Linear Algebra Done Right Notes

Steven Labalme

September 8, 2021

Contents

1 Vector Spaces **1**
- 1.A $\mathbb{R}^n$ and $\mathbb{C}^n$ 1
- 1.B Definition of Vector Space 4
- 1.C Subspaces 5

2 Finite-Dimensional Vector Spaces **7**
- 2.A Span and Linear Independence 7
- 2.B Bases 10
- 2.C Dimension 11

3 Linear Maps **14**
- 3.A The Vector Space of Linear Maps 14
- 3.B Null Spaces and Ranges 16
- 3.C Matrices 19
- 3.D Invertibility and Isomorphic Vector Spaces 21
- 3.E Products and Quotients of Vector Spaces 24
- 3.F Duality 27

4 Polynomials **33**

5 Eigenvalues, Eigenvectors, and Invariant Subspaces **36**
- 5.A Invariant Subspaces 36

References **39**

List of Figures

1.1 $x \in \mathbb{R}^2$ can be conceived as a point or a vector. 3
1.2 Vector addition. 3
1.3 A vector and its additive inverse. 3
1.4 Scalar multiplication. 3

3.1 Visualizing $v + U$. 25

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.

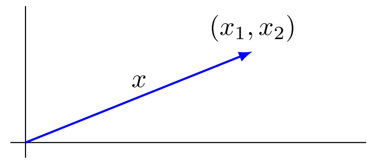

Figure 1.1: $x \in \mathbb{R}^2$ can be conceived as a point or a vector.

- 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.

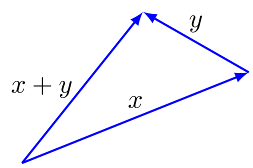

Figure 1.2: Vector addition.

- "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).

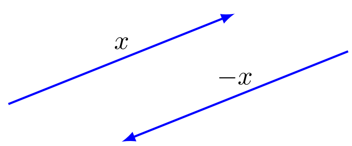

Figure 1.3: A vector and its additive inverse.

- 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:

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

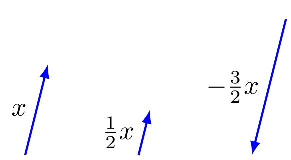

Figure 1.4: Scalar multiplication.

- **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*

$$v = a_1 v_1 + \cdots + a_n v_n$$

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 $v = a_1 v_1 + \cdots + a_n v_n$ where $a_1, \dots, a_n \in \mathbb{F}$, as desired. Now suppose for the sake of contradiction $v = c_1 v_1 + \cdots + c_n v_n$ as well, where $c_1, \dots, c_n \in \mathbb{F}$ and $c_j \neq a_j$ for some $i \in [n]$. Then

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

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 $v = a_1 v_1 + \cdots + a_n v_n$. 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 $a_1 v_1 + \cdots + a_n v_n = 0$ 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

$$v = \underbrace{a_1u_1 + \cdots + a_mu_m}_{u} + \underbrace{b_1w_1 + \cdots + b_nw_n}_{w}$$

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 $v = a_1u_1 + \cdots + a_mu_m$ where $a_1, \dots, a_m \in \mathbb{F}$. Similarly, we have that $v = b_1w_1 + \cdots + b_mw_n$ where $b_1, \dots, b_n \in \mathbb{F}$. It follows that

$$\begin{aligned} 0 &= v - v \\ &= a_1u_1 + \cdots + a_mu_m - b_1w_1 - \cdots - b_nw_n \end{aligned}$$

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, $v = a_1u_1 + \cdots + a_mu_m = 0$, 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., $\text{span}(u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k) \subset U_1 + U_2$), 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 $U_1 + U_2 \subset \text{span}(u_1, \dots, u_m, v_1, \dots, v_j, w_1, \dots, w_k)$). 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. ■

Exercises

1 Suppose V is finite-dimensional and U is a subspace of V such that $\dim U = \dim V$. Prove that $U = V$.

Proof. Since V is finite dimensional and U is a subspace of V, we have by Theorem 2.9 that there is a subspace W of V such that $V = U \oplus W$. It follows by Theorem 2.14 that

$$\dim W = \dim V - \dim U + \dim\{0\} = 0$$

Thus, $W = \{0\}$. Therefore,

$$\begin{aligned} V &= U \oplus W \\ &= \{u + w : u \in U, w \in W\} \\ &= \{u + 0 : u \in U\} \\ &= U \end{aligned}$$

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* $\operatorname{range} T$ *is a subspace of W.*

Proof. To prove that $\operatorname{range} T$ 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 $\operatorname{range} T = W$. *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* $\operatorname{range} T$ *is finite-dimensional and*

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

Proof. By Theorem 3.3, $\operatorname{null} T$ is a subspace of V finite-dimensional. Thus, by Theorem 2.4, $\operatorname{null} T$ is finite-dimensional. It follows by Theorem 2.7 that we may let $u_1, \dots, u_m$ be a basis of $\operatorname{null} T$. 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 $\operatorname{range} T$ 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 $\operatorname{range} T$ 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

$$\begin{aligned} c_1Tv_1 + \cdots + c_nTv_n &= 0 \\ T(c_1v_1 + \cdots + c_nv_n) &= 0 \end{aligned}$$

It follows that $c_1v_1 + \cdots + c_nv_n \in \operatorname{null} T$. Thus, since $u_1, \dots, u_m$ is a basis of $\operatorname{null} T$ by Theorem 2.5, we have that

$$\begin{aligned} c_1v_1 + \cdots + c_nv_n &= d_1u_1 + \cdots + d_mu_m \\ 0 &= d_1u_1 + \cdots + d_mu_m - c_1v_1 - \cdots - c_nv_n \end{aligned}$$

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 $c = (c_1, \dots, c_n)$ 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.

3.D Invertibility and Isomorphic Vector Spaces

9/6: - **Invertible** (linear map): A linear map $T \in \mathcal{L}(V, W)$ such that there exists a linear map $S \in \mathcal{L}(V, W)$ such that ST equals the identity map on V and TS equals the identity map on W.

- **Inverse** (of $T \in \mathcal{L}(V, W)$): The linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I_V$ and $TS = I_W$. *Denoted by* $\boldsymbol{T^{-1}}$.

- We now justify the use of the word "the" in the definition of the inverse.

  Theorem 3.11. *An invertible linear map has a unique inverse.*

  Proof. Suppose $T \in \mathcal{L}(V, W)$ is invertible and S_1, S_2 are inverses of T. Then

  $$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2$$

  as desired. ■

- We now give a criterion for invertibility.

  Theorem 3.12. *A linear map is invertible if and only if it is injective and surjective.*

  Proof. Let $T \in \mathcal{L}(V, W)$. Suppose first that T is invertible.

  To prove that T is injective, it will suffice to show that for all $u, v \in V$, $Tu = Tv$ implies that $u = v$. Let u, v be arbitrary elements of V that satisfy $Tu = Tv$. Then

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

  as desired.

  To prove that T is surjective, it will suffice to show that $\operatorname{range} T = W$. Since $\operatorname{range} T \subset W$, we need only show that $W \subset \operatorname{range} T$. Let $w \in W$ be arbitrary. Since $w = T(T^{-1}w)$ where $T^{-1}w \in V$, we have that $w \in \operatorname{range} T$, as desired.

  Now suppose that T is injective and surjective. To prove that T is invertible, we will define a function $S : W \to V$, prove that it is a linear map, prove that $TS = I_W$, and prove that $ST = I_V$. Let Sw be the unique element of V such that $T(Sw) = w$ (the surjectivity of T guarantees that there exists *an* element of V that T maps to w, and the injectivity of T guarantees the uniqueness of said element).

  To prove that S is a linear map, it will suffice to show that S is additive and homogenous. To verify additivity, first note that the additivity of T implies that

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

  But since the above equation implies that $Sw_1 + Sw_2$ is the unique element of V that T maps to $w_1 + w_2$, we have by the definition of S that $S(w_1 + w_2) = Sw_1 + Sw_2$. The proof is symmetric for homogeneity.

  To prove that $TS = I_W$, we need only appeal to the definition of S, which states that $(TS)w = T(Sw) = w$ for all $w \in W$. It immediately follows that $TS = I_W$.

  To prove that $ST = I_V$, first note that for all $v \in V$,

  $$T((ST)v) = (TS)(Tv) = I(Tv) = Tv$$

  It follows by the injectivity of T that $(ST)v = v$, i.e., that $ST = I_V$, as desired. ■

- **Isomorphism**: An invertible linear map.

- **Isomorphic** (vector spaces): Two vector spaces such that there exists an isomorphism from one vector space onto the other one.

- Isomorphic vector spaces have the same dimension.

Theorem 3.13. *Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.*

Proof. Suppose V, W are isomorphic finite-dimensional vector spaces over $\mathbb{F}$. Then there exists an isomorphism $T : V \to W$. By the definition of isomorphism, T is an invertible linear map, meaning by Theorem 3.12 that T is injective and surjective. Thus, since there exists an injective linear map $T : V \to W$, the contrapositive of Theorem 3.7 asserts that $\dim V \leq \dim W$. Additionally, since there exists a surjective linear map $T : V \to W$, the contrapositive of Theorem 3.8 asserts that $\dim V \geq \dim W$. Therefore, we have that $\dim V = \dim W$, as desired.

Now suppose that $\dim V = \dim W$. Let $v_1, \dots, v_n$ be a basis of V, and let $w_1, \dots, w_n$ be a basis of W. By Theorem 3.1, there exists a unique linear map $T : V \to W$ such that $Tv_j = w_j$ for each $j = 1, \dots, n$. To prove that T is an isomorphism, Theorem 3.12 tells us that it will suffice to show that it is injective and surjective. To show that T is surjective, it will suffice to show that $\operatorname{range} T = W = \operatorname{span}(w_1, \dots, w_n)$. But since $Tv_j = w_j \in \operatorname{range} T$ for all $j = 1, \dots, n$, $\operatorname{range} T \subset W$, and $\operatorname{range} T$ is a vector space (see Theorem 3.5), we have that $\operatorname{range} T = \operatorname{span}(w_1, \dots, w_n) = W$, as desired. To prove that T is injective, Theorem 3.4 tells us that it will suffice to show that $\operatorname{null} T = \{0\}$, i.e., that $\dim \operatorname{null} T = 0$. But since $\dim \operatorname{range} T = \dim W = \dim V$, we have by the Fundamental Theorem of Linear Maps that

$$\begin{aligned}
\dim \operatorname{null} T + \dim \operatorname{range} T &= \dim V \\
&= \dim W \\
&= \dim \operatorname{range} T \\
\dim \operatorname{null} T &= 0
\end{aligned}$$

as desired. ∎

- This result implies that every finite-dimensional vector space of dimension n is isomorphic to $\mathbb{F}^n$.
- It also allows us to formalize the link between linear maps from V to W and matrices in $\mathbb{F}^{m,n}$.

Theorem 3.14. *Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$.*

Proof. We have already established that $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$ and that $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$, so we already know that $\mathcal{M}$ is a linear map. To prove that it is invertible, Theorem 3.12 tells us that it will suffice to show that $\mathcal{M}$ is injective and surjective.

To show that $\mathcal{M}$ is injective, Theorem 3.4 tells us that it will suffice to verify that $\operatorname{null} \mathcal{M} = \{0\}$. Let $T \in \mathcal{L}(V, W)$ be arbitrary. If $\mathcal{M}(T) = 0$ (0 denoting the zero matrix), then

$$Tv_k = 0w_1 + \cdots + 0w_m = 0$$

for all $k = 1, \dots, n$. But since $v_1, \dots, b_n$ is a basis of V, this implies that $T = 0$ (0 denoting the zero transformation), as desired.

To show that $\mathcal{M}$ is surjective, it will suffice to verify that $\operatorname{range} \mathcal{M} = \mathbb{F}^{m,n}$. Clearly $\operatorname{range} \mathcal{M} \subset \mathbb{F}^{m,n}$, so we focus on the other direction. Let $A \in \mathbb{F}^{m,n}$ be arbitrary. Define $T \in \mathcal{L}(V, W)$ by

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

for $k = 1, \dots, n$. It follows by the definition of a matrix of a linear transformation that $\mathcal{M}(T) = A$, as desired. ∎

- We can now determine the dimension of the vector space of linear maps from one finite-dimensional vector space to another.

  Theorem 3.15. *Suppose V and W are finite-dimensional. Then $\mathcal{L}(V,W)$ is finite-dimensional and*

  $$\dim \mathcal{L}(V,W) = (\dim V)(\dim W)$$

  Proof. By Theorem 3.14, $\mathcal{L}(V,W)$ and $\mathbb{F}^{m,n}$ are isomorphic. Thus, by Theorem 3.13, $\mathcal{L}(V,W)$ and $\mathbb{F}^{m,n}$ have the same dimension. Therefore, we have that

  $$\begin{aligned}\dim \mathcal{L}(V,W) &= \dim \mathbb{F}^{m,n} \\ &= mn \\ &= (\dim V)(\dim W)\end{aligned}$$

  as desired. ∎

- **Matrix** (of $v \in V$ with respect to the basis $v_1,\dots,v_n$ of V): The n-by-1 matrix $\mathcal{M}(v)$ whose entries $A_{j,1}$ are defined by

  $$v = A_{1,1}v_1 + \cdots + A_{n,1}v_n$$

- We now show that the columns of the matrix of T are directly related to the effect T has on basis vectors.

  Theorem 3.16. *Suppose $T \in \mathcal{L}(V,W)$, $v_1,\dots,v_n$ is a basis of V, and $w_1,\dots,w_m$ is a basis of W. Let $1 \leq k \leq n$. Then*

  $$\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$$

  Proof. As an element of W, $Tv_k = c_1w_1 + \cdots + c_mw_m$ for some $c_1,\dots,c_m \in \mathbb{F}$. By the definition of the matrix of T, the values in column k are $c_1,\dots,c_m$. Similarly, by the definition of the matrix of Tv_k, the values in its one column are $c_1,\dots,c_m$, as desired. ∎

- Linear maps act like matrix multiplication.

  Theorem 3.17. *Suppose $T \in \mathcal{L}(V,W)$, $v \in V$, $v_1,\dots,v_n$ is a basis of V, and $w_1,\dots,w_m$ is a basis of W. Then*

  $$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$$

  Proof. Let $v = c_1v_1 + \cdots + c_nv_n$. Then by the linearity of T, $Tv = c_1Tv_1 + \cdots + c_nTv_n$. It follows by the linearity of $\mathcal{M}$, Theorem 3.16, and the fact that a matrix times a vector is equal to a linear combination of columns that

  $$\begin{aligned}\mathcal{M}(Tv) &= c_1\mathcal{M}(Tv_1) + \cdots + c_n\mathcal{M}(Tv_n) \\ &= c_1\mathcal{M}(T)_{\cdot,1} + \cdots + c_n\mathcal{M}(T)_{\cdot,n} \\ &= \mathcal{M}(T)\mathcal{M}(v)\end{aligned}$$

  as desired. ∎

- "Each m-by-n matrix A induces a linear map from $\mathbb{F}^{n,1}$ to $\mathbb{F}^{m,1}$, namely the matrix multiplication function that takes $x \in \mathbb{F}^{n,1}$ to $Ax \in \mathbb{F}^{m,1}$" (Axler, 2015, p. 85).
- **Operator**: A linear map from a vector space to itself.
- $\boldsymbol{\mathcal{L}(V)}$: The set of all operators on V.
  - Mathematically, $\mathcal{L}(V) = \mathcal{L}(V,V)$.

- On infinite-dimensional vector spaces, neither injectivity nor surjectivity, alone, implies invertibility.
- However, only one does on finite-dimensional vector spaces.

Theorem 3.18. *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivalent.*

(*a*) *T is invertible;*
(*b*) *T is injective;*
(*c*) *T is surjective.*

Proof. First, suppose that T is invertible. Then by Theorem 3.12, T is injective, as desired.

Second, suppose that T is injective. Then by Theorem 3.4, $\operatorname{null} T = \{0\}$. It follows by the Fundamental Theorem of Linear Maps that

$$\begin{aligned}\dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T \\ &= \dim V\end{aligned}$$

Thus, since $\operatorname{range} T$ has the same dimension as V and is a subspace of V (by Theorem 3.5), $\operatorname{range} T = V$. Therefore, T is surjective, as desired.

Third, suppose that T is surjective. Then $\operatorname{range} T = V$. It follows that $\dim \operatorname{range} T = \dim V$. Thus, by the Fundamental Theorem of Linear Maps, we have that

$$\begin{aligned}\dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \\ &= 0\end{aligned}$$

Consequently, by Theorem 3.4, T is injective. Therefore, by Theorem 3.12, T is invertible, as desired. ■

3.E Products and Quotients of Vector Spaces

- **Product** (of $V_1, \dots, V_m$): The set

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

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

$$(u_1, \dots, u_m) + (v_1, \dots, v_m) = (u_1 + v_1, \dots, u_m + v_m)$$

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

$$\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m)$$

- The product of n vector spaces over $\mathbb{F}$ is a vector space over $\mathbb{F}$, with addition and scalar multiplication defined as above.
- We can, for example, identify $\mathbb{R}^2 \times \mathbb{R}^3$ with $\mathbb{R}^5$ by constructing an isomorphism from every vector $((x_1, x_2), (x_3, x_4, x_5)) \in \mathbb{R}^2 \times \mathbb{R}^3$ to the vector $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$.
- The dimension of the product of vector spaces is equal to the sum of the dimensions of the component vector spaces.

Theorem 3.19. *Suppose $V_1, \dots, V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and*

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

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

- We now relate products and direct sums.

  Theorem 3.20. *Suppose that $U_1, \dots, U_m$ are subspaces of V. Define a linear map $\Gamma : U_1 \times \cdots \times U_m \to U_1 + \cdots + U_m$ by*

  $$\Gamma(u_1, \dots, u_m) = u_1 + \cdots + u_m$$

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

  Proof. Suppose first that Γ is injective. Then the only way to write 0 as a sum $u_1 + \cdots + u_m$, where each $u_j \in U_j$, is by taking each $u_j = 0$. It follows by the condition on direct sums that $U_1 + \cdots + U_m$ is a direct sum. The proof is symmetric in the reverse direction. ∎

  - Note that since Γ is surjective by the definition of $U_1 + \cdots + U_m$, the condition that Γ is injective could be changed to the condition that Γ is invertible.

- We can now prove that the dimensions add up in a direct sum.

  Theorem 3.21. *Suppose V is finite-dimensional and $U_1, \dots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is a direct sum if and only if*

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

  Proof. Suppose first that $U_1 + \cdots + U_m$ is a direct sum. Then by Theorem 3.20, there exists an invertible linear map Γ from $U_1 \times \cdots \times U_m$ to $U_1 + \cdots + U_m$. Thus, by Theorem 3.13, $U_1 \times \cdots \times U_m$ and $U_1 + \cdots + U_m$ have the same dimension. Therefore,

  $$\begin{aligned} \dim(U_1 + \cdots + U_m) &= \dim(U_1 \times \cdots \times U_m) \\ &= \dim U_1 + \cdots + \dim U_m && \text{Theorem 3.19} \end{aligned}$$

  as desired.

  The proof is symmetric in the other direction. ∎

- **Sum** (of $v \in V$ and U a subspace of V): The subset of V defined by

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

- **Affine subset** (of V): A subset of V of the form $v + U$ for some $v \in V$ and some subspace U of V.
- **Parallel** (subset to U): An affine subset $v + U$ of V.
- **Quotient space**: The set of all affine subsets of V parallel to U.

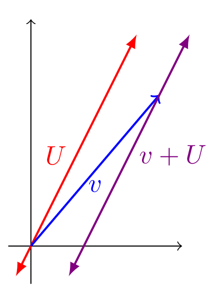

Figure 3.1: Visualizing $v + U$.

– Symbolically,

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

- Two affine subsets parallel to U are equal or disjoint.

Theorem 3.22. *Suppose U is a subspace of V and $v, w \in V$. Then the following are equivalent.*

(*a*) $v - w \in U$;

(*b*) $v + U = w + U$;

(*c*) $(v + U) \cap (w + U) \neq \emptyset$.

Proof. First, suppose that $v - w \in U$. Let $x \in v + U$ be arbitrary. Then $x = v + u$ for some $u \in U$. Now since $v - w \in U$, $u \in U$, and U is a subspace, we have that $v - w + u \in U$. Thus, $x = w - w + v + u = w + (v - w + u) \in w + U$. The proof is symmetric in the other direction. Therefore, $v + U = w + U$, as desired.

Second, suppose that $v + U = w + U$. Since U is nonempty ($0 \in U$ by definition), we know that $v + U \neq \emptyset \neq w + U$. Therefore, $(v + U) \cap (w + U) \supset \{0\} \neq \emptyset$, as desired.

Third, suppose that $(v + U) \cap (w + U) \neq \emptyset$. Then there exists x such that $x \in v + U$ and $x \in w + U$. It follows that $x = v + u_1$ and $x = w + u_2$ for some $u_1, u_2 \in U$. Thus, by transitivity, $v + u_1 = w + u_2$. Therefore, $v - w = u_2 - u_1 \in U$, as desired. ■

- **Sum** (of $v + U, w + U \in V/U$): The affine subset $(v + w) + U$.
- **Product** (of $v + U \in V/U$ and $\lambda \in \mathbb{F}$): The affine subset $(\lambda v) + U$.
- We now verify that the above operations are well-defined and prove that the quotient space is a vector space.

Theorem 3.23. *Suppose U is a subspace of V. Then V/U, with the operations of addition and scalar multiplication as defined above, is a vector space.*

Proof. The way affine subsets are defined, we may have $v + U = \hat{v} + U$ and yet have $v \neq \hat{v}$. Thus, we must first guarantee that the operations of addition and scalar multiplication, as defined above, are well-defined, i.e., that if $v + U = \hat{v} + U$ and $w + U = \hat{w} + U$, then $(v + w) + U = (\hat{v} + \hat{w}) + U$ and $(\lambda v) + U = (\lambda \hat{v}) + U$. Let's begin.

To confirm that addition as defined above is a well-defined operation, let $v, \hat{v}, w, \hat{w} \in V$ be such that $v + U = \hat{v} + U$ and $w + U = \hat{w} + U$. Then by Theorem 3.22, $v - \hat{v} \in U$ and $w - \hat{w} \in U$. It follows since U is a subspace that $(v - \hat{v}) + (w - \hat{w}) \in U$. Consequently, $(v + w) - (\hat{v} + \hat{w}) \in U$, so by Theorem 3.22 again, $(v + w) + U = (\hat{v} + \hat{w}) + U$, as desired.

Similarly, $v + U = \hat{v} + U$ implies $v - \hat{v} \in U$, implies $\lambda(v - \hat{v}) = \lambda v - \lambda \hat{v} \in U$, implies $(\lambda v) + U = (\lambda \hat{v}) + U$, as desired.

The remaining proof that V/U is a vector space is straightforward; note that $0 + U$ is the identity element and $(-v) + U$ is the additive inverse. ■

- **Quotient map**: The linear map $\pi : V \to V/U$ defined by $\pi(v) = v + U$ for all $v \in V$.
- We now give a formula for the dimension of a quotient space.

Theorem 3.24. *Suppose V is finite-dimensional and U is a subspace of V. Then*

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

Proof. Let π be the quotient map from V to V/U. From Theorem 3.22, we know that in order for $w+U = 0+U$, we must have $v-0 = v \in U$. Thus, $\pi(u) = 0$ if and only if $u \in U$, meaning $\operatorname{null}\pi = U$. Additionally, we clearly have that $\operatorname{range}\pi = V/U$. Therefore, by the Fundamental Theorem of Linear Maps, we have that

$$\begin{aligned}\dim V &= \dim\operatorname{null}\pi + \dim\operatorname{range}\pi \\ &= \dim U + \dim V/U \\ \dim V/U &= \dim V - \dim U\end{aligned}$$

as desired. ■

- Lastly, consider the fact that we can add any vector in the null space of a linear map T to an argument passed to T without changing its output. In other words, if $T \in \mathcal{L}(V,W)$, $v \in V$, and $u \in \operatorname{null}T$, then $T(v+u) = Tv + Tu = Tv$. We formalize this concept with the following definition.
- $\boldsymbol{\tilde{T}}$: The function from $V/(\operatorname{null}T)$ to W defined by $\tilde{T}(v + \operatorname{null}T) = Tv$, where $T \in \mathcal{L}(V,W)$.
- We now state a few basic results about $\tilde{T}$.

  Theorem 3.25. *Suppose $T \in \mathcal{L}(V,W)$. Then*

  (a) *$\tilde{T}$ is a linear map from $V/(\operatorname{null}T)$ to W;*
  (b) *$\tilde{T}$ is injective;*
  (c) *$\operatorname{range}\tilde{T} = \operatorname{range}T$;*
  (d) *$V/(\operatorname{null}T)$ is isomorphic to $\operatorname{range}T$.*

3.F Duality

9/7:

- **Linear functional** (on v): A linear map from V to $\mathbb{F}$.
  - In other words, a linear functional is an element of $\mathcal{L}(V,\mathbb{F})$.
- **Dual space** (of V): The vector space of all linear functionals on V. *Denoted by* $\boldsymbol{V'}$. *Also known as* $\boldsymbol{V^*}$. *Given by*

$$V' = \mathcal{L}(V,\mathbb{F})$$

- We now give a definition of the dimension of the dual space.

  Theorem 3.26. *Suppose V is finite-dimensional. Then V' is also finite-dimensional and*

$$\dim V' = \dim V$$

Proof. By Theorem 3.15, we have that

$$\begin{aligned}\dim V' &= \dim\mathcal{L}(V,\mathbb{F}) \\ &= (\dim V)(\dim\mathbb{F}) \\ &= (\dim V)(1) \\ &= \dim V\end{aligned}$$

as desired. ■

- **Dual basis** (of a basis $v_1,\dots,v_n$ of V): The list $\varphi_1,\dots,\varphi_n$ of elements of V', where each φ_j is the linear functional on V such that

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

where $v_1,\dots,v_n$ is a basis of V.

- We now verify that the dual basis of a basis of V is actually a basis of the dual space.

  Theorem 3.27. *Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.*

  Proof. Let $v_1, \dots, v_n$ be a basis of V, and let $\varphi_1, \dots, \varphi_n$ be the corresponding dual basis. Since the dual basis has length equal to the dimension of V' (by Theorem 3.26), Theorem 2.12 tells us that it will suffice to show that $\varphi_1, \dots, \varphi_n$ is linearly independent to confirm that it is a basis of V'. To do so, suppose

  $$a_1\varphi_1 + \cdots + a_n\varphi_n = 0$$

  where $a_1, \dots, a_n \in \mathbb{F}$ and 0 denotes the zero transformation. Since $(a_1\varphi_1 + \cdots + a_n\varphi_n)(v_j) = a_j$ for $j = 1, \dots, n$, we have that for any vector $c_1v_1 + \cdots + c_nv_n \in V$,

  $$(a_1\varphi_1 + \cdots + a_n\varphi_n)(c_1v_1 + \cdots + c_nv_n) = c_1a_1 + \cdots + c_na_n$$

  Therefore, the only way to guarantee that $c_1a_1 + \cdots + c_na_n = 0$ is to let $a_1 = \cdots = a_n = 0$, as desired. ∎

- **Dual map** (of $T \in \mathcal{L}(V, W)$): The linear map $T' \in \mathcal{L}(W', V')$ defined by

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

  for all $\varphi \in W'$. *Also known as* $\boldsymbol{T^*}$.

- We now prove some algebraic properties of dual maps.

  Theorem 3.28.

  (a) $(S + T)' = S' + T'$ *for all* $S, T \in \mathcal{L}(V, W)$.

  Proof. Let $S, T \in \mathcal{L}(V, W)$ be arbitrary. To prove that $(S + T)' = S' + T'$, it will suffice to show that $(S + T)'(\varphi) = (S' + T')(\varphi)$ for all $\varphi \in W'$. Let $\varphi \in W'$ be arbitrary. However, before we go into the main equality, it will be useful if we verify that $\varphi \circ (S + T) = \varphi \circ S + \varphi \circ T$. To do so, it will suffice to show that $(\varphi \circ (S + T))(v) = (\varphi \circ S + \varphi \circ T)(v)$ for all $v \in V$. Let $v \in V$ be arbitrary. Then

  $$\begin{aligned}
  (\varphi \circ (S + T))(v) &= \varphi((S + T)(v)) \\
  &= \varphi(S(v) + T(v)) \\
  &= \varphi(S(v)) + \varphi(T(v)) \\
  &= (\varphi \circ S)(v) + (\varphi \circ T)(v) \\
  &= (\varphi \circ S + \varphi \circ T)(v)
  \end{aligned}$$

  Now we can show that

  $$\begin{aligned}
  (S + T)'(\varphi) &= \varphi \circ (S + T) \\
  &= \varphi \circ S + \varphi \circ T \\
  &= S'(\varphi) + T'(\varphi) \\
  &= (S' + T')(\varphi)
  \end{aligned}$$

  as desired. ∎

  (b) $(\lambda T)' = \lambda T'$ *for all* $\lambda \in \mathbb{F}$ *and all* $T \in \mathcal{L}(V, W)$.

  Proof. The proof is symmetric to the proof of part (a). ∎

  (c) $(ST)' = T'S'$ *for all* $T \in \mathcal{L}(U, V)$ *and* $S \in \mathcal{L}(V, W)$.

Proof. Let $\varphi \in W'$ be arbitrary. Then

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

as desired. ■

- **Annihilator** (of $U \subset V$): The set

$$U^0 = \{\varphi \in V' : \varphi(u) = 0 \ \forall \ u \in U\}$$

- The annihilator is a subspace.

Theorem 3.29. *Suppose $U \subset V$. Then U^0 is a subspace of V'*

Proof. To prove that U^0 is a subspace of V', it will suffice to show that $0 \in U^0$, $\varphi, \psi \in U^0$ implies $\varphi + \psi \in U^0$, and $\varphi \in U^0$ and $\lambda \in \mathbb{F}$ imply $\lambda\varphi \in U^0$. Let's begin.

Since $0(u) = 0$ for all $u \in U$, $0 \in U^0$.

Let $\varphi, \psi \in U^0$ be arbitrary. Let $u \in U$ be arbitrary. Then $(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0$, as desired.

The proof is symmetric for scalar multiplication. ■

- Dimension of the annihilator.

Theorem 3.30. *Suppose V is finite-dimensional and U is a subspace of V. Then*

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

Proof. Let $i \in \mathcal{L}(U, V)$ be the identity map $i(u) = u$ for all $u \in U$. Then $i' : V' \to U'$ is a linear map. It follows by the Fundamental Theorem of Linear Maps that

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

Since $i'(\varphi) = \varphi \circ i = \varphi$ for all $\varphi \in V'$, and $U^0 = \{\varphi \in V' : \varphi = 0\}$, we have that $i'(\varphi) = 0$ for all $\varphi \in U^0$. Thus, $U^0 = \operatorname{null} i'$. Additionally, we have that $\dim V = \dim V'$ by Theorem 3.26. Lastly, let $\psi \in U'$ be arbitrary. Define $\psi \in V'$ by

$$\psi(v) = \begin{cases} \varphi(v) & v \in U \\ 0 & v \notin U \end{cases}$$

Thus, $i'(\psi) = \psi \circ i = \varphi$. It follows that $\varphi \in \operatorname{range} i'$. Consequently, $\operatorname{range} i' = U'$, so $\dim U = \dim U' = \dim \operatorname{range} i'$ by Theorem 3.26. Therefore, we have from the first equation and the three substitutions that

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

as desired.[1] ■

- We now describe the null space of T'.

Theorem 3.31. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then*

(*a*) $\operatorname{null} T' = (\operatorname{range} T)^0$.

Proof. First, let $\varphi \in \operatorname{null} T'$ be arbitrary. Then $T'(\varphi) = \varphi \circ T = 0$. It follows that $0 = (\varphi \circ T)(v) = \varphi(Tv)$ for all $v \in V$. But this means that φ is a linear functional that maps every element of $\operatorname{range} T$ to 0, i.e., that $\varphi \in (\operatorname{range} T)^0$. The proof is symmetric in the other direction. ■

[1] Note that we may also prove this by constructing a basis of U extending it to a basis of V, and showing that the extended portion of the dual basis is a basis of U^0.

(b) $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$.

Proof. We have that

$$\begin{aligned} \dim \operatorname{null} T' &= \dim(\operatorname{range} T)^0 && \text{Theorem 3.31a} \\ &= \dim W - \dim \operatorname{range} T && \text{Theorem 3.30} \\ &= \dim W - (\dim V - \dim \operatorname{null} T) && \text{Fundamental Theorem of Linear Maps} \\ &= \dim \operatorname{null} T + \dim W - \dim V \end{aligned}$$

as desired. ■

– Note that the proof of part (a) does not use the hypothesis that V, W are finite-dimensional, so the argument holds for infinite-dimensional vector spaces as well.

- T surjective is equivalent to T' injective.

Theorem 3.32. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is surjective if and only if T' is injective.*

Proof. Suppose first that T is surjective. Then $\operatorname{range} T = W$. It follows by Theorem 3.30 that

$$\dim(\operatorname{range} T)^0 = \dim W - \dim \operatorname{range} T = 0$$

meaning that $(\operatorname{range} T)^0 = \{0\}$. Thus, by Theorem 3.31a, $\operatorname{null} T' = \{0\}$. Therefore, by Theorem 3.4, T' is injective, as desired.

The proof is symmetric in the other direction. ■

- We now describe the range space of T'.

Theorem 3.33. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then*

(a) $\dim \operatorname{range} T' = \dim \operatorname{range} T$.

Proof. We have that

$$\begin{aligned} \dim \operatorname{range} T' &= \dim W' - \dim \operatorname{null} T' && \text{Fundamental Theorem of Linear Maps} \\ &= \dim W - \dim \operatorname{null} T' && \text{Theorem 3.26} \\ &= \dim W - \dim(\operatorname{range} T)^0 && \text{Theorem 3.31a} \\ &= \dim \operatorname{range} T && \text{Theorem 3.30} \end{aligned}$$

as desired. ■

(b) $\operatorname{range} T' = (\operatorname{null} T)^0$.

Proof. First, let $\varphi \in \operatorname{range} T'$ be arbitrary. Then there exists $\psi \in W'$ such that $\varphi = T'(\psi)$. Now let $v \in \operatorname{null} T$ be arbitrary. It follows that

$$\varphi(v) = (T'(\psi))(v) = (\psi \circ T)(v) = \psi(Tv) = \psi(0) = 0$$

Therefore, $\varphi \in (\operatorname{null} T)^0$, as desired.

Second, we have that

$$\begin{aligned} \dim \operatorname{range} T' &= \dim \operatorname{range} T && \text{Theorem 3.33a} \\ &= \dim V - \dim \operatorname{null} T && \text{Fundamental Theorem of Linear Maps} \\ &= \dim(\operatorname{null} T)^0 && \text{Theorem 3.30} \end{aligned}$$

Therefore, since Theorem 3.5 implies that $\operatorname{range} T'$ is a subspace of $(\operatorname{null} T)^0$ and $\dim \operatorname{range} T' = \dim(\operatorname{null} T)^0$, Exercise 2.C.1 asserts that $\operatorname{range} T' = (\operatorname{null} T)^0$, as desired. ■

- T injective is equivalent to T' surjective.

Theorem 3.34. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is injective if and only if T' is surjective.*

Proof. Suppose first that T is injective. Then by Theorem 3.4, $\operatorname{null} T = \{0\}$. Thus, since Theorem 3.2 asserts that $\varphi(0) = 0$ for any linear functional, we have that every linear functional is in the annihilator of $\operatorname{null} T$, i.e., that $(\operatorname{null} T)^0 = V'$. It follows by Theorem 3.33b that $\operatorname{range} T' = V'$. Therefore, T' is surjective, as desired.

The proof is symmetric in the other direction. ∎

9/8: - **Transpose** (of an m-by-n matrix A): The matrix obtained from A by interchanging the rows and columns. More specifically, the n-by-m matrix A^t whose entries are given by $(A^t)_{k,j} = A_{j,k}$. *Denoted by* $\boldsymbol{A^t}$.

- Properties of the transpose:

$$(A + C)^t = A^t + C^t \qquad\qquad (\lambda A)^t = \lambda A^t$$

- Transpose of a product.

Theorem 3.35. *If A is an m-by-n matrix and C is an n-by-p matrix, then*

$$(AC)^t = C^t A^t$$

Proof. We have that

$$\begin{aligned}
((AC)^t)_{k,j} &= (AC)_{j,k} \\
&= \sum_{r=1}^{n} A_{j,r} C_{r,k} \\
&= \sum_{r=1}^{n} (C^t)_{k,r} (A^t)_{r,j} \\
&= (C^t A^t)_{k,j}
\end{aligned}$$

for all $1 \leq k \leq p$ and $1 \leq j \leq m$, as desired. ∎

- We now show that the transpose and the dual map are essentially the same object.

Theorem 3.36. *Suppose $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(T') = (\mathcal{M}(T))^t$.*

Proof. Let $v_1, \dots, v_n$ be a basis of V, and let $\varphi_1, \dots, \varphi_n$ be the corresponding dual basis of V'. Similarly, let $w_1, \dots, w_m$ be a basis of W, and let $\psi_1, \dots, \psi_m$ be the corresponding dual basis of W'. Let $A = \mathcal{M}(T)$ and $C = \mathcal{M}(T')$. Let $1 \leq j \leq m$ and $1 \leq k \leq n$ be arbitrary. Then we have from the definition of $\mathcal{M}(T')$ that

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

from the definition of T' that

$$\begin{aligned}
(\psi \circ T)(v_k) &= \sum_{r=1}^{n} C_{r,j} \varphi_r(v_k) \\
&= C_{k,j}
\end{aligned}$$

and from the definition of $\mathcal{M}(T)$ that

$$
\begin{aligned}
(\psi \circ T)(v_k) &= \psi_j(Tv_k) \\
&= \psi_j\left(\sum_{r=1}^{m} A_{r,k} w_r\right) \\
&= \sum_{r=1}^{m} A_{r,k} \psi_j(w_r) \\
&= A_{j,k}
\end{aligned}
$$

Therefore, from the last two results, we have by transitivity that $A_{j,k} = C_{k,j}$ for all $1 \leq j \leq m$ and $1 \leq k \leq n$. It follows that $C = A^t$, i.e., that $\mathcal{M}(T') = (\mathcal{M}(T))^t$, as desired. ■

- **Row rank** (of a matrix A): The dimension of the span of the rows of A in $\mathbb{F}^{1,n}$.
- **Column rank** (of a matrix A): The dimension of the span of the columns of A in $\mathbb{F}^{m,1}$.
- The dimension of $\operatorname{range} T$ equals the column rank of $\mathcal{M}(T)$.

Theorem 3.37. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then $\dim \operatorname{range} T$ equals the column rank of $\mathcal{M}(T)$.*

Proof. Let $v_1, \dots, v_n$ be a basis of V, and let $w_1, \dots, w_m$ be a basis of W. Since $Tv = c_1 Tv_1 + \cdots + c_n Tv_n$ for all $Tv \in \operatorname{range} T$ (because $v = c_1 v_1 + \cdots + c_n Tv_n$ for some $c_1, \dots, c_n \in \mathbb{F}$ for all $v \in V$, and T is a linear map), we have that $\operatorname{range} T = \operatorname{span}(Tv_1, \dots, Tv_n)$. Additionally, since $\mathcal{M}$ is an isomorphism from $\operatorname{span}(Tv_1, \dots, Tv_n)$ to $\operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$, Theorem 3.13 asserts that $\dim \operatorname{span}(Tv_1, \dots, Tv_n) = \dim \operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))$. Therefore,

$$
\begin{aligned}
\dim \operatorname{range} T &= \dim \operatorname{span}(Tv_1, \dots, Tv_n) \\
&= \dim \operatorname{span}(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n))
\end{aligned}
$$

where the latter value is the column rank, as desired. ■

- Row rank equals column rank.

Theorem 3.38. *Suppose $A \in \mathbb{F}^{m,n}$. Then the row rank of A equals the column rank of A.*

Proof. Let $T : \mathbb{F}^{n,1} \to \mathbb{F}^{m,1}$ be defined by $Tx = Ax$. It follows that $\mathcal{M}(T) = A$. Thus,

$$
\begin{aligned}
\text{column rank } A &= \text{column rank } \mathcal{M}(T) \\
&= \dim \operatorname{range} T && \text{Theorem 3.37} \\
&= \dim \operatorname{range} T' && \text{Theorem 3.33a} \\
&= \text{column rank } \mathcal{M}(T') && \text{Theorem 3.37} \\
&= \text{column rank } A^t && \text{Theorem 3.36} \\
&= \text{row rank } A
\end{aligned}
$$

as desired. ■

- **Rank** (of A): The column rank of A.

Chapter 4

Polynomials

9/8:

- **Real part** (of $a+bi \in \mathbb{C}$): The number a. *Denoted by* $\mathbf{Re}\,\boldsymbol{z}$.
- **Imaginary part** (of $a+bi \in \mathbb{C}$): The number b. *Denoted by* $\mathbf{Im}\,\boldsymbol{z}$.
- **Complex conjugate** (of $z \in \mathbb{C}$): The number $\operatorname{Re} z - (\operatorname{Im} z)i$. *Denoted by* $\bar{\boldsymbol{z}}$.
- **Absolute value** (of $z \in \mathbb{C}$): The number $\sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}$. *Denoted by* $|\boldsymbol{z}|$.
- $z = \bar{z}$ if and only if $z \in \mathbb{R}$.
- Properties of complex numbers.

  Theorem 4.1. *Suppose $w, z \in \mathbb{C}$. Then*

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

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

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

  additivity and multiplicativity of the complex conjugate
  $\overline{w+z} = \bar{w} + \bar{z}$ *and* $\overline{wz} = \bar{w}\bar{z}$.

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

  real and imaginary parts are bounded by $|z|$
  $|\operatorname{Re} z| \leq |z|$ *and* $|\operatorname{Im} z| \leq |z|$.

  absolute value of the complex conjugate
  $|\bar{z}| = |z|$.

  multiplicativity of absolute value
  $|wz| = |w|\,|z|$.

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

- If a polynomial is the zero function, then all coefficients are 0.
  - It follows that the coefficients of a polynomial are uniquely determined.
- **Division Algorithm** (for integers): If p, s are nonnegative integers with $s \neq 0$, then there exist nonnegative integers q, r such that $r < s$ and

$$p = sq + r$$

- Analogously,

  Theorem 4.2 (Division Algorithm for Polynomials). *Suppose that $p, s \in \mathcal{P}(\mathbb{F})$, with $s \neq 0$. Then there exist unique polynomials $q, r \in \mathcal{P}(\mathbb{F})$ such that*

  $$p = sq + r$$

  and $\deg r < \deg s$.

  Proof. Let $n = \deg p$ and $m = \deg s$. We divide into two cases ($n < m$ and $n \geq m$). If $n < m$, then take $q = 0$ and $r = p$.

  On the other hand, if $n \geq P$, then let $T : \mathcal{P}_{n-m}(\mathbb{F}) \times \mathcal{P}_{m-1}(\mathbb{F}) \to \mathcal{P}_n(()\mathbb{F})$ be defined by

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

  We can easily confirm that T is a linear map.

  We now seek to prove that $\operatorname{null} T = \{(0,0)\}$. Let $(q, r) \in \operatorname{null} T$ be arbitrary. Then $sq + r = 0$. It follows that all coefficients of the polynomial $sq+r$ are zero. Consequently, $q = 0$ and $r = 0$, as desired. Therefore, $\dim \operatorname{null} T = 0$. Additionally, Theorem 3.19 implies that

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

  It follows by the Fundamental Theorem of Linear Maps that

  $$\begin{aligned} \dim \operatorname{range} T &= \dim(\mathcal{P}_{n-m}(\mathbb{F}) \times \mathcal{P}_{m-1}(\mathbb{F})) - \dim \operatorname{null} T \\ &= n + 1 \\ &= \dim \mathcal{P}_n(\mathbb{F}) \end{aligned}$$

  Thus, by Exercise 2.C.1, $\operatorname{range} T = \mathcal{P}_n(\mathbb{F})$. Therefore, since $p \in \mathcal{P}_n(\mathbb{F})$, we know that there exists $q \in \mathcal{P}_{n-m}(\mathbb{F})$ and $r \in \mathcal{P}_{m-1}(\mathbb{F})$ such that $p = T(q, r) = sq + r$.

  Additionally, we know that q, r are unique: If there exist q', r' such that $T(q', r') = p$, then $T(q-q', r-r') = p - p = 0$, implying since $\operatorname{null} T = \{(0,0)\}$ that $q - q' = 0$ and $r - r' = 0$, i.e., that $q = q'$ and $r = r'$. ■

- **Zero** (of $p \in \mathcal{P}(\mathbb{F})$): A number $\lambda \in \mathbb{F}$ such that $p(\lambda) = 0$. *Also known as* **root**.

- **Factor** (of $p \in \mathcal{P}(\mathbb{F})$): A polynomial $s \in \mathcal{P}(\mathbb{F})$ such that there exists a polynomial $q \in \mathcal{P}(\mathbb{F})$ satisfying $p = sq$.

- We now relate zeroes and factors.

  Theorem 4.3. *Suppose $p \in \mathcal{P}(\mathbb{F})$ and $\lambda \in \mathbb{F}$. Then $p(\lambda) = 0$ if and only if there is a polynomial $q \in \mathcal{P}(\mathbb{F})$ such that*

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

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

- Putting bounds on the number of zeroes a polynomial can have.

  Theorem 4.4. *Suppose $p \in \mathcal{P}(\mathbb{F})$ is a polynomial with degree $m \geq 0$. Then p has at most m distinct zeros in $\mathbb{F}$.*

- We cannot prove the following without complex analysis, but we will state it, regardless.

  Theorem 4.5 (Fundamental Theorem of Algebra). *Every nonconstant polynomial with complex coefficients has a zero.*

- The following proceeds immediately from the Fundamental Theorem of Algebra.

  Theorem 4.6. *If $p \in \mathcal{P}(\mathbb{C})$ is a nonconstant polynomial, then p has a unique factorization (except for the order of the factors) of the form*

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

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

- We now explore some of the differences between $\mathbb{R}$ and $\mathbb{C}$.

  Theorem 4.7. *Suppose $p \in \mathcal{P}(\mathbb{C})$ is a polynomial with real coefficients. If $\lambda \in \mathbb{C}$ is a zero of p, then so is $\bar{\lambda}$.*

  Theorem 4.8. *Suppose $b, c \in \mathbb{R}$. Then there is a polynomial factorization of the form*

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

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

  Theorem 4.9. *Suppose $p \in \mathcal{P}(\mathbb{R})$ is a nonconstant polynomial. Then p has a unique factorization (except for the order of factors) of the form*

  $$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M)$$

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

Chapter 5

Eigenvalues, Eigenvectors, and Invariant Subspaces

5.A Invariant Subspaces

9/8: - Let $T \in \mathcal{L}(V)$, and let V be decomposable into a direct sum of proper subspaces as follows.

$$V = U_1 \oplus \cdots \oplus U_m$$

  - To understand T, we need only understand each each **restriction** of T to a U_j.
  - Since $T|_{U_j}$ may not map U_j onto itself in every case, to use operator-based tools, we need to consider only direct sum decompositions into subspaces that T maps onto themselves, or **invariant subspace**.

- **Invariant subspace** (of V under T): A subspace U of V such that $u \in U$ implies $Tu \in U$, where $T \in \mathcal{L}(V)$.
  - In other words, U is invariant under T iff $T|_U \in \mathcal{L}(U)$.
- Some invariant subspaces under $T \in \mathcal{L}(V)$: $\{0\}$, V, null T, and range T.
- **Invariant subspace problem**: The most famous unsolved problem in functional analysis, dealing with invariant subspaces of operators on infinite-dimensional vector spaces.
- To begin our study of invariant subspaces, we consider the simplest possible type of invariant subspace: those with dimension 1.
- Every 1-dimensional subspace of V is of the form span(v) for some $v \in V$.
  - If span(v) is invariant under $T \in \mathcal{L}(V)$, then $Tv \in \text{span}(v)$.
  - If $Tv \in \text{span}(v)$, then there exists $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$.
- **Eigenvalue** (of T): A number $\lambda \in \mathbb{F}$ such that there exists a nonzero vector $v \in V$ satisfying the equation $Tv = \lambda v$. *Also known as* **characteristic value**.
- "T has a 1-dimensional invariant subspace if and only if T has an eigenvalue" (Axler, 2015, p. 134).
- We now give some conditions λ can satisfy to be deemed an eigenvalue.

  Theorem 5.1. *Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, $I \in \mathcal{L}(V)$ is the identity operator on V, and $\lambda \in \mathbb{F}$. Then the following are equivalent.*

  (a) λ is an eigenvalue of T.

(*b*) *$T - \lambda I$ is not injective.*

(*c*) *$T - \lambda I$ is not surjective.*

(*d*) *$T - \lambda I$ is not invertible.*

Proof. Suppose first that λ is an eigenvalue of T. Then

$$\begin{aligned} Tv &= \lambda v \\ Tv &= \lambda I v \\ Tv - \lambda I v &= 0 \\ (T - \lambda I)v &= 0 \end{aligned}$$

for some $v \in V$ such that $v \neq 0$. It follows that $v \in \text{null}(T - \lambda I)$, so by Theorem 3.4, $T - \lambda I$ is not injective, as desired. The proof is symmetric in the other direction. Therefore, conditions (a) and (b) are equivalent.

To prove that (a), (b), (c), and (d) are equivalent at this point, it will suffice to show that (b), (c), and (d) are equivalent. But we have this by Theorem 3.12, as desired. ∎

- **Eigenvector** (of T): A nonzero vector $v \in V$ such that there exists a $\lambda \in \mathbb{F}$ satisfying the equation $Tv = \lambda v$.
- Since $Tv = \lambda v$ iff $(T - \lambda I)v = 0$, "a vector $v \in V$ with $v \neq 0$ is an eigenvector of T corresponding to λ if and only if $v \in \text{null}(T - \lambda I)$" (Axler, 2015, p. 135).
- Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Theorem 5.2. *Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and $v_1, \dots, v_m$ are corresponding eigenvectors. Then $v_1, \dots, v_m$ is linearly independent.*

Proof. Suppose for the sake of contradiction that $v_1, \dots, v_m$ is linearly dependent. Then by the Linear Dependence Lemma, we may let k be the smallest positive integer such that $v_k \in \text{span}(v_1, \dots, v_{k-1})$. It follows that

$$v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}$$

for some $a_1, \dots, a_{k-1} \in \mathbb{F}$. Thus, applying T, we have that

$$\begin{aligned} Tv_k &= a_1 Tv_1 + \cdots + a_{k-1} Tv_{k-1} \\ \lambda_k v_k &= a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1} \end{aligned}$$

If we multiply the first equation by λ_k and subtract the above equation from it, we get that

$$0 = a_1(\lambda_k - \lambda_1)v_1 + \cdots + a_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1}$$

But since k is the smallest positive integer j such that $v_j \in \text{span}(v_1, \dots, v_{j-1})$, we know that $v_1, \dots, v_{k-1}$ are linearly independent. Thus, $a_1(\lambda_k - \lambda_1) = \cdots = a_{k-1}(\lambda_k - \lambda_{k-1}) = 0$. More specifically, since all eigenvalues are distinct (i.e., $\lambda_k - \lambda_j \neq 0$ for any $j = 1, \dots, k-1$), we must have that $a_1 = \cdots = a_{k-1} = 0$. But this implies that

$$\begin{aligned} v_k &= a_1 v_1 + \cdots + a_{k-1} v_{k-1} \\ &= 0 \end{aligned}$$

contradicting the fact that v_k, as an eigenvector, is nonzero. ∎

- We now put a bound on the number of eigenvalues.

Theorem 5.3. *Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.*

Proof. Let $T \in \mathcal{L}(V)$ have distinct eigenvalues $\lambda_1, \dots, \lambda_m$ and corresponding eigenvectors $v_1, \dots, v_m$. Then by Theorem 5.2, $v_1, \dots, v_m$ is linearly independent. It follows by Theorem 2.3 that $m \leq \dim V$ ■

- **Restriction operator** (of $T : V \to W$ to $U \subset V$): The function $T|_U : U \to W$ defined by $T|_U(u) = Tu$ for all $u \in U$. *Denoted by* $\boldsymbol{T|_U}$.
  - The fact that U is invariant under T is what allows us to consider $T|_U$ to be in $\mathcal{L}(U)$ as opposed to just $\mathcal{L}(V)$.
- **Quotient operator**: The operator $T/U \in \mathcal{L}(V/U)$ defined by $(T/U)(v + U) = Tv + U$ for all $v \in V$.
- Axler (2015) verifies that the restriction operator and the quotient operator actually *are* operators, in general, as defined.

References

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