e, e' \in V$ are both additive identities. Then

$$\begin{aligned} e &= e + e' \\ &= e'. \end{aligned}$$

Thus $e = e'$. □

Proposition 1.3.2: Unique Additive Inverse

Every vector $v \in V$ has a unique additive inverse.

Proof. Let $v \in V$, and suppose $w, w' \in V$ are both additive inverses of v. Then

$$\begin{aligned} 0 &= v + w \\ w' &= (w + v) + w' \\ w' &= w + (v + w') \\ w' &= w + 0 \\ w' &= w. \end{aligned}$$

Thus $w = w'$. □

Let us also define a notion of subtraction: we say $v - w = v + (-w)$.

Proposition 1.3.3: -v

For any $v \in V$,

$$-v = (-1) \cdot v.$$

Proof. Let $v, -v \in V$ where $-v$ is the inverse of v. Then

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

Since every $v \in V$ has a unique additive inverse, $-v = (-1) \cdot v$. □

Proposition 1.3.4: 0 Times a Vector

For every $v \in V$, $0v = 0$.

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

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

Adding the additive inverse of $0v$ to both sides, we get $0v = 0$. $\square$

Proposition 1.3.5: Scalar Times 0

For every $a \in \mathbb{F}$, $a\mathbf{0} = \mathbf{0}$.

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

$$a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0} + a\mathbf{0}.$$

Adding the additive inverse to both sides yields $a\mathbf{0} = \mathbf{0}$. $\square$

§1.4 Subspaces

Subspaces can greatly expand our examples of vector spaces.

Definition 1.4.1: Subspaces

A subset $U \subseteq V$ is a **subspace** (or a **linear subspace**) of V if U is also a vector space.
U is a subspace of V if and only if

1. $\mathbf{0} \in U$.
2. For all $u, w \in U$, $u + w \in U$.
3. For all $u \in U$, $\lambda \in \mathbb{F}$, $\lambda \cdot u \in U$.

That is, addition and scalar multiplication are **closed** in U, and the identity element exists.
We see that these three properties are enough for U to satisfy the six properties of vector spaces: associativity, commutativity, and distributivity are automatically satisfied, as they hold on the larger space V (and so also hold on the subspace U); addition and scalar multiplication make sense in U, and the additive identity exists; the third condition guarantees the additive inverse ($-v = -1v$).

Example 13. *What are the subspaces of* $\mathbb{R}^2$ *and* $\mathbb{R}^3$*?*

Solution: It turns out that there are only three valid types of subspaces of $\mathbb{R}^2$:

1. The zero vector $\mathbf{0} = (0, 0)$.

2. All lines through the origin ($y = \alpha x$).
3. $\mathbb{R}^2$ itself.

Similarly, there are only four valid types of subspaces of $\mathbb{R}^2$:

1. The zero vector $\mathbf{0} = (0, 0, 0)$.
2. All lines through through the origin.
3. All planes through the origin.
4. $\mathbb{R}^3$ itself.

Let us now do a rough sketch of a proof that the list of subspaces of $\mathbb{R}^2$ is complete.

Proof. Let W be a subspace of R^2. If W has no nonzero vectors, then $W = \{\mathbf{0}\}$. If W has a non-zero vector $v \in V \setminus \{\mathbf{0}\}$, then W must contain the line through v passing through $\mathbf{0}$.
Moreover, if W contains some $w \in V$ not on the line, we have the ability to "turn" the coordinate plane, such that any $u \in V$ can be formed by $\alpha v + \beta w$. □

§1.4.1 Sums of Subspaces

With vector spaces, we are primarily only interested in subspaces, not arbitrary subsets. Thus, the notion of the sum of subspaces is useful.

Definition 1.4.2: Sum of Subsets

Suppose $U_1, \ldots, U_m$ are subsets of V. The **sum** of $U_1, \ldots, U_m$, denoted $U_1 + \ldots + U_m$, is the set of all possible sums of elements of $U_1, \ldots, U_m$. Precisely,

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

Example 14. *Suppose $V = \mathbb{R}^3$. Let $U_1 = \{(x, 0, 0) \in \mathbb{R}^3 \mid x \in \mathbb{R}\}$ be the subspace containing elements with only x components, and $U_2 = \{(0, y, 0) \in \mathbb{R}^3 \mid y \in \mathbb{R}\}$ be the subspace containing elements with only y components. Then*

$$U_1 + U_2 = \{(x, y, 0)\mathbb{R}^3 \mid x, y \in \mathbb{R}\},$$

or the xy-plane.

Are these sums of subspaces actually subspaces themselves? Indeed, it is the smallest subspace containing all of the individual subspaces.

Proposition 1.4.1: Sum of Subspaces

Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \ldots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.

Proof. Clearly, $0 \in U_1 + \ldots + U_m$ and addition and scalar multiplication in $U_1 + \ldots + U_m$ is closed. Thus $U_1 + \ldots + U_m$ is a subspace of V.

To show that it is the smallest, observe first that $U_1, \ldots, U_m$ are all contained in $U_1 + \ldots + U_m$ (for U_j, simply set $u_i = 0$ for any $i \neq j$). Additionally, every subspace of V containing $U_1, \ldots, U_m$ contains $U_1 + \ldots + U_m$ as well, since subspaces must contain all finite sums of their elements (in this case, $u_i \in U_i$). Thus, since $U_1 + \ldots + U_m$ contains every individual subspace, and any subspace containing $U_1, \ldots, U_m$ also contains $U_1 + \ldots + U_m$, we have that $U_1 + \ldots + U_m$ is the smallest subspace containing $U_1, \ldots, U_m$. □

§1.4.2 Direct Sums

Suppose $U_1, \ldots, U_m$ are subspaces of V. Every element of $U_1 + \ldots + U_m$ can be written as

$$u_1 + \ldots + u_m,$$

where each u_j is in U_j. Like the concept of injectivity, we are interested in the case when each vector in $U_1 + \ldots + U_m$ can only be written in one way. We call these **direct sums**.

Definition 1.4.3: Direct Sum

Suppose $U_1, \ldots, U_m$ are subspaces of V. The sum $U_1 + \ldots + U_m$ is a **direct sum** if each element of $U_1 + \ldots + U_m$ can be written in only one way as a sum $u_1 + \ldots + u_m$, where $u_j \in U_j$. We denote this sum

$$U_1 \oplus \ldots \oplus U_m.$$

Two theorems are useful in determining if a sum of subspaces is a direct sum. Their proofs are left as an exercise for the reader.

Theorem 1.4.1: Condition for a Direct Sum

Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \ldots + U_m$ is a direct sum if and only if the only way to write

$$0 = u_1 + \ldots + u_m$$

is by setting each $u_j = 0$.

Proof. One direction is easy. To show the other direction, assume there are multiple ways to write a vector v, and perform arithmetic $0 = v - v$ to arrive at $u_j = 0$. □

Theorem 1.4.2: Direct Sum of Two Subspaces

Suppose U, W are subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.

Proof. If we know direct sum, then there is only one way to write $0 = v + -v$ ($v \in U \cap W$). For the other direction, try writing $0 = u + w$ for some $u \in U$, $w \in W$, and showing that $u = w = 0$ necessarily. □

Chapter 2

Finite-Dimensional Vector Spaces

§2.1 Span and Linear Independence

Suppose a friend imagines a subspace $W \subseteq \mathbb{R}^3$. You know that $(1,0,0),(0,1,0) \in W$. What else do you know must be in W? Well, first, $\mathbf{0} = (0,0,0) \in W$ by definition. But moreover, anything in the form $\{(a,b,0) \mid a,b \in \mathbb{R}\}$ (the xy-plane) must be in W, since any point on the plane can be made by $\alpha \cdot a + \beta \cdot b$ (we will later see that $(1,0)$ and $(0,1)$ are **basis vectors** of $\mathbb{R}^2$).

Definition 2.1.1: Linear Combination and Span

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

$$\lambda_1 v_1 + \ldots + \lambda_n v_n, \text{ where } \lambda_i \in \mathbb{F}.$$

The **span** (or **linear span**) of $v_1, \ldots, v_n$, is the set of all linear combinations of $v_1, \ldots, v_n$:

$$\operatorname{span}(v_1, \ldots, v_n) = \{a_1 v_1 + \ldots + a_m v_m \mid a_i \in \mathbb{F}\}.$$

The span of no vectors is $\{0\}$.

Proposition 2.1.1: Span is Smallest Subspace

The span of $v_1, \ldots, v_m$ is the smallest subspace of V containing $v_1, \ldots, v_m$. Precisely:

1. $\operatorname{span}(v_1, \ldots, v_m)$ is a subspace of V.
2. Any subspace W of V containing $v_1, \ldots, v_m$ also contains $\operatorname{span}(v_1, \ldots, v_m)$.

Proof. Let $v_1, \ldots, v_m$ be a list of vectors in V.

$\operatorname{span}(v_1, \ldots, v_m)$ is clearly a subspace of V: achieve $\mathbf{0} \in \operatorname{span}(v_1, \ldots, v_m)$ by setting each $a_j = 0$, and since $a_j + b_j, \lambda a_j \in \mathbb{F}$, $\operatorname{span}(v_1, \ldots, v_m)$ is closed under addition and scalar multiplication.

Now, we show that $\operatorname{span}(v_1, \ldots, v_m)$ is the smallest subspace containing $v_1, \ldots, v_m$. Every vector v_j is a linear combination of $v_1, \ldots, v_m$ (take, for $i \neq j$, $a_i = 0$); thus $\operatorname{span}(v_1, \ldots, v_m)$ contains each v_j. Additionally, every subspace U of V that contains each $v_j \in U$ is closed under addition and scalar multiplication, so U contains every linear combination of $v_1, \ldots, v_m$; thus U contains $\operatorname{span}(v_1, \ldots, v_m)$. So, since $\operatorname{span}(v_1, \ldots, v_m)$

contains every vector v_j, and any subspace U of V that contains every vector v_j also contains $\operatorname{span}(v_1, \ldots, v_m)$, the span is the smallest subspace containing every v_j. □

Definition 2.1.2: Spanning a Vector Space

If $\operatorname{span}(v_1, \ldots, v_m) = V$, then $v_1, \ldots, v_m$ **spans** V, and $v_1, \ldots, v_m$ are a **spanning set**.

We now make one of the key definitions of linear algebra.

Definition 2.1.3: Finite Dimensional Vector Spaces

If V is spanned by a **finite** list of vectors $v_1, \ldots, v_m$ then V is **finite-dimensional**.

If V is not finite-dimensional, then V is **infinite-dimensional**.

Example 15. *Let $\mathcal{P}(\mathbb{F})$ be the set (indeed, vector space) of polynomials over a field $\mathbb{F}$. Show $\mathcal{P}(\mathbb{F})$ is infinite-dimensional.*

Solution: Let $p \in \mathcal{P}(\mathbb{F})$, and let m denote the highest degree polynomial in $\mathcal{P}(\mathbb{F})$. Then p has at most degree m; thus a polynomial p^{m+1} is not spanned by any list of vectors in $\mathcal{P}(\mathbb{F})$; thus $\mathcal{P}(\mathbb{F})$ is finite-dimensional.

§2.1.1 Linear Independence

As with sums/direct sums, we are interested if a vector has a unique linear combination; that is, given a list $v_1, \ldots, v_m \in V$, and $v \in \operatorname{span}(v_1, \ldots, v_m)$, are there unique $a1, \ldots, a_m \in \mathbb{F}$ such that

$$v = a_1v_1 + \ldots + a_mv_m?$$

In other words, is there only one way to create a certain vector given a span? Suppose there's more than one way; then there exists $b_1, \ldots, b_m \in \mathbb{F}$ such that

$$v = b_1v_1 + \ldots + b_mv_m;$$

then

$$0 = (a_1 - b_1)v_1 + \ldots + (a_mb_m)v_m.$$

If the only way to do this is the obvious way, where $a_i - b_i = 0$, then the representation is unique. We call this **linear independence**.

Definition 2.1.4: Linear Independence

A list of vectors $v_1, \ldots, v_m \in V$ is **linearly independent** if the only choice of $a_1, \ldots, a_m \in \mathbb{F}$ that makes $a_1v_1 + \ldots + a_mv_m$ equal 0 is $a_i = 0$.

A list of vectors in V is **linearly dependent** if it is not linearly independent.

That is, there exist non-zero $a_i \in \mathbb{F}$ such that

$$0 = \sum_{i=1}^{m} a_i v_i.$$

An empty list of vectors () is linearly independent.

Example 16. *1. A list of one vector $v \in V$ is linearly independent if and only if v is non-zero.*

2. *A list of two vectors $v1, v_2 \in V$ is linearly independent if and only if one vector is not a scalar combination of the other vector; that is, $v_1 \neq \lambda v_2$ for some $\lambda \in \mathbb{F}$.*

3. $(1,0,0),\ (0,1,0) \in \mathbb{R}^3$ *is linearly independent.*

4. $(1,-1,0), (-1,0,1), (0,1,-1) \in \mathbb{R}^3$ *is linearly dependent. In particular,* $(1,-1,0) + (-1,0,1) + (0,1,-1) = \mathbf{0}$. *Alternatively, we can write* $(-1,0,1)$ *as a linear combination of the other two:*

$$(-1,0,1) = -1 \cdot (1,-1,0) - (0,1,-1).$$

Intuitively, a list of vectors is linearly independent if none of its vectors are a linear combination of the other vectors; each vector is "independent" of the other vectors. In other words, a vector is linearly independent if it is not in the span of the other vectors. Its negation is also important, and is arguably more intuitive: a list of vectors is linearly dependent **iff** it is in the span of the other vectors (it is "dependent" on the other vectors). Formally, this gives rise to an important lemma, and theorem.

Lemma 2.1.1: Linear Dependence Lemma

Suppose that $v_1, \ldots, v_m \in V$ is a linearly dependent list of vectors. Then there exists some $j \in \{1, \ldots, m\}$ such that:

1. $v_j \in \text{span}(v_1, ..., v_{j-1})$
2. If the j^{th} term is removed from the list, the span of the remaining vectors $v_1, \ldots, \hat{v_j}$[a]$, \ldots, v_m$ equals $\text{span}(v_1, \ldots, v_m)$.

In other words, removing the linearly dependent vector has no effect on the overall span of the vectors.

[a] here, hat means "with v_j removed"

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

$$a_1 v_1 + \ldots + a_m v_m = 0.$$

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

$$v_j = -\frac{a_1}{a_j} v_1 - \ldots - \frac{a_{j-1}}{a_j};$$

hence v_j is in the span of $v_1, \ldots, v_{j-1}$.

Now, suppose $u \in \operatorname{span}(v_1, \ldots, v_m)$. Then there exist $b_1, \ldots, b_m \in \mathbb{F}$ such that

$$u = b_1 v_1 + \ldots + b_m v_m.$$

If we replace v_j with 2.1.1, the resulting list consists only of $v_1, \ldots, \hat{v_j}, \ldots, v_m$; thus we see that u is in the span of the list. □

Theorem 2.1.1: Length of Linearly Independent List and Span

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. Left as an exercise for the reader. Try starting with $u_1, \ldots, u_m \in V$ a list of linearly independent vectors, and $v_1, \ldots, v_n \in V$ a spanning list of V, and show that $m \leq n$. Use the Linear Dependence Lemma to iteratively add u_i and remove w_j; eventually, we are left with a list with all u_i, and optionally some w_j.
To see why we cannot have more u than w, if that were the case, then $u_1, \ldots, u_n$ would span V, but $u_{n+1}, \ldots, u_m$ would be linearly independent, a contradiction. Thus $m \leq n$. □

Intuitively, every subspace of a finite-dimensional vector space is also finite-dimensional.

Proposition 2.1.2: 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 construct a spanning list of U:

- If $U = \{0\}$, then U is finite-dimensional and we are done, so choose a non-zero $v_1 \in U$.
- If $U = \operatorname{span}(v_1, \ldots, v_{j-1})$, then U is finite-dimensional and we are done; otherwise, if $U \neq \operatorname{span}(v_1, \ldots, v_{j-1})$, choose a vector $v_j \in U$ such that $v_j \notin \operatorname{span}(v_1, \ldots, v_{j-1})$ (equivalently, $v_1, \ldots, v_j$ is linearly independent).

After the process, we are left with a linearly independent spanning list of U. Since U is a subspace of V, this linearly independent list cannot be longer than the length of V's basis (aka spanning list), and so U is finite-dimensional as well. □

§2.2 Bases

Spanning lists and linearly independent lists go hand in hand; now, we bring these concepts together.

Definition 2.2.1: Basis

A **basis** of V is a list of vectors in V that is both linearly independent and spans V.

Example 17. 1. $e_1, e_2, \ldots, e_n = \{(1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,\ldots,1)\}$ *is the* ***standard basis*** *for* $\mathbb{F}^n$.

2. $(1,2),\ (3,5)$ *is another basis for* $\mathbb{R}^2$; *however,* $(1,2),\ (3,5),\ (4,13)$ *spans* $\mathbb{R}^2$ *but is not linearly independent.*

3. $1, x, \ldots, x^m$ *is a basis of* $\mathcal{P}_m(\mathbb{R})$.

Bases are incredibly useful in constructing unique vectors in a vector space.

Proposition 2.2.1: Criterion for Bases

A list $v_1, \ldots, v_n \in V$ is a basis of V if and only if every $v \in V$ can be written *uniquely* in the form

$$v = a_1v_1 + \ldots + a_nv_n,$$

where $a_1, \ldots, a_n \in \mathbb{F}$.

Proof. First, suppose $v_1, \ldots, v_n$ is a basis of V. Let $v \in V$. Since $V = \text{span}(v_1, \ldots, v_n)$, we have, for $a_i \in \mathbb{F}$

$$v = a_1v_1 + \ldots + a_nv_n.$$

Suppose, for $c_i \in \mathbb{F}$, that

$$v = c_1v_1 + \ldots + c_nv_n.$$

This implies

$$(v - v) = 0 = \sum_{i=1}^{n}(a_i - c_i)v_i.$$

Since $v_1, ..., v_n$ is linearly independent, all $a_i - c_i = 0$, and so $a_i = c_i$.
Now, suppose $v \in V$ can be represented uniquely by $a_1v_1+\ldots+a_nv_n$. Clearly, $v_1, \ldots, v_n$ spans V; and given

$$0 = a_1v_1 + \ldots + a_nv_n,$$

since the representation is unique, we must have $a_i = 0$ (since otherwise, if any $a_i \neq 0$, then the representation wouldn't be unique); hence $v_1, \ldots, v_n$ is linearly independent as well, and so is a basis of V. □

Spanning lists may not be bases of V due to linear independence, while linearly independent lists may not be bases due to spanning. Thus, we look for ways to create bases from spanning/linearly independent lists.
For spanning lists, we get the idea that we can discard "useless" vectors while maintaining span.

Proposition 2.2.2: Spanning List contains a Basis

Every spanning list in a vector space V can be reduced to a basis of V.

Proof. Let $B = v_1, \ldots, v_n$ span V; we iteratively remove "useless" vectors until left with a basis.

- If $v_1 = 0$, delete v_1 from B; otherwise, leave B unchanged.
- If $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$, delete v_j from B. Otherwise, leave B unchanged.

After iterating n times (through the entire list), we are left with a list B that spans V (since we only removed linearly dependent vectors in the span of the other vectors). Moreover, B is linearly independent, since no vector $v_i \in B$ is in the span of the previous vectors. Hence B is a basis of V. □

An easy corollary follows:

Corollary 2.2.1: 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; using the previous result, we reduce this list B to a basis. □

Now, we work with linearly independent lists; we can add "uncovered" vectors until such a list spans V while maintaining linear independence.

Proposition 2.2.3: Linearly Independent List Extends to a Basis

Every linearly independent list of vectors in a finite-dimensional vector space V can be extended to a basis of V.

Proof. Let $u_1, \ldots, u_m \in V$ be a linearly independent list, and let $w_1, \ldots, w_n \in V$ be a basis for V. Thus the list

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

spans V. Using the procedure before, we reduce this list to a basis of V; this basis has all of the u's, since $u_1, \ldots, u_m$ is linearly independent, and some of the w's. □

For example, suppose we have $(2, 3, 4), (9, 6, 8) \in \mathbb{R}^3$. Using $e_1, e_2, e_3 \in \mathbb{R}^3$ as the standard basis, the procedure results in a basis $(2, 3, 4), (9, 6, 8), (0, 1, 0)$.

We finish with some subspaces of V; intuitively, two subspaces can be combined to form V.

Proposition 2.2.4: 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; so let $u_1, \ldots, u_m$ be a basis of U. Since $u_1, \ldots, u_m \in V$ is linearly independent, extend the list to a basis $u_1, \ldots, u_m, w_1, \ldots, w_n$ of V, and let $W = \operatorname{span}(w_1, \ldots, w_n)$. To prove $V = U \oplus W$, we need to show

$$V = U + W \text{ and } U \cap W = \{0\}.$$

For any $v \in V$, since $u_1, \ldots, u_m, w_1, \ldots, w_n$ is a basis for v, we have $a_i, b_i \in \mathbb{F}$ such that

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

Since $\sum_{i=1}^m a_i u_i \in U$, $\sum_{j=1}^n b_j w_j \in W$, we have $v = u + w$, $u \in U$, $w \in W$. Thus $v \in U + W$, and so $V = U + W$.
Now, suppose $v \in U \cap W$. Then we have, for $a_i, b_i \in \mathbb{F}$,

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

and so

$$\sum_{i=1}^m a_i u_i - \sum_{j=1}^n b_j w_j = 0.$$

Since $u_1, \ldots, u_m, w_1, \ldots, w_n$ is a basis and so linearly independent, every $a_i, b_j = 0$, and so $v = 0(u_1 + \ldots + u_m) = 0(w_1 + \ldots + w_n) = 0$. Hence $U \cap W = \{0\}$, and so

$$V = U \oplus W.$$

□

§2.3 Dimension

With a space such as $\mathbb{R}^2$ or $\mathbb{R}^3$, we get the notion of two or three dimensions. For each, we see that their bases are that length (e.g. $(1,0),(0,1)$ is length two, $(1,0,0),(0,1,0),(0,0,1)$ is length three); hence, it seems that dimension is dependent on length of basis. However, in a finite-dimensional vector space, this would only make sense if every basis had the same length. Fortunately, this is the case:

Proposition 2.3.1: Basis Length does not Depend on Basis

Any two bases of a finite-dimensional vector space have the same length.

Proof. Let B_1, B_2 be bases of a finite-dimensional vector space V. Then B_1 is linearly independent, and B_2 spans V. Hence the length of B_1 is less than or equal to the length of B_2. Swapping roles (e.g. B_1 spans V, B_2 linearly independent), we see that the length of B_1 is greater than or equal to the length of B_2. Hence their lengths are equal. □

Now, we can formally define dimension:

Definition 2.3.1: Dimension

The **dimension** of a finite-dimensional vector space V, denoted $\dim V$, is the length of any basis B of V.

For example, $\mathcal{P}_m\mathbb{F}$ has dimension $m+1$, because the basis $1, x, \ldots, x^m \in \mathcal{P}_m(\mathbb{F})$ has length $m+1$.
As expected, the dimension of a subspace is less than or equal to the dimension of the vector space.

Proposition 2.3.2: Dimension of a Subspace

Let V be a finite-dimensional vector space, and let U be a subspace of V. Then any basis $u_1, \ldots, u_m \in U$ is a linearly independent list in V, and any basis $v1, \ldots, v_n \in$

V is a spanning list of V, so $m \leq n$, or $\dim U \leq \dim V$.

Bases require two properties: linearly independent and spanning. It turns out that given any two out of three properties from length, linearly independent, and spanning, we can deduce whether a list is a basis. Clearly, if a list is linearly independent and spanning, it has the right length (e.g. $\dim V$), but the other two conditions (right length + lin. ind., or right length + spanning) may not be as obvious.

Proposition 2.3.3: Linearly Independent List of Right Length is a Basis

Given a finite-dimensional vector space V, every linearly independent list of vectors in V with length $\dim V$ is a basis for V.

Proof. Let $n = \dim V$ and $v_1, \ldots, v_n \in V$ be a linearly independent list in V, and suppose $v_1, \ldots, v_n$ does not span V. Then there exists some vector $v \in V$ such that $v \notin \operatorname{span}(v_1, \ldots, v_n)$, so $v_1, \ldots, v_n, v$ is linearly independent. However, the length of every linearly independent list in V is less than or equal to the length of any spanning set of V; and since $n + 1 > n$ (V has a basis a.k.a. spanning list of length n), this is a contradiction. Thus $v_1, \ldots, v_n$ spans V, and therefore is a basis.

See Axler for an alternative proof; their logic makes more mental leaps. □

Proposition 2.3.4: Spanning List of 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. Let $v_1, \ldots, v_n$ span V, and suppose $v_1, \ldots, v_n$ is linearly dependent. Then there exists $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$, so $v_1, \ldots, \hat{v}_j, \ldots, v_n$ spans V. However, again the length of every linearly independent list is less than or equal to the length of any spanning list of V; and since $n-1 < n$ (V has a basis a.k.a. linearly independent list of length n), this is a contradiction. Thus $v_1, \ldots, v_n$ is linearly independent, and therefore is a basis. □

Finally, we find the dimension of the sum of two subspaces of V. Intuitively, we keep all basis vectors of U_1, U_2, while discarding any "duplicates" (imagine a Venn Diagram!).

Proposition 2.3.5: Dimension of a Sum

If $U_1,\ U_2$ 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. Let $u_1, \ldots, u_m$ be a basis for $U_1 \cap U_2$; thus $\dim U_1 \cap U_2 = m$. Because $u_1, \ldots, u_m$ is a basis, it is linearly independent in both U_1 and U_2; thus extend the list to a basis of U_1, $u_1, \ldots, u_m, v_1, \ldots, v_j$, and a basis of U_2, $u_1, \ldots, u_m, w_1, \ldots, w_k$. Thus $\dim U_1 = m + j$, and $\dim U_2 = m + k$.
We now show that

$$u_1, \ldots, u_m, v_1, \ldots, v_j, w_1, \ldots, w_k$$

is a basis for $U_1 + U_2$; this will show that $\dim(U_1 + U_2) = m + j + k = \dim U_1 + \dim U_2 - \dim U_1 \cap U_2$.

Clearly span$(u_1, \ldots, u_m, v_1, \ldots, v_j, w_1, \ldots, w_k)$ contains U_1 and U_2 (since any vector in either U_1 or U_2 could be made with a combination), and hence equals $U_1 + U_2$. To show linear independence, suppose

$$a_1u_1 + \ldots + a_mu_m + b_1v_1 + \ldots + b_jv_j + c_1w_1 + \ldots + c_kw_k = 0.$$

Rewriting, we get

$$c_1w_1 + \ldots + c_kw_k = -a_1u_1 - \ldots - a_mu_m - b_1v_1 - \ldots - b_jv_j,$$

and so $c_iw_i \in U_1$. Since all w_i are in U_2, this implies $c_1w_1 + \ldots + c_kw_k \in U_1 \cap U_2$. Since $u_1, ..., u_m$ is a basis for $U_1 \cap U_2$, we can rewrite as

$$c_1w_1 + \ldots + c_kw_k = d_1u_1 + \ldots + d_mu_m.$$

However, $u_1, \ldots, u_m, w_1, \ldots, w_k$ is linearly independent, so all $c_i, d_i = 0$. Thus we get, from the original equation,

$$a_1u_1 + \ldots + a_mu_m + b_1v_1 + \ldots + b_jv_j = 0.$$

Since this list is a basis, all $a_i, b_i = 0$. Thus all $a, b, c = 0$, and so the list is linearly independent.
Thus $u_1, \ldots, u_m, v_1, \ldots, v_j, w_1, \ldots, w_k$ is a basis for $U_1 + U_2$. □

Chapter 3

Linear Maps

§3.1 Linear Maps

Definition 3.1.1: Linear Maps

Let V, W be vector spaces over a field $\mathbb{F}$. A function

$$\begin{aligned} T : V &\longrightarrow W \\ v &\longmapsto T(v) \in W. \end{aligned}$$

is a **linear map** if it satisfies, given $v_1, v_2 \in V,\ \lambda \in \mathbb{F}$:

1. $T(v_1 + v_2) = T(v_1) + T(v_2) \in W$.
2. $T(\lambda v) = \lambda T(v)$.

Proposition 3.1.1: Linear Maps Preserve 0

If $T : V \to W$ is a linear map, then $T(\mathbf{0}) = \mathbf{0}$.

Proof. We have

$$\begin{aligned} T(0) &= T(0 + 0) \\ &= T(0) + T(0). \end{aligned}$$

Adding the additive inverse of $T(0)$ to both sides, we have

$$0 = T(0).$$

$\square$

Proposition 3.1.2: Combination of Linearity Properties

A function $T : V \to W$ is linear if and only if

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$$

for all $v_1, v_2 \in V,\ \alpha, \beta \in \mathbb{F}$.

Example 18. *Let V, W be any vector spaces over $\mathbb{F}$.*

1. *The* ***zero map***

$$\begin{aligned} 0 : V &\longrightarrow W \\ v &\longmapsto 0(v) = 0 \end{aligned}$$

   is a linear map.

2. *The* ***identity map***

$$\begin{aligned} I : V &\longrightarrow V \\ v &\longmapsto I(v) = v \end{aligned}$$

   is a linear map.

3. *Any linear map*

$$\begin{aligned} T : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto T(x) = ax \end{aligned}$$

   is a linear map.

Proposition 3.1.3: Linear Maps in R

Let $T : \mathbb{R} \to \mathbb{R}$ be a linear map. Then there is some $a \in \mathbb{R}$ such that $T(x) = ax$ for all $x \in \mathbb{R}$.

Proof. Let $a = T(1)$. Then for any $x \in \mathbb{R}$,

$$T(x) = T(x \cdot 1) = x \cdot T(1) = ax.$$

$\square$

Example 19. *Say $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map such that $T(1, 0) = (2, 1)$ and $T(0, 1) = (1, -1)$. What else do we know?*

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