Linear Algebra Done Right Notes

Steven Labalme

September 5, 2021

Contents

List of Figures

Chapter 1

Vector Spaces

1.A $\mathbb{R}^n$ and $\mathbb{C}^n$

10/27:

- Assumed familiarity with the set $\mathbb{R}$ of real numbers.
- **Complex number**: An ordered pair (a, b), where $a, b \in \mathbb{R}$, but we will write this as $a + bi$.
  - The set of all complex number is denoted by $\mathbb{C}$:

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}^{[1]}$$

  - Definitions of **addition** and **multiplication** on $\mathbb{C}$ are given, but I know these.
- Properties of complex arithmetic:
  - **Commutativity**: $\alpha + \beta = \beta + \alpha$ and $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbb{C}$.
  - **Associativity**: $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ and $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbb{C}$.
  - **Identities**: $\lambda + 0 = \lambda$ and $\lambda 1 = \lambda$ for all $\lambda \in \mathbb{C}$.
  - **Additive inverse**: For every $\alpha \in \mathbb{C}$, there exists a unique $\beta \in \mathbb{C}$ such that $\alpha + \beta = 0$.
  - **Multiplicative inverse**: For every $\alpha \in \mathbb{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbb{C}$ such that $\alpha\beta = 1$.
  - **Distributive property**: $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbb{C}$.
- "The properties above are proved using the familiar properties of real numbers and the definitions of complex addition and multiplication" (Axler, 2015, p. 3).
- $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$.
  - Any theorem proved with $\mathbb{F}$ holds when $\mathbb{F}$ is replaced with $\mathbb{R}$ and when $\mathbb{F}$ is replaced with $\mathbb{C}$.
- **Scalar**: A number or magnitude. This word is commonly used to differentiate a quantity from a **vector** quantity.
- Subtraction and division are defined.
- Properties of exponents are defined.
- The set $\mathbb{R}^2$, which can be conceived as a plane, is the set of all **ordered pairs** of real numbers:

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

[1] The complex numbers equal the set of numbers $a + bi$ such that a and b are elements of the real numbers.

- The set $\mathbb{R}^3$, which can be conceived as ordinary space, is the set of all **ordered triples** of real numbers:

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

- "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:

$$(x_1, \dots, x_n)$$

  Two lists are equal if and only if they have the same length and the same elements in the same order" (Axler, 2015, p. 5).
- **Ordered pair**: A list of length 2.
- **Ordered triple**: A list of length 3.
- ***n*-tuple**: A list of length n.
- Although lists are sometimes discussed without specifying their length, a list must, by definition, have a finite length, i.e. $(x_1, x_2, \dots)$ is not a list.
- A list of length 0 looks like this: $()$.
  - Such an object is defined to avoid trivial exceptions to theorems.
- Lists vs. **sets**: In lists, order matters and repetitions have meaning. In sets, order and repetitions are irrelevant.
- "$\mathbb{F}^n$ is the set of all lists of length n of elements of $\mathbb{F}$:

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

  For $(x_1, \dots, x_n) \in \mathbb{F}^n$ and $j \in \{1, \dots, n\}$, we say that x_j is the j^{th} **coordinate** of $(x_1, \dots, x_n)$" (Axler, 2015, p. 6).
- For help in conceiving higher dimensional spaces, consider reading Abbott (1952). This is an amusing account of how $\mathbb{R}^3$ would be perceived by creatures living in $\mathbb{R}^2$.
- **Addition** (in $\mathbb{F}^n$): Add corresponding coordinates:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

- For a simpler notation, use a single letter to denote a list of n numbers.
  - **Commutativity** (of addition in $\mathbb{F}^n$): If $x, y \in \mathbb{F}^n$, then $x + y = y + x$.
  - However, the proof still requires the more formal, cumbersome list notation.
- **0**: The list of length n whose coordinates are all 0:

$$0 = (0, \dots, 0)$$

  - Although the ambiguity in the use of "0" on the left vs. right side of the equation may seem confusing, context can always differentiate between which definition is needed.
- A picture can help visualize $\mathbb{R}^2$ because $\mathbb{R}^2$ can be sketched on 2-dimensional surfaces such as paper.
  - A typical element of $\mathbb{R}^2$ is a point $x = (x_1, x_2)$.
  - However, points are generally though of as an arrow starting at the origin and ending at x, as shown below.

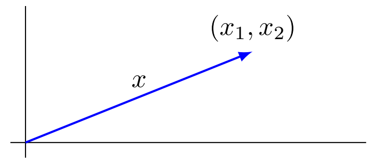

Figure 1.1: $x \in \mathbb{R}^2$ can be conceived as a point or a vector.

- When thought of as an arrow, x is called a **vector**.
- When translated without varying length or direction, it is still the same vector.
- Remember that these pictures are aids — although we cannot visualize higher dimensional vector spaces, the algebraic elements are as rigorously defined as those of $\mathbb{R}^2$.

- Addition has a simple geometric interpretation in $\mathbb{R}^2$.
- If we want to add $x + y$, slide y so that its initial point coincides with the terminal point of x. The sum is the vector from the tail of x to the head of y.

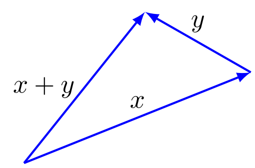

Figure 1.2: Vector addition.

- "For $x \in \mathbb{F}^n$, the **additive inverse** of x, denoted $-x$, is the vector $-x \in \mathbb{F}^n$ such that

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

In other words, if $x = (x_1, \ldots, x_n)$, then $-x = (-x_1, \ldots, -x_n)$" (Axler, 2015, p. 9).

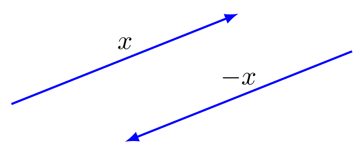

Figure 1.3: A vector and its additive inverse.

- For $x \in \mathbb{R}^2$, $-x$ is the vector parallel to x with the same length but in the opposite direction.

- **Product (scalar multiplication)**: When multiplying $\lambda \in \mathbb{F}$ and $x \in \mathbb{F}^n$, multiply each coordinate of x by λ:

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

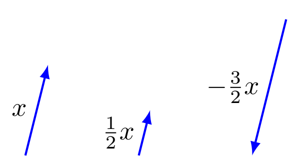

Figure 1.4: Scalar multiplication.

- **Field**: A "set containing at least two distinct elements called 0 and 1, along with operations of addition and multiplication satisfying all the properties" of complex arithmetic (see earlier in this section) (Axler, 2015, p. 10).

1.B Definition of Vector Space

- **Addition (on a set V)**: "A function that assigns an element $u + v \in V$ to each pair of elements $u, v \in V$" (Axler, 2015, p. 12).
- **Scalar multiplication (on a set V)**: "A function that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbb{F}$ and each $v \in V$" (Axler, 2015, p. 12).
- **Vector space**: "A set V along with an addition and a scalar multiplication on V such that the following properties hold:" (Axler, 2015, p. 12).

  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 $a, b \in \mathbb{F}$

  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 $a, b \in \mathbb{F}$ and all $u, v \in V$

- To be more precise, V depends on $\mathbb{F}$, so sometimes we say V is a **vector space over** $\mathbb{F}$.
  - For example, $\mathbb{R}^n$ is only a vector space over $\mathbb{R}$, not $\mathbb{C}$.
- **Real vector space**: A vector space over $\mathbb{R}$.
- **Complex vector space**: A vector space over $\mathbb{C}$.
- $\mathbb{F}^\infty$ is a vector space.
- $\mathbb{F}^S$ denotes the set of functions from S to $\mathbb{F}$.
  - For example, $\mathbb{R}^{[0,1]}$ is the "set of real-valued functions on the interval $[0, 1]$" (Axler, 2015, p. 14).
  - You can think of $\mathbb{F}^n$ as $\mathbb{F}^{\{1,2,\dots,n\}}$.
- Elementary properties of vector spaces:

  Theorem 1.1 (Unique additive identity). *A vector space has a unique additive identity.*

  Proof. Suppose 0 and $0'$ are both additive identities in V. Then

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

  The first equality holds due to 0 being an additive identity. The second holds due to commutativity. The third holds due to $0'$ being an additive identity. Thus, $0 = 0'$, and V has only one additive identity. ■

  Theorem 1.2 (Unique additive inverse). *Each element $v \in V$ has a unique additive inverse.*

Proof. Same idea:

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

■

Theorem 1.3 (The number 0 times a vector)**.** *$0v = 0 \; \forall \; v \in V$, where 0 on the left side is a scalar and 0 on the right side is a vector (the additive identity of V).*

Proof. Since this property asserts something about both scalar multiplication and the additive identity, the distributive property (the only part of the definition of a vector space that connects scalar multiplication and vector addition) must be used in the proof.

$$\begin{aligned} 0v &= (0 + 0)v \\ 0v &= 0v + 0v \\ 0v - 0v &= 0v + 0v - 0v \\ 0 &= 0v \end{aligned}$$

■

Theorem 1.4 (A number times the vector 0)**.** *$a0 = 0 \; \forall \; a \in \mathbb{F}$, where 0 is a vector.*

Proof. Same as above. ■

Theorem 1.5 (The number -1 times a vector)**.** *$(-1)v = -v \; \forall \; v \in V$, where -1 is a scalar and $-v$ is the additive inverse of v.*

Proof.

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

■

1.C Subspaces

- **Subspace**: A subset U of V that is a vector space under the same definition of addition and scalar multiplication as on V, e.g., satisfies the following three conditions.

  additive identity
  $0 \in U$

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

  closed under scalar multiplication
  $a \in \mathbb{F}$ and $u \in U$ implies $au \in U$

- The other conditions can be derived from the above 3.
- When we look at subspaces within the differentiable functions, the logical foundation of calculus appears.
- The subspaces of $\mathbb{R}^2$ are $\{0\}$, $\mathbb{R}^2$, and any straight line through the origin.
- The subspaces of $\mathbb{R}^3$ are $\{0\}$, $\mathbb{R}^3$, any straight line through the origin, and any flat plane through the origin.
- **Sum of subsets**: If $U_1, \dots, U_n$ are subsets of V, their **sum** (denoted $U_1 + \cdots + U_n$) is the set of all possible sums of elements of $U_1, \dots, U_n$:

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

- The sum of subspaces is the smallest containing subspace.
  - Clearly, the sum of subspaces is a subspace (satisfies 3 tenets).
  - The sum of subspaces contains every original element (u_1 plus the 0 from u_2, etc.). Any subspace containing all of these elements must contain every finite sum of them (by definition). Thus, no smaller subspace can be created than that of the sum of every element.
- **Direct sum**: A sum of subspaces where each element of $U_1 + \cdots + U_m$ can be written in only one way as a sum $u_1 + \cdots + u_m$.
  - $U_1 \oplus \cdots \oplus U_m$ denotes $U_1 + \cdots + U_m$ if $U_1 + \cdots + U_m$ is a direct sum.
- A sum of subspaces is a direct sum if and only if the only way to write 0 as a sum of elements is by summing the 0 of each subset.
- A sum of subspaces U and W is a direct sum if and only if $U \cap W = \{0\}$.

Chapter 2

Finite-Dimensional Vector Spaces

2.A Span and Linear Independence

9/3:

- **Linear combination** (of a list $v_1, \dots, v_m$ of vectors in V): A vector of the form $a_1v_1 + \cdots + a_mv_m$, where $a_1, \dots, a_m \in \mathbb{F}$.
- **Span** (of $v_1, \dots, v_m$): The set of all linear combinations of a list of vectors $v_1, \dots, v_m$ in V. *Also known as* **linear span**. *Denoted by* $\text{span}(v_1, \dots, v_m)$. *Given by*

$$\text{span}(v_1, \dots, v_m) = \{a_1v_1 + \cdots + a_mv_m : a_1, \dots, a_m \in \mathbb{F}\}$$

  - We define $\text{span}() = \{0\}$.
- Span as a subspace.

  Theorem 2.1. *The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.*

  Proof. Let $v_1, \dots, v_m \in V$ be a list of vectors. We will first prove that $\text{span}(v_1, \dots, v_m)$ is a subspace of V. We will then prove that it contains all of the vectors in the list. Lastly, we will prove that any proper subset of $\text{span}(v_1, \dots, v_m)$ either doesn't contain all the vectors in the list or is not a subspace of V. Let's begin.

  To prove that $\text{span}(v_1, \dots, v_m)$ is a subspace of V, it will suffice to show that $\text{span}(v_1, \dots, v_m)$ contains the additive identity, $\text{span}(v_1, \dots, v_m)$ is closed under addition, and $\text{span}(v_1, \dots, v_m)$ is closed under scalar multiplication. By the definition of $\text{span}(v_1, \dots, v_m)$, we know that $0v_1 + \cdots + 0v_m = 0 \in \text{span}(v_1, \dots, v_m)$. If $a_1v_1 + \cdots + a_mv_m \in \text{span}(v_1, \dots, v_m)$ and $b_1v_1 + \cdots + b_mv_m \in \text{span}(v_1, \dots, v_m)$, then naturally $(a_1v_1 + \cdots + a_mv_m) + (b_1v_1 + \cdots + b_mv_m) = (a_1 + b_1)v_1 + \cdots + (a_m + b_m)v_m \in \text{span}(v_1, \dots, v_m)$. Lastly, if $a_1v_1 + \cdots + a_mv_m \in \text{span}(v_1, \dots, v_m)$ and $\lambda \in \mathbb{F}$, then naturally $\lambda(a_1v_1 + \cdots + a_mv_m) = (\lambda a_1)v_1 + \cdots + (\lambda a_m)v_m \in \text{span}(v_1, \dots, v_m)$.

  By setting every $a_i = 0$ except $a_j = 1$, we can guarantee that $v_j \in \text{span}(v_1, \dots, v_m)$ for all $j \in [m]$.

  Suppose for the sake of contradiction that there exists a smaller subspace U of V that contains $v_1, \dots, v_m$. It follows that there exists a vector $u \in \text{span}(v_1, \dots, v_m)$ such that $u \notin U$. Since $u \in \text{span}(v_1, \dots, v_m)$, $u = a_1v_1 + \cdots + a_mv_m$ for some $a_1, \dots, a_m \in \mathbb{F}$. However, by definition, $v_1, \dots, v_m \in U$, so since U is closed under addition and scalar multiplication, their linear combination $a_1v_1 + \cdots + a_mv_m = u \in U$, a contradiction. ■

- If $\text{span}(v_1, \dots, v_m) = V$, we say that $v_1, \dots, v_m$ **spans** V.
- **Finite-dimensional vector space**: A vector space such that some list of vectors in it spans the space.

- Recall that by definition, every list has finite length.

- **Polynomial** (with coefficients in $\mathbb{F}$): A function $p : \mathbb{F} \to \mathbb{F}$ such that there exist $a_0, \dots, a_m \in \mathbb{F}$ such that
$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$
for all $z \in \mathbb{F}$.
- $\boldsymbol{\mathcal{P}(\mathbb{F})}$: The set of all polynomials with coefficients in $\mathbb{F}$.
  - $\mathcal{P}(\mathbb{F})$, under the usual addition and scalar multiplication, is a vector space over $\mathbb{F}$.
  - Thus, $\mathcal{P}(\mathbb{F})$ is a subspace of $\mathbb{F}^{\mathbb{F}}$.
- We will later prove that the coefficients of a polynomial uniquely determine it.
- **Degree** (of a polynomial p): The number m, where $p = a_0 + a_1 z + \cdots + a_m z^m$ and $a_m \neq 0$. *Denoted by* $\deg p = m$.
  - The polynomial $p(z) = 0$ is said to have degree $-\infty$.
- $\boldsymbol{\mathcal{P}_m(\mathbb{F})}$: The set of all polynomials with coefficients in $\mathbb{F}$ and degree at most m, where m is a nonnegative integer.
  - $\mathcal{P}_m(\mathbb{F}) = \operatorname{span}(1, z, \dots, z^m)$ is a finite-dimensional vector space for all nonnegative integers m.
- **Infinite-dimensional vector space**: A vector space that is not finite dimensional.
  - $\mathcal{P}(\mathbb{F})$ is infinite-dimensional.
- **Linearly independent** (list $v_1, \dots, v_m$): A list $v_1, \dots, v_m$ of vectors in V such that the only choice of $a_1, \dots, a_m \in \mathbb{F}$ that makes $a_1 v_1 + \cdots + a_m v_m = 0$ is $a_1 = \cdots = a_m = 0$.
  - We also let the empty list be linearly independent.
- $v_1, \dots, v_m$ is linearly independent if and only if each vector in $\operatorname{span}(v_1, \dots, v_m)$ has only one representation as a linear combination of $v_1, \dots, v_m$.
- If some vectors are removed from a linearly independent list, the remaining list is also linearly independent.
  - Suppose $v_1, \dots, v_m$ is linearly independent. Suppose $v_1, \dots, v_n$ is not linearly independent, with $n < m$. Then $a_1 v_1 + \cdots + a_n v_n = 0$ for some $a_1, \dots, a_n \in \mathbb{F}$ such that $a_i \neq 0$ for all $i \in [n]$. But then $a_1 v_1 + \cdots + a_n v_n + 0 v_{n+1} + \cdots + 0 v_m = 0$, a contradiction.
- **Linearly dependent** (list $v_1, \dots, v_m$): A list $v_1, \dots, v_m$ of vectors in V that is not linearly independent.
  - In other words, $v_1, \dots, v_m$ are linearly dependent if there exist $a_1, \dots, a_m \in \mathbb{F}$, not all 0, such that $a_1 v_1 + \cdots + a_m v_m = 0$.
- The following is an important and oft-used lemma.

  Lemma 2.2 (Linear Dependence Lemma)**.** *Suppose $v_1, \dots, v_m$ is a linearly dependent list in V. Then there exists $j \in \{1, \dots, m\}$ such that the following hold:*

  (a) $v_j \in \operatorname{span}(v_1, \dots, v_{j-1})$*;*

  (b) if the j^{th} term is removed from $v_1, \dots, v_m$, the span of the remaining list equals $\operatorname{span}(v_1, \dots, v_m)$*.*

Proof. We divide into two cases (the list is $v_1 = 0$, and the list is $v_1, \dots, v_m$).

If the list is $v_1 = 0$, then the list is linearly dependent. Choose $j = 1$. Clearly, $v_1 \in \operatorname{span}() = \{0\}$ by definition. Additionally, $\operatorname{span}() = \{0\} = \{a_1 0 : a_1 \in \mathbb{F}\} = \operatorname{span}(v_1)$, as desired.

Since $v_1, \dots, v_m$ is linearly dependent, there exist $a_1, \dots, a_m \in \mathbb{F}$, not all 0, such that $a_1 v_1 + \cdots + a_m v_m = 0$. Let j be the largest element of $\{1, \dots, m\}$ such that $a_j \neq 0$. Then

$$\begin{aligned} 0 &= a_1 v_1 + \cdots + a_j v_j + 0 v_{j+1} + \cdots + 0 v_m \\ -a_j v_j &= a_1 v_1 + \cdots + a_{j-1} v_{j-1} \\ v_j &= -\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1} \end{aligned}$$

It follows that $v_j \in \operatorname{span}(v_1, \dots, v_{j-1})$, as desired.

Now clearly $\operatorname{span}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m) \subset \operatorname{span}(v_1, \dots, v_m)$. In the other direction, suppose $u = c_1 v_1 + \cdots + c_m v_m \in \operatorname{span}(v_1, \dots, v_m)$. Then

$$u = c_1 v_1 + \cdots + c_{j-1} v_{j-1} + c_j \left(-\frac{a_1}{a_j} v_1 - \cdots - \frac{a_{j-1}}{a_j} v_{j-1} \right) + c_{j+1} v_{j+1} + \cdots + c_m v_m \in \operatorname{span}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m)$$

as desired. ■

- We next prove an immediate consequence of the Linear Dependence Lemma.

  Theorem 2.3. *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 that $u_1, \dots, u_m$ is linearly independent in V, and that $w_1, \dots, w_n$ spans V. We must prove that $m \leq n$. To do so, it will suffice to use the following m-step process.

  Step 1: Let $B = w_1, \dots, w_n$. Adding any $v \in V$ 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, since $u_1 \neq 0$ (it's part of a linearly independent list, and thus cannot be written as $0u_i$ for any u_i), the Linear Dependence Lemma asserts that we can remove one of the w_i's such that the new list B consisting of u_1 and the remaining w_i's spans V.

  Step j: The list B from step $j-1$ spans V. Thus, as before, adjoin vector u_j to B, placing it just after $u_1, \dots, u_{j-1}$. It follows by the Linear Dependence Lemma that one of the following vectors (i.e., one of the w_i's) is in $\operatorname{span}(u_1, \dots, u_j)$, so we can remove it and know that the list comprised of $u_1, \dots, u_j$ followed by the remaining w_i's spans V.

  After step m, we have added all of 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.[1] ■

  - This theorem allows us to prove results such as no list of 4 or more vectors is linearly independent in $\mathbb{R}^3$ (since $(1,0,0), (0,1,0), (0,0,1)$ spans $\mathbb{R}^3$), and no list of fewer than 4 vectors spans $\mathbb{R}^4$ (since $(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)$ is linearly independent in $\mathbb{R}^4$).

- We now can rigorously prove that subspaces of finite-dimensional vector spaces are also finite-dimensional.

  Theorem 2.4. *Every subspace of a finite-dimensional vector space is finite-dimensional.*

  Proof. Let V be finite-dimensional, and suppose for the sake of contradiction that U is infinite-dimensional subspace of V. Since V is finite-dimensional, there exists a list of vectors $v_1, \dots, v_m$ such that $\operatorname{span}(v_1, \dots, v_m) = V$. To arrive at a contradiction, we will construct a linearly independent list of vectors in U of length $m+1$, contradicting Theorem 2.3.

[1] We should be able to do this more rigorously via induction on m.

Since U is infinite-dimensional, there is no list of vectors in U spans it. Thus, if we choose $u_1 \in U$, we know that $\operatorname{span}(u_1) \neq U$. It follows since $\operatorname{span}(u_1) \subset U$ (as we know from the closure of U) that there exists $u_2 \in U$ such that $u_2 \notin \operatorname{span}(u_1)$. However, we will still have that $\operatorname{span}(u_1, u_2) \neq U$. More importantly, though, since $u_2 \notin \operatorname{span}(u_1)$ and $u_1 \notin \operatorname{span}()$, the Linear Dependence Lemma implies that u_1, u_2 is linearly independent. We can clearly continue in this fashion up to $u_1, \dots, u_{m+1}$, as desired. ■

2.B Bases

- **Basis** (of V): A list of vectors in V that is linearly independent and spans V.
- **Standard basis** (of $\mathbb{F}^n$): The list $(1,0,\dots,0),(0,1,0,\dots,0),\dots,(0,\dots,0,1)$.
- Determining whether a list of vectors is a basis:

Theorem 2.5. *A list $v_1, \dots, v_n$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form*

$$v = a_1v_1 + \cdots + a_nv_n$$

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

Proof. Suppose first that $v_1, \dots, v_n$ is a basis of V. Let $v \in V$ be arbitrary. We will first show that v can be written in the desired form, and then we will show that this writing is unique. Let's begin. By the definition of a basis, $v_1, \dots, v_n$ spans V. Thus, $\operatorname{span}(v_1, \dots, v_n) = V$. It follows that $v \in \operatorname{span}(v_1, \dots, v_n)$, which implies by the definition of span that $v = a_1v_1 + \cdots + a_nv_n$ where $a_1, \dots, a_n \in \mathbb{F}$, as desired. Now suppose for the sake of contradiction $v = c_1v_1 + \cdots + c_nv_n$ as well, where $c_1, \dots, c_n \in \mathbb{F}$ and $c_j \neq a_j$ for some $i \in [n]$. Then

$$\begin{aligned} 0 &= v - v \\ &= (a_1v_1 + \cdots + a_nv_n) - (c_1v_1 + \cdots + c_nv_n) \\ &= (a_1 - c_1)v_1 + \cdots + (a_n - c_n)v_n \end{aligned}$$

Since at least $a_j - c_j \neq 0$ but the above sum still does equal 0, we have that $v_1, \dots, v_n$ are not linearly independent, a contradiction.

Now suppose that every $v \in V$ can be written uniquely in the form $v = a_1v_1 + \cdots + a_nv_n$. To prove that $v_1, \dots, v_n$ is a basis of V, it will suffice to show that $v_1, \dots, v_n$ spans V and is linearly independent. Let's start with the first claim. Clearly, $\operatorname{span}(v_1, \dots, v_n) \subset V$, and since every $v \in V$ may be written as a linear combination of $v_1, \dots, v_n$, we know that every $v \in V$ is an element of $\operatorname{span}(v_1, \dots, v_n)$, as desired. On the other hand, we know that $0 = 0v_1 + \cdots + 0v_n$ and 0 can only be written in this unique form. Thus, the only choice of $a_1, \dots, a_m \in \mathbb{F}$ that makes $a_1v_1 + \cdots + a_nv_n = 0$ is $a_1 = \cdots = a_n = 0$, proving that $v_1, \dots, v_n$ is linearly independent. ■

- Finding the basis in a spanning list.

Theorem 2.6. *Every spanning list in a vector space can be reduced to a basis of the vector space.*

Proof. Let $v_1, \dots, v_n$ span V. We induct on n. For the base case $n = 0$, if () spans V, then since () is linearly independent by definition, we are done. Now suppose inductively that we have proven that every spanning list in V of length n can be reduced to a basis of V; we wish to prove that every spanning list in V of length $n+1$ can be reduced to a basis of V. Let $v_1, \dots, v_{n+1}$ span V. If $v_1, \dots, v_{n+1}$ is linearly independent, we are done. If $v_1, \dots, v_{n+1}$ is linearly dependent, then by the Linear Dependence Lemma, we can remove a vector from the list without changing its span. Our new list only has length n, so by the inductive hypothesis, it will reduce to a basis of V. ■

- Proving the existence of a basis in a finite-dimensional vector space.

Theorem 2.7. *Every finite-dimensional vector space has a basis.*

Proof. Let V be finite-dimensional. As such, there exists a list $v_1, \dots, v_n$ of vectors in V that spans V. It follows by Theorem 2.6 that some sublist of $v_1, \dots, v_n$ is a basis of V, as desired. ■

- Extending a linearly independent list into a basis (basically the dual of Theorem 2.6).

Theorem 2.8. *Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.*

Proof. Let $u_1, \dots, u_m$ be a linearly independent list of vectors in V. By Theorem 2.7, V has a basis $w_1, \dots, w_n$. It follows that $u_1, \dots, u_m, w_1, \dots, w_n$ spans V. Thus, by Theorem 2.6, which removes the first linearly dependent vector in $u_1, \dots, u_m, w_1, \dots, w_n$ (necessarily one of the w_i's since $u_1, \dots, u_m$ are linearly independent) via the Linear Dependence Lemma, there exists a sublist of $u_1, \dots, u_m, w_1, \dots, w_n$ containing $u_1, \dots, u_m$ that is a basis of V. ■

- Finding orthogonal complements.

Theorem 2.9. *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. Since V is finite-dimensional, Theorem 2.4 asserts that U is finite-dimensional. Thus, by Theorem 2.7, U has a basis $u_1, \dots, u_m$. It follows by Theorem 2.8 that there exist $w_1, \dots, w_n \in V$ such that $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V. Let $W = \operatorname{span}(w_1, \dots, w_n)$.

To prove that $U \oplus W = V$, it will suffice to show that

$$U + W = V \qquad\qquad U \cap W = \{0\}$$

To prove the first equation, it will suffice to show that for any vector $v \in V$, $v = u + w$ for $u \in U$ and $w \in W$. But since $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V, we have by the definition of span that

$$v = \underbrace{a_1 u_1 + \cdots + a_m u_m}_{u} + \underbrace{b_1 w_1 + \cdots + b_n w_n}_{w}$$

as desired.

To prove the second equation, let $v \in U \cap W$ be arbitrary. Then since $v \in U$ and $u_1, \dots, u_m$ is a basis of U, we have that $v = a_1 u_1 + \cdots + a_m u_m$ where $a_1, \dots, a_m \in \mathbb{F}$. Similarly, we have that $v = b_1 w_1 + \cdots + b_m w_n$ where $b_1, \dots, b_n \in \mathbb{F}$. It follows that

$$\begin{aligned} 0 &= v - v \\ &= a_1 u_1 + \cdots + a_m u_m - b_1 w_1 - \cdots - b_n w_n \end{aligned}$$

But since $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V, $u_1, \dots, u_m, w_1, \dots, w_n$ is linearly independent. It follows that $a_1 = \cdots = a_m = b_1 = \cdots = b_n = 0$. Therefore, $v = a_1 u_1 + \cdots + a_m u_m = 0$, as desired. ■

  - Note that this same basic idea extends to the infinite-dimensional case, but that proof requires considerably more advanced tools.

2.C Dimension

- It would be nice to define the dimension of a vector space as the length of a basis, but in order to do this we must first show that all bases have the same length.

Theorem 2.10. *Any two bases of a finite-dimensional vector space have the same length.*

Proof. Let B_1, B_2 be two arbitrary bases of V. Since B_1 is linearly independent in V and B_2 spans V, Theorem 2.3 asserts that $\operatorname{len} B_1 \leq \operatorname{len} B_2$. Similarly, since B_2 is linearly independent in V and B_1 spans V, Theorem 2.3 asserts that $\operatorname{len} B_2 \leq \operatorname{len} B_1$. Therefore, $\operatorname{len} B_1 = \operatorname{len} B_2$, as desired. ■

- **Dimension** (of V finite-dimensional): The length of any basis of V. *Denoted by* $\mathbf{dim}\, \boldsymbol{V}$.
- We can now give the expected inequality regarding the dimension of subspaces with respect to the dimension of the vector space.

  Theorem 2.11. *If V is finite-dimensional, and U is a subspace of V, then* $\dim U \leq \dim V$.

  Proof. Since V is finite dimensional, Theorem 2.4 asserts that U is finite-dimensional. Thus, Theorem 2.7 implies that they have bases $B_U = u_1, \dots, u_m$ and $B_V = v_1, \dots, v_n$. Therefore, since B_U is linearly independent in V and B_V spans V, Theorem 2.3 asserts that $\dim U = \operatorname{len} B_U \leq \operatorname{len} B_V = \dim V$, as desired. ■

- Note that while there exists an order-preserving bijection over addition and scalar multiplication by elements of $\mathbb{R}$ between $\mathbb{R}^2$ and $\mathbb{C}$, $\dim \mathbb{R}^2 = 2 \neq 1 = \dim \mathbb{C}$. Thus, when we talk about the dimension of a vector space, the role played by the choice of $\mathbb{F}$ cannot be neglected.
- Consider a list of vectors in a vector space that has the same length as a basis of the vector space. To check that this list is a basis, we can show that it suffices to show that the list is linearly independent *or* spans the vector space, as we will now prove.

  Theorem 2.12. *Suppose V is finite-dimensional. Then every linearly independent list of vectors in V with length* $\dim V$ *is a basis of V.*

  Proof. Let $\dim V = n$, and let $v_1, \dots, v_n$ be linearly independent. By Theorem 2.8, we can extend $v_1, \dots, v_n$ to a basis of V. However, since every basis of V has length n by Theorem 2.10, we need not add any vectors to $v_1, \dots, v_n$ to make it a basis; in other words, $v_1, \dots, v_n$ already is a basis. ■

  Theorem 2.13. *Suppose V is finite-dimensional. Then every spanning list of vectors in V with length* $\dim V$ *is a basis of V.*

  Proof. The proof is symmetric to the proof of Theorem 2.12. ■

- Lastly, we tackle the dimension of a sum of subspaces, which bears a resemblance to the principle of inclusion-exclusion.

  Theorem 2.14. *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. By Theorem 2.7, $U_1 \cap U_2$ (which we can prove is a subspace in its own right) has a basis, which we may denote $u_1, \dots, u_m$. Since $u_1, \dots, u_m$ is linearly independent in U_1, Theorem 2.8 asserts that it can be extended to a basis $u_1, \dots, u_m, v_1, \dots, v_j$ of U_1. Similarly, it can be extended to a basis $u_1, \dots, u_m, w_1, \dots, w_k$ of U_2.

  To prove that $u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k$ is a basis of $U_1 + U_2$, it will suffice to show that it is linearly independent and spans $U_1 + U_2$.

  To show that $u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k$ is linearly independent, it will suffice to verify that

  $$\begin{aligned} a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_jv_j + c_1w_1 + \cdots + c_kw_k = 0 \\ c_1w_1 + \cdots + c_kw_k = -a_1u_1 - \cdots - a_mu_m - b_1v_1 - \cdots - b_jv_j \end{aligned}$$

  implies that all of the coefficients equal 0. Suppose that the above equation holds. Then since $c_1w_1 + \cdots + c_kw_k$ can be written as a linear combination of the basis vectors of U_1, $c_1w_1 + \cdots + c_kw_k \in U_1$.

Additionally, since $c_1w_1 + \cdots + c_kw_k$ is a linear combination of vectors in U_2, $c_1w_1 + \cdots + c_kw_k \in U_2$. Thus, $c_1w_1 + \cdots + c_kw_k \in U_1 \cap U_2$. It follows that $c_1w_1 + \cdots + c_kw_k$ can be written as a linear combination of $u_1, \dots, u_m$, i.e.,

$$\begin{aligned} c_1w_1 + \cdots + c_kw_k &= d_1u_1 + \cdots + d_mu_m \\ 0 &= d_1u_1 + \cdots + d_mu_m - c_1w_1 - \cdots - c_kw_k \end{aligned}$$

for some $d_1, \dots, d_m \in \mathbb{F}$. But since $u_1, \dots, u_m, w_1, \dots, w_k$ is linearly independent as the basis of U_2, the above equation implies that $c_1 = \cdots = c_k = 0$. This implies that

$$0 = -a_1u_1 - \cdots - a_mu_m - b_1v_1 - \cdots - b_jv_j$$

meaning since $u_1, \dots, u_m, v_1, \dots, v_j$ is linearly independent as the basis of U_1, the above equation implies that $a_1 = \cdots = a_m = b_1 = \cdots = b_j = 0$, as desired.

To show that $u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k$ spans $U_1 + U_2$, it will suffice to show that all vectors in the list are elements of $U_1 + U_2$ (i.e., $\operatorname{span}(u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k) \subset U_1 + U_2$), and that every vector in $U_1 + U_2$ can be written as a linear combination of $u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k$ (i.e., that $U_1 + U_2 \subset \operatorname{span}(u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k)$). Since every vector in the list is an element of U_1 or U_2, we can show that it is an element of $U_1 + U_2$ by adding it to the additive identity of the other space. On the other hand, let $x \in U_1 + U_2$. Then $x = x_1 + x_2$, where $x_1 \in U_1$ and $x_2 \in U_2$. It follows that $x_1 = a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_jv_j$ and $x_2 = a_1'u_1 + \cdots + a_m'u_m + c_1w_1 + \cdots + c_kw_k$. Therefore, $x = (a_1 + a_1')u_1 + \cdots + (a_m + a_m')u_m + b_1v_1 + \cdots + b_jv_j + c_1w_1 + \cdots + c_kw_k$, as desired.

Having established that $u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k$ is a basis of $U_1 + U_2$, we have that

$$\begin{aligned} \dim(U_1 + U_2) &= m + j + k \\ &= (m + j) + (m + k) - m \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \end{aligned}$$

as desired. ■

Chapter 3

Linear Maps

3.A The Vector Space of Linear Maps

9/5:

- **Linear map** (from V to W): A function $T : V \to W$ with the following properties. *Also known as* **linear transformation**.

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

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

  - Note that for linear maps, Tv means the same as the more standard functional notation $T(v)$.

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

- **Zero map**: The function $0 \in \mathcal{L}(V, W)$ that takes each element of some vector space to the additive identity of another vector space. *Defined by*

$$0v = 0$$

- **Identity map**: The function $I \in \mathcal{L}(V, V)$ on some vector space that takes each element to itself. *Defined by*

$$Iv = v$$

- We can also formalize more esoteric linear processes as linear maps.
  - For example, $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$ can be thought of as the differentiation map $Dp = p'$. This formalizes the fact that $(f + g)' = f' + g'$ and $(\lambda f)' = \lambda f'$.
  - We can do the same with integration: Let $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathbb{R})$ be described by $Tp = \int_0^1 p(x)\,\mathrm{d}x$. This formalizes the fact that integrals are additive and homogenous.
  - Axler (2015) gives a number more examples.

- We now prove that there exists a unique linear map from a vector space of dimension n to any n vectors in another vector space.

  Theorem 3.1. *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 define a function $T : V \to W$. We then show that T satisfies the specified property. After that, we show that it is a linear map. Lastly, we show that it is unique. Let's begin.

Let $T : V \to W$ be defined by

$$T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n$$

for all $c_1v_1 + \cdots + c_nv_n \in V$. Note that this definition is valid since, by Theorem 2.5, each $v \in V$ can be written in the form $c_1v_1 + \cdots + c_nv_n$ where $c_1, \dots, c_n \in \mathbb{F}$.

To prove that $Tv_j = w_j$ for all $j = 1, \dots, n$, let each c_i in the above definition equal 0 save c_j, which we set equal to 1. Then we have

$$\begin{aligned} T(0v_1 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} + \cdots + 0v_n) &= 0w_1 + \cdots + 0w_{j-1} + 1w_j + 0w_{j+1} + \cdots + 0w_n \\ T(v_j) &= w_j \end{aligned}$$

as desired.

To prove that T is a linear map, it will suffice to verify additivity and homogeneity, which we may do as follows. Let $u, v \in V$ with $u = a_1v_1 + \cdots + a_nv_n$ and $v = c_1v_1 + \cdots + c_nv_n$, and let $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} T(u + v) &= T((a_1 + c_1)v_1 + \cdots + (a_n + c_n)v_n) \\ &= (a_1 + c_1)w_1 + \cdots + (a_n + c_n)w_n \\ &= Tu + Tv \end{aligned}$$

and

$$\begin{aligned} T(\lambda v) &= T(\lambda c_1v_1 + \cdots + \lambda c_nv_n) \\ &= \lambda c_1w_1 + \cdots + \lambda c_nw_n \\ &= \lambda Tv \end{aligned}$$

as desired.

Now suppose $\tilde{T} \in \mathcal{L}(V, W)$ satisfies $\tilde{T}v_j = w_j$ for all $j = 1, \dots, n$. To prove that $T = \tilde{T}$, it will suffice to show that $\tilde{T}(c_1v_1 + \cdots + c_nv_n) = T(c_1v_1 + \cdots + c_nv_n)$ for all $c_1v_1 + \cdots + c_nv_n \in V$. Let $c_1v_1 + \cdots + c_nv_n \in V$ be arbitrary. We know that $\tilde{T}(v_j) = w_j$ for all $j = 1, \dots, n$. It follows since $\tilde{T}$ is a linear map (specifically, since it's homogenous) that $c_jw_j = c_j\tilde{T}(v_j) = \tilde{T}(c_jv_j)$ for all $j = 1, \dots, n$. Similarly, the additivity of $\tilde{T}$ implies that

$$\begin{aligned} T(c_1v_1 + \cdots + c_nv_n) &= c_1w_1 + \cdots + c_nw_n \\ &= \tilde{T}(c_1v_1) + \cdots + \tilde{T}(c_nv_n) \\ &= \tilde{T}(c_1v_1 + \cdots + c_nv_n) \end{aligned}$$

as desired. ■

- **Sum** (of $S, T \in \mathcal{L}(V, W)$): The linear map $(S + T) \in \mathcal{L}(V, W)$ defined by $(S + T)(v) = Sv + Tv$ for all $v \in V$.
- **Product** (of $T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$): The linear map $(\lambda T) \in \mathcal{L}(V, W)$ defined by $(\lambda T)(v) = \lambda(Tv)$ for all $v \in V$.
- It follows that, under these definitions of addition and multiplication, $\mathcal{L}(V, W)$ is a vector space.
- **Product** (of $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$): The linear map $ST \in \mathcal{L}(U, W)$ defined by $(ST)(u) = S(Tu)$ for all $u \in U$.
  - Note that the product is just function composition, but most mathematicians do write ST instead of $S \circ T$.
- Linear maps (with the correct corresponding domains) satisfy the associativity, identity, and distributive properties. However, multiplication of linear maps is not necessarily commutative.

- $(T_1T_2)T_3 = T_1(T_2T_3)$.
- $TI_V = I_WT = T$ (note that if $T \in \mathcal{L}(V,W)$, $I_V \in \mathcal{L}(V,V)$ and $I_W \in \mathcal{L}(W,W)$).
- $(S_1+S_2)T = S_1T + S_2T$ and $S(T_1+T_2) = ST_1 + ST_2$.

- Linear maps send 0 to 0.

Theorem 3.2. *Suppose $T \in \mathcal{L}(V,W)$. Then $T(0) = 0$.*

Proof. By additivity, we have

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

as desired. ∎

3.B Null Spaces and Ranges

- **Null space** (of $T \in \mathcal{L}(V,W)$): The subset of V consisting of those vectors that T maps to 0. *Also known as* **kernel**. *Denoted by* $\textbf{null}\,\boldsymbol{T}$. *Given by*

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

- The null space is a subspace.

Theorem 3.3. *Suppose $T \in \mathcal{L}(V,W)$. Then* $\operatorname{null} T$ *is a subspace of V.*

Proof. To prove that $\operatorname{null} T$ is a subspace of V, it will suffice to show that $0 \in \operatorname{null} T$, $u, v \in \operatorname{null} T$ implies that $u+v \in \operatorname{null} T$, and $u \in \operatorname{null} T$ and $\lambda \in \mathbb{F}$ imply $\lambda u \in \operatorname{null} T$. Let's begin.

By Theorem 3.2, $T(0) = 0$. Therefore, $0 \in \operatorname{null} T$, as desired.

Let $u, v \in \operatorname{null} T$ be arbitrary. Then by additivity

$$T(u+v) = Tu + Tv = 0 + 0 = 0$$

so $u+v \in \operatorname{null} T$, as desired.

Let $u \in \operatorname{null} T$ and $\lambda \in \mathbb{F}$ be arbitrary. Then by homogeneity,

$$T(\lambda u) = \lambda Tu = \lambda 0 = 0$$

so $\lambda u \in \operatorname{null} T$, as desired. ∎

- **Injective** (function): A function $T : V \to W$ such that $Tu = Tv$ implies $u = v$. *Also known as* **one-to-one**.

- If 0 is the only vector that gets mapped to 0, then T is injective.

Theorem 3.4. *Let $T \in \mathcal{L}(V,W)$. Then T is injective if and only if* $\operatorname{null} T = \{0\}$.

Proof. Suppose first that T is injective. To prove that $\operatorname{null} T = \{0\}$, it will suffice to show that $0 \in \operatorname{null} T$ and for every $v \in \operatorname{null} T$, $v = 0$. By Theorem 3.3, $0 \in \operatorname{null} T$. Now let $v \in \operatorname{null} T$ be arbitrary. By the definition of the null space, we have $Tv = 0$. By Theorem 3.2, we have $T(0) = 0$. Thus, by transitivity, we have that $Tv = T(0)$. It follows by injectivity that $v = 0$, as desired.

Now suppose that $\operatorname{null} T = \{0\}$. To prove that T is injective, it will suffice to show that if $Tu = Tv$, then $u = v$. Suppose $u, v \in V$ satisfy $Tu = Tv$. Then

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

so $(u-v) \in \operatorname{null} T = \{0\}$. It follows that $u - v = 0$, i.e., that $u = v$, as desired. ∎

- **Range** (of $T \in \mathcal{L}(V,W)$): The subset of W consisting of those vectors that are of the form Tv for some $v \in V$. *Also known as* **image**. *Denoted by* **range T**. *Given by*

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

- The range is a subspace.

Theorem 3.5. *Suppose $T \in \mathcal{L}(V,W)$. Then* range T *is a subspace of W.*

Proof. To prove that range T is a subspace of W, it will suffice to show that $0 \in \operatorname{range} T$, $w_1, w_2 \in \operatorname{range} T$ implies that $(w_1 + w_2) \in \operatorname{range} T$, and $w \in \operatorname{range} T$ and $\lambda \in \mathbb{F}$ imply $\lambda w \in \operatorname{range} T$. Let's begin.

By the definition of a vector space, $0 \in V$. By Theorem 3.2, $T(0) = 0$. Therefore, $0 \in \operatorname{range} T$, as desired.

Let $w_1, w_2 \in \operatorname{range} T$ be arbitrary. Then there exist $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$. It follows by additivity that

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2$$

Therefore, since $v_1 + v_2 \in V$, we have that $(w_1 + w_2) \in \operatorname{range} T$, as desired.

Let $w \in \operatorname{range} T$ and $\lambda \in \mathbb{F}$ be arbitrary. Then there exists $v \in V$ such that $Tv = w$. It follows by homogeneity that

$$T(\lambda v) = \lambda Tv = \lambda w$$

Therefore, since $\lambda v \in V$, we have that $\lambda w \in \operatorname{range} T$, as desired. ■

- **Surjective** (function): A function $T : V \to W$ such that range $T = W$. *Also known as* **onto**.

- We now prove a very important theorem.

Theorem 3.6 (Fundamental Theorem of Linear Maps)**.** *Suppose V is finite-dimensional and $T \in \mathcal{L}(V,W)$. Then* range T *is finite-dimensional and*

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

Proof. By Theorem 3.3, null T is a subspace of V finite-dimensional. Thus, by Theorem 2.4, null T is finite-dimensional. It follows by Theorem 2.7 that we may let $u_1, \dots, u_m$ be a basis of null T. As a basis of a subspace of V, $u_1, \dots, u_m$ is a linearly independent list of vectors in V. Consequently, by Theorem 2.8, we may extend it to a basis $u_1, \dots, u_m, v_1, \dots, v_n$ of V.

Having established this terminology, we can now see that to prove that range T is finite-dimensional, it will suffice to show that $Tv_1, \dots, Tv_n$ spans it. To show that $\operatorname{span}(Tv_1, \dots, Tv_n) = \operatorname{range} T$, it will suffice to show that every $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{span}(Tv_1, \dots, Tv_n)$ is an element of range T and that every $Tv \in \operatorname{range} T$ is an element of $\operatorname{span}(Tv_1, \dots, Tv_n)$. Let $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{span}(Tv_1, \dots, Tv_n)$ be arbitrary. Then

$$\begin{aligned} b_1Tv_1 + \cdots + b_nTv_n &= T(b_1v_1 + \cdots + b_nv_n) \\ &= T(0u_1 + \cdots + 0u_m + b_1v_1 + \cdots + b_nv_n) \end{aligned}$$

Therefore, since $0u_1 + \cdots + 0u_m + b_1v_1 + \cdots + b_nv_n \in V$ by V's closure under addition and scalar multiplication, we have that $b_1Tv_1 + \cdots + b_nTv_n \in \operatorname{range} T$, as desired. Now let $Tv \in \operatorname{range} T$ be arbitrary. Since $v \in V$ and $u_1, \dots, u_m, v_1, \dots, v_n$ is a basis of V, Theorem 2.5 implies that $v = a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n$ for some $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{F}$. Therefore,

$$\begin{aligned} Tv &= T(a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n) \\ &= T(a_1u_1 + \cdots + a_mu_m) + T(b_1v_1 + \cdots + b_nv_n) \\ &= a_1Tu_1 + \cdots + a_mTu_m + b_1Tv_1 + \cdots + v_nTv_n \\ &= a_10 + \cdots + a_m0 + b_1Tv_1 + \cdots + v_nTv_n \\ &= b_1Tv_1 + \cdots + v_nTv_n \end{aligned}$$

where each $Tu_j = 0$ because each $u_j \in \operatorname{null} T$, so $Tv \in \operatorname{span}(Tv_1, \dots, Tv_n)$, as desired.

Before we can verify the equation from the theorem, we need to establish one last fact: that $Tv_1, \dots, Tv_n$ is linearly independent. Suppose $c_1, \dots, c_n \in \mathbb{F}$ make

$$\begin{aligned} c_1 Tv_1 + \cdots + c_n Tv_n &= 0 \\ T(c_1 v_1 + \cdots + c_n v_n) &= 0 \end{aligned}$$

It follows that $c_1 v_1 + \cdots + c_n v_n \in \operatorname{null} T$. Thus, since $u_1, \dots, u_m$ is a basis of $\operatorname{null} T$ by Theorem 2.5, we have that

$$\begin{aligned} c_1 v_1 + \cdots + c_n v_n &= d_1 u_1 + \cdots + d_m u_m \\ 0 &= d_1 u_1 + \cdots + d_m u_m - c_1 v_1 - \cdots - c_n v_n \end{aligned}$$

for some $d_1, \dots, d_m \in \mathbb{F}$. But since $u_1, \dots, u_m, v_1, \dots, v_n$ is linearly independent as the basis of V, the above equation implies that $c_1 = \cdots = c_n = 0$, as desired.

Having established that $u_1, \dots, u_m, v_1, \dots, v_n$ is a basis of V, $u_1, \dots, u_m$ is a basis of $\operatorname{null} T$, and $Tv_1, \dots, Tv_n$ spans $\operatorname{range} T$ and is linearly independent in $\operatorname{range} T$ (i.e., is a basis of $\operatorname{range} T$), we have that

$$\begin{aligned} \dim V &= m + n \\ &= \dim \operatorname{null} T + \dim \operatorname{range} T \end{aligned}$$

as desired. ■

- We can now prove that a linear map to a "smaller" vector space cannot be injective.

Theorem 3.7. *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. Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned} \dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T && \text{Fundamental Theorem of Linear Maps} \\ &\geq \dim V - \dim \operatorname{range} T && \text{Theorem 2.11} \\ &> 0 \end{aligned}$$

It follows that $\operatorname{null} T$ has a basis consisting of a list of one or more vectors. As a linearly independent list, naturally none of these vectors will be equal to the zero vector. Thus, since $\operatorname{null} T$ contains vectors other than 0, Theorem 3.4 implies that T is not injective. ■

- Similarly, we can prove that a linear map to a "bigger" vector space cannot be surjective.

Theorem 3.8. *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. Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned} \dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T && \text{Fundamental Theorem of Linear Maps} \\ &\leq \dim V && < \dim W \end{aligned}$$

Therefore, $\operatorname{range} T \neq W$, so T cannot be surjective. ■

- Theorems 3.7 and 3.8 allow us to express questions about systems of linear equations in terms of linear maps.

- For example, a question about a **homogenous** system of linear equations could be, "does there exist a nonzero solution to the homogenous system $\sum_{k=1}^{n} A_{1,k}x_k = 0, \dots, \sum_{k=1}^{n} A_{m,k}x_k = 0$?"
- If we define $T : \mathbb{F}^n \to \mathbb{F}^m$ by

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

we can express the system of equations as $T(x_1, \dots, x_n) = 0$ and ask instead, "is $\dim \operatorname{null} T > 0$?"

- **Homogenous** (system of linear equations): A system of m linear equations $\sum_{k=1}^{n} A_{1,k}x_k = c_1$ through $\sum_{k=1}^{n} A_{m,k}x_k = c_m$ such that the constant term $c_j = 0$ for all $j = 1, \dots, m$.
- Continuing with the linear equations example, we can rigorously show the following.

Theorem 3.9. *A homogenous system of linear equations with more variables than equations has nonzero solutions.*

Proof. In terms of the above, $T : \mathbb{F}^n \to \mathbb{F}^m$ where $n > m$. Thus, by Theorem 3.7, T is not injective. Consequently, by Theorem 3.4, $\dim \operatorname{null} T > 0$. Therefore, the system has nonzero solutions. ∎

Theorem 3.10. *An inhomogenous system of linear equations with more equations than variables has no solution for some choice of constant terms.*

Proof. In terms of the above, $T : \mathbb{F}^n \to \mathbb{F}^m$ where $m > n$. We want to know if there exists $(c_1, \dots, c_m) \in \mathbb{F}^m$ such that $T(x_1, \dots, x_n) \neq (c_1, \dots, c_m)$ for any $(x_1, \dots, x_n) \in \mathbb{F}^n$. In other words, we want to know if there exists $(c_1, \dots, c_m) \in \mathbb{F}^m$ such that $(c_1, \dots, c_m) \notin \operatorname{range} T$, i.e., if $\operatorname{range} T \neq \mathbb{F}^m$. But since $n < m$, Theorem 3.8 asserts that T is not surjective, meaning that $\operatorname{range} T \neq W$, as desired. ∎

- Note that while the past two results are typically proven with Gaussian elimination, the abstract approach taken here leads to cleaner proofs.

3.C Matrices

- ***m*-by-*n* matrix**: A rectangular array A of elements of $\mathbb{F}$ with m rows and n columns

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}$$

where m and n are positive integers.

  - The notation $A_{j,k}$ denotes the entry in row j, column k of A. In other words, the first index refers to the row nunber and the second index refers to the column number.
- **Matrix** (of $T \in \mathcal{L}(V, W)$ with respect to the bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W): The m-by-n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m$$

  - If the bases are not clear from context, then the notation $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ is used.
  - Another way of wording the definition states that the k^{th} column of $\mathcal{M}(T)$ consists of the scalars needed to write Tv_k as a linear combination of $w_1, \dots, w_m$.
  - Assuming standard bases, we "can think of the k^{th} column of $\mathcal{M}(T)$ as the T applied to the k^{th} standard basis vector" (Axler, 2015, p. 71).

- **Sum** (of two m-by-n matrices A, C): The m-by-n matrix $A+C$ defined by $(A+C)_{j,k} = A_{j,k} + C_{j,k}$.
  - Symbolically,
$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{pmatrix} = \begin{pmatrix} A_{1,1}+C_{1,1} & \cdots & A_{1,n}+C_{1,n} \\ \vdots & & \vdots \\ A_{m,1}+C_{m,1} & \cdots & A_{m,n}+C_{m,n} \end{pmatrix}$$
- Suppose $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.
- **Product** (of an m-by-n matrix A and $\lambda \in \mathbb{F}$): The m-by-n matrix λA defined by $(\lambda A)_{j,k} = \lambda A_{j,k}$.
  - Symbolically,
$$\lambda \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}$$
- Suppose $\lambda \in \mathbb{F}$ and $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.
- $\boldsymbol{\mathbb{F}^{m,n}}$: The set of all m-by-n matrices with entries in $\mathbb{F}$, where m and n are positive integers.
- We have that $\dim \mathbb{F}^{m,n} = mn$.
  - Note that a basis of $\mathbb{F}^{m,n}$ is the set of all m-by-n matrices that have 0s everywhere save a 1 in a single place.
- **Product** (of an m-by-n matrix A and an n-by-p matrix C): The m-by-p matrix AC defined by $(AC)_{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k}$.
  - We may derive this by noting that if $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = C$, $T : U \to V$ and $S : V \to W$, and $u_1, \dots, u_p$, $v_1, \dots, v_n$, and $w_1, \dots, w_m$ are bases, then
$$\begin{aligned} (ST)u_k &= S\left(\sum_{r=1}^{n} C_{r,k} v_r\right) \\ &= \sum_{r=1}^{n} C_{r,k} S v_r \\ &= \sum_{r=1}^{n} C_{r,k} \sum_{j=1}^{m} A_{j,r} w_j \\ &= \sum_{j=1}^{m} \left(\sum_{r=1}^{n} A_{j,r} C_{r,k}\right) w_j \end{aligned}$$
  - Matrix multiplication is not commutative, but is distributive and associative.
- Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. Then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.
- If A is an m-by-n matrix, then...
  - We let $A_{j,\cdot}$ denote the 1-by-n matrix consisting of row j of A;
  - We let $A_{\cdot,k}$ denote the m-by-1 matrix consisting of column k of A.
- Thus, if A is an m-by-n matrix and C is an n-by-p matrix, then $(AC)_{j,k} = A_{j,\cdot} C_{\cdot,k}$ for all $1 \leq j \leq m$ and $1 \leq k \leq p$.
- Similarly, $(AC)_{\cdot,k} = AC_{\cdot,k}$.

- Lastly, suppose A is an m-by-n matrix and $c = (c_1, \ldots, c_n)$ is an n-by-1 matrix. Then $Ac = c_1 A_{\cdot,1} + \cdots + c_n A_{\cdot,n}$.
  - In other words, Ac is a linear combination of the columns of A, with the scalars that multiply the columns coming from c.

References

Abbott, E. A. (1952). *Flatland: A romance of many dimensions* (sixth). Dover.
Axler, S. (2015). *Linear algebra done right* (Third). Springer.