Honors Linear Algebra and Multivariable Calculus

Math 340

R. Ebrahimian
University of Maryland, College Park

September 13, 2024

Notations

- $\in$, belongs to.
- $\forall$, for all.
- $\exists$, there exists or for some.
- D_f, the domain of function f.
- $\operatorname{Im} f$ or R_f, the image of function f.
- $\mathbb{N}$, the set of nonnegative integers.
- $\mathbb{Z}^+$, the set of positive integers.
- $\mathbb{Q}$, the set of rational numbers.
- $\mathbb{R}$, the set of real numbers.
- $A \subseteq B$, set A is a subset of set B.
- $A \subsetneqq B$, set A is a proper subset of set B.
- $A \cup B$, the union of sets A and B.
- $A \cap B$, the intersection of sets A and B.
- $\bigcup\limits_{i=1}^{n} A_i$, the union of sets $A_1, A_2, \ldots, A_n$.
- $\bigcap\limits_{i=1}^{n} A_i$, the intersection of sets $A_1, A_2, \ldots, A_n$.
- $A_1 \times A_2 \times \cdots \times A_n$, the Cartesian product of sets $A_1, A_2, \ldots, A_n$.
- $\emptyset$, the empty set.
- $f^{-1}(T)$, the inverse image (or pre-image) of set T under function f.
- $f(S)$, the image of set S under function f.
- $\operatorname{span} \mathcal{S}$, the subspace spanned by set $\mathcal{S}$.
- $\dim V$, the dimension of vector space V.
- $\langle \mathbf{v}, \mathbf{w} \rangle$, the inner product of vectors $\mathbf{v}$ and $\mathbf{w}$.
- $\mathbf{v} \cdot \mathbf{w}$, the standard inner product of vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$.
- $||\mathbf{v}||$, the norm of vector $\mathbf{v}$.
- $\det A$, the determinant of a square matrix A.

- $D_{\mathbf{u}}f(\mathbf{x}_0)$, the directional derivative of f at $\mathbf{x}_0$ with respect to the nonzero vector $\mathbf{u}$.
- $f_x, D_1 f, \dfrac{\partial f}{\partial x}$, the partial derivative of f with respect to x.
- $\mathbf{u} \times \mathbf{v}$, the cross product of $\mathbf{u}$ and $\mathbf{v}$.
- ∇f, the gradient of a scalar function f.
- curl $\mathbf{F}$, the curl of a vector field $\mathbf{F}$.
- div $\mathbf{F}$, the divergence of a vector field $\mathbf{F}$.

Contents

These note may contain occasional typos or errors. Feel free to email me at ebrahimi@umd.edu if you notice a typo or an error.

Week 1

1.1 Sets

YouTube Video: https://youtu.be/nU8U1G6BNqk

A **set** is a well-defined collection of unordered **elements**. Each set is usually defined either by listing all of its elements or by a property as below:

$$S = \{s_1, s_2, \ldots, s_n\} \qquad \text{or} \qquad S = \{s \mid s \text{ satisfies property } \mathcal{P}\}$$

Note that the order of elements in a set and repetition do not matter. So, $\{1,1,2\}$, $\{1,2\}$, $\{1,2,1,2,2\}$ and $\{2,1\}$ are all the same set.

Notation: Instead of "x is an element of the set A" or "x belongs to the set A", we write "$x \in A$".

Definition 1.1. Let A and B be two sets for which the following statement is true:

"If $x \in A$, then $x \in B$."

Then, we say A is a **subset** of B, in which case we write $A \subseteq B$.
We say a subset A of a set B is **proper** if $A \neq B$, in which case we write $A \subsetneqq B$ or $A \subsetneq B$.
The **union** of A and B, denoted by $A \cup B$, is the set consisting of all elements that are in A or B (or both).
The **intersection** of A and B, denoted by $A \cap B$, is the set consisting of all elements that are in both A and B. In other words

$$A \cup B = \{x \mid x \in A, \text{ or } x \in B\}, \qquad \text{and} \qquad A \cap B = \{x \mid x \in A, \text{ and } x \in B\}.$$

The union and intersection of n sets is defined similarly:

$$\bigcup_{i=1}^{n} A_i = \{x \mid x \in A_i, \text{ for some } i\}, \qquad \text{and} \qquad \bigcap_{i=1}^{n} A_i = \{x \mid x \in A_i, \text{ for all } i\}.$$

The union and intersection of infinitely many sets $A_1, A_2, \ldots$ are defined similarly and they are denoted by $\bigcup\limits_{n=1}^{\infty} A_n$ and $\bigcap\limits_{n=1}^{\infty} A_n$.

The **empty set** or the **null set** is the set with no elements. It is denoted by $\emptyset$ or $\{\}$.

Remark. The word "or" in mathematics is not exclusive. In other words, for two statements p and q, the statement "p or q" means "p or q or both". For example, "$x \in A$ or $x \in B$" means, "x is an element of A or x is an element of B or x is an element of both A and B."

Remark. Sometimes sets are labeled by elements of another set. For example, instead of $\bigcup_{n=0}^{\infty} A_n$ we may write $\bigcup_{n\in\mathbb{N}} A_n$ and instead of $\bigcup_{n=-\infty}^{\infty} A_n$ we may write $\bigcup_{n\in\mathbb{Z}} A_n$. This is especially useful when there are too many sets to label them using only integers. For example, in the union $\bigcup_{r\in\mathbb{R}} A_r$, there is a set A_r corresponding to every real number r.

Definition 1.2. We say two sets A and B are **equal** if and only if $A \subseteq B$ and $B \subseteq A$, in which case we write $A = B$.

Example 1.1. Prove that for every three sets A, B, and C we have $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

Definition 1.3. An **ordered pair** (a, b) is two objects a and b with a specified order. Two ordered pairs (a, b) and (c, d) are the same if and only if $a = c$ and $b = d$. An n**-tuple** $(a_1, a_2, \ldots, a_n)$ is n objects $a_1, a_2, \ldots, a_n$ with a specified order. Two n-tuples $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ are equal if and only if $a_i = b_i$ for $i = 1, \ldots, n$.

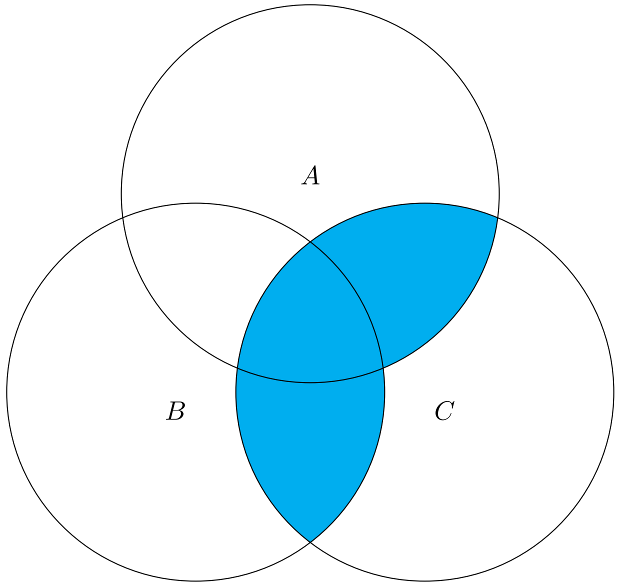
Definition 1.4. The **Cartesian product** of n sets $A_1, A_2, \ldots, A_n$, denoted by $A_1 \times A_2 \times \cdots \times A_n$, is the set of all n-tuples $(a_1, a_2, \ldots, a_n)$ for which $a_i \in A_i$ for all i. The Cartesian product of n copies of a set A is denoted by A^n.

Example 1.2. Every point on the plane can be represented by an element of the set $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$. Every point on the n-dimensional space can be represented by an element of the set

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in \mathbb{R}\}.$$

Definition 1.5. We say two sets A and B are **disjoint** if $A \cap B = \emptyset$. We say sets $A_1, A_2, \ldots, A_n$ are **disjoint** if $\bigcap_{i=1}^{n} A_i = \emptyset$. We say $A_1, A_2, \ldots, A_n$ are **pairwise disjoint** if for every $i \neq j$, A_i and A_j are disjoint. Similarly, these notions are defined for an infinite collection of sets A_i with $i \in I$.

To understand sets we often picture them as ovals or circles. The following shows the **Venn diagram** of $(A \cup B) \cap C$.



Definition 1.6. For two sets A, B, the **difference** $A - B$ consists of all elements of A that are not in B.

$$A - B = \{x \in A \mid x \notin B\}.$$

When dealing with sets, we often assume all of our sets are subsets of a given larger set U. This set is called a **universal set**. For example in number theory the universal set is usually $\mathbb{Z}$, the set of integers. In calculus we deal with real numbers and thus our universal set is typically $\mathbb{R}$.

Assume A is a subset of the universal set U. The **complement** of A in U is the set consisting of all elements of U that are not in A. The complement of A is denoted by A^c.

Theorem 1.1 (De Morgan's Laws). *Given n subsets $A_1, \ldots, A_n$ of a universal set U we have:*

(a) $\left(\bigcap_{j=1}^{n} A_j\right)^c = \bigcup_{j=1}^{n} A_j^c.$

(b) $\left(\bigcup_{j=1}^{n} A_j\right)^c = \bigcap_{j=1}^{n} A_j^c.$

Remark. Similar to above, given a nonempty set of indices I and a collection of subsets A_i of a universal set U, for every $i \in I$, we have:

(a) $\left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c.$

(b) $\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c.$

1.2 Functions

YouTube Video: https://youtu.be/lwQG_dMXQOA

Definition 1.7. Given two nonempty sets A and B, a **function** or a **mapping** $f : A \to B$ is a rule that assigns to every element $a \in A$ an element $f(a) \in B$. The set A is called the **domain** of f and is denoted by D_f. The set B is called the **co-domain** of f. The **range** or **image** of f, denoted by R_f or $\text{Im } f$, is the set $\text{Im } f = \{f(a) \mid a \in A\}$.

Two functions f and g are called **equal** if they have the same domain, the same co-domain, and $f(x) = g(x)$ for all x in their common domain.

f is called **surjective** or **onto** if for every $b \in B$ there is $a \in A$ for which $f(a) = b$.

f is called **injective** or **one-to-one** if whenever $f(a_1) = f(a_2)$ we also have $a_1 = a_2$.

f is called **bijective** if it is injective and surjective.

The **composition** $f \circ g$ of two functions f, g with $R_g \subseteq D_f$, is a function from D_g to the co-domain of f given by $f \circ g(x) = f(g(x))$, for all $x \in D_g$.

The function $\text{id}_A : A \to A$ defined by $\text{id}_A(a) = a$, for all $a \in A$ is called the **identity** function of A.

A function $f : A \to B$ is called **invertible** if and only if there is a function $g : B \to A$ for which $f \circ g = \text{id}_B$ and $g \circ f = \text{id}_A$. The function g is called the **inverse** of f and is denoted by f^{-1}.

Example 1.3 (Projection)**.** The function $\pi_1 : A \times B \to A$ defined by $\pi_1(a, b) = a$ is called the projection onto the first component. Similarly, the function $\pi_i : A_1 \times \cdots \times A_n \to A_i$ defined by $\pi_i(a_1, \ldots, a_n) = a_i$ is called the **projection** onto the i-th component.

Definition 1.8. Given a function $f : A \to B$, and a subset S of A, the **image** of S under f is the set $f(S) = \{f(s) \mid s \in S\}$. If T is a subset of B, then the **pre-image** or **inverse image** of T under f is the set $f^{-1}(T) = \{a \in A \mid f(a) \in T\}$.

Note that the pre-image and image are both sets.

Example 1.4. Let $f : A \times B \to B$ be the projection onto the second component. For every $b \in B$ find the pre-image of $\{b\}$ under f.

Example 1.5. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by $f(x, y) = 2x + 3y$. For every real number b, evaluate and describe $f^{-1}(\{b\})$. How do these pre-images change as b changes?

Theorem 1.2 (Properties of Pre-image)**.** *Suppose $f : A \to B$ is a function, $S \subseteq A$, $T \subseteq B$, and $T_i \subseteq B$ for every $i \in I$, where I is a nonempty set of indices. Then*

(a) $S \subseteq f^{-1}(f(S))$, *and* $f(f^{-1}(T)) \subseteq T$.

(b) $f^{-1}(\bigcup_{i \in I} T_i) = \bigcup_{i \in I} f^{-1}(T_i)$.

(c) $f^{-1}(\bigcap_{i \in I} T_i) = \bigcap_{i \in I} f^{-1}(T_i)$.

Proof. (a) Suppose $s \in S$. By definition of image, $f(s) \in f(S)$. Therefore, by definition of pre-image $s \in f^{-1}(f(S))$. This completes the proof of $S \subseteq f^{-1}(f(S))$.

Suppose $x \in f(f^{-1}(T_i))$. By definition $x = f(y)$, for some $y \in f^{-1}(T_i)$. Therefore, $f(y) \in T_i$, which means $x \in T_i$. This means $f(f^{-1}(T_i)) \subseteq T_i$.

Parts (b) and (c) are left as exercises. $\square$

1.3 Proofs

YouTube Video: https://youtu.be/sW823lmew64

In writing proofs you should note the following:

- You cannot prove a *universal statement* (statements involving *for every* or *for all*) by examples. For example if you are asked to prove "The sum of every two odd integers is even." your proof may not be "3 is odd, 5 is odd, 3+5=8 is even. Therefore, the sum of every two odd integers is even."

  On the other hand, for *existential statements* (when a statement is asking you to show something exists), giving an example and showing that the example satisfies all the required conditions is enough.

- Do not use the same variable for two different things.

- You may not assume anything but what is given in the assumptions.

- All steps must be justified and the justifications must all be clearly stated.

- You may only use known facts. These are typically things that have been previously proved as theorems or are facts stated in definitions.

- To prove a statement of the form "p if and only if q" we will need to prove both "If p, then q" and "If q, then p".

To prove a *statement* (usually of the form "If p then q"), there are three main methods of proof. We will look at each one via examples.

1.3.1 Direct Proof

In this method we start from the assumption and by taking logical steps we end up with the conclusion.

Example 1.6. Prove that the function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=x^3+2x$ is one-to-one.

Solution. By definition of one-to-one, we need to prove if $f(x)=f(y)$, then $x=y$.

Suppose $f(x)=f(y)$. Then $x^3+2x=y^3+2y$. Therefore, $x^3-y^3+2(x-y)=0$, which implies $(x-y)(x^2+xy+y^2+2)=0$. This means either $x=y$ or $x^2+xy+y^2+2=0$. If the second equality holds, by the quadratic formula we obtain $x=\dfrac{-y\pm\sqrt{y^2-4(y^2+2)}}{2}$. The discriminant is $-3y^2-8$ which is negative. Therefore, this equality is impossible, and hence $x=y$. This means f is one-to-one. $\square$

1.3.2 Proof by Contradiction

In this method, we assume the conclusion is false while the assumption is true. After taking logical steps we obtain a contradiction. A contradiction is a statement that is false: Either it violates a fact in math such as a theorem or a definition or it violates the assumptions. When using proof by contradiction make sure to clearly specify you are using this method. You could say " On the contrary assume..." or "By way of contradiction assume..." or simply state "We will use proof by contradiction."

Example 1.7. Prove that there are infinitely many primes.

Solution. On the contrary assume there are only a finite number of primes, and let $p_1, p_2, \ldots, p_n$ be the list of all primes. Since the integer $d = p_1 \cdots p_n + 1$ is more than one, d has a prime factor. Since $p_1, p_2, \ldots, p_n$ is the list of all primes, one of the p_i's must divide d. On the other hand p_i divides $p_1 p_2 \cdots p_n$. Therefore, p_i must divide $d - p_1 p_2 \cdots p_n = 1$. This is a contradiction. Therefore, the initial assumption must be false, and thus there must exist infinitely many primes. $\square$

1.3.3 Proof by Induction

To prove a statement $P(n)$ (i.e. a statement that depends on a positive integer n) we will:

- Prove $P(1)$ (basis step); and
- Assume $P(n)$ holds for some $n \geq 1$, and then prove $P(n+1)$ (inductive step).

If you need to use $P(n-1)$ in your proof of $P(n+1)$, then the basis step must involve two consecutive integers, e.g. $P(1)$ and $P(2)$.

Often times we use what is called **strong induction** which involves assuming $P(1), \ldots, P(n)$ and then proving $P(n+1)$ in addition to proving the basis step.

When employing the method of mathematical induction keep in mind to always start your proof by "We will prove *the statement* by induction on *the variable*". Replace "the statement" and "the variable" accordingly. Also, clearly separate the basis step and the inductive step.

Example 1.8. Prove that the sum of the first n positive odd integers is n^2.

1.4 $\mathbb{R}^n$ as a Vector Space

YouTube Video: https://youtu.be/Bwpk4fPJmoU

As we saw earlier, elements of $\mathbb{R}^n$ are n-tuples of the form $(x_1, x_2, \ldots, x_n)$, where x_j's are real numbers. Each one of these elements is called a **vector** and these vectors can be added componentwise as follows:

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Each vector can also be scaled by any real number c (also called a **scalar**) as follows:

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

This vector addition and scalar multiplication satisfy the following properties.

(I) (Closure) For every two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and every scalar $c \in \mathbb{R}$, both $\mathbf{x} + \mathbf{y}$ and $c\mathbf{x}$ are in $\mathbb{R}^n$.

(II) (Associativity) For every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, and every $a, b \in \mathbb{R}$, we have $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$, and $a(b\mathbf{x}) = (ab)\mathbf{x}$.

(III) (Commutativity) For every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

(IV) (Additive Identity) For every $\mathbf{x} \in \mathbb{R}^n$ we have $\mathbf{x} + (0, 0, \ldots, 0) = \mathbf{x}$. (The vector $(0, 0, \ldots, 0)$ is called the **zero vector** and is denoted by $\mathbf{0}$.)

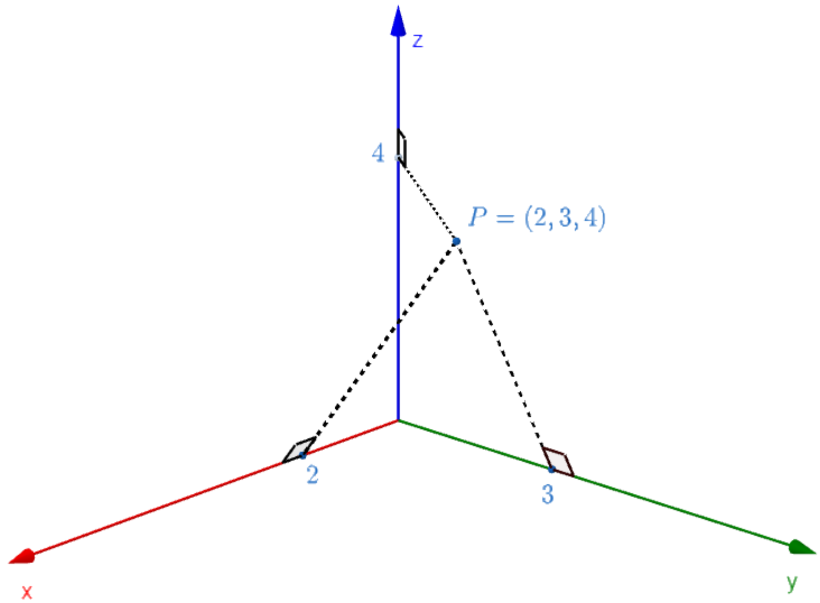
(V) (Additive Inverse) For every $\mathbf{x} \in \mathbb{R}^n$, there is a vector $\mathbf{y} \in \mathbb{R}^n$ for which $\mathbf{x} + \mathbf{y} = \mathbf{0}$. (This vector $\mathbf{y}$ is called the **additive inverse** of $\mathbf{x}$ and is denoted by $-\mathbf{x}$. It is given by $-(x_1, x_2, \ldots, x_n) = (-x_1, -x_2, \ldots, -x_n)$.)

(VI) (Distributivity) For every $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $(a+b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$, and $a(\mathbf{x}+\mathbf{y}) = a\mathbf{x} + a\mathbf{y}$.

(VII) (Multiplicative Identity) For every $\mathbf{x} \in \mathbb{R}^n$ we have $1\mathbf{x} = \mathbf{x}$.

The seven properties I-VII listed above are called **vector space** properties of $\mathbb{R}^n$. This is often phrased as "$\mathbb{R}^n$ is a vector space." Note that sometimes we refer to elements of $\mathbb{R}^n$ as **points**. This is only for conceptualizing these objects. The math does not change. When elements of $\mathbb{R}^n$ are seen as points, the zero vector is referred to as the **origin**.

Geometrically, elements of $\mathbb{R}^2$ can be represented by points on a plane. Elements of $\mathbb{R}^3$ can be represented by points in a 3D space. To do that, we need three axes, x-, y-, and z-**axes**. These three axes must satisfy the right-hand rule. The coordinates of each point can be found by dropping perpendiculars to the axes.



The set of all points with positive coordinates, is called the **first octant**.

There are three planes each containing two of the x-, y-, and z- axes. Each of the three xy-, xz- and yz-planes is called a **coordinate plane**.

Theorem 1.3. *The distance between two points* (x_1, y_1, z_1) *and* (x_2, y_2, z_2) *in* $\mathbb{R}^3$ *is given by*

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

1.5 Warm-ups

Example 1.9. How many elements does the set $\{2, 1, 3, 2\}$ have? How about the set $\{3, 2, 1\}$? How are these two sets related?

Solution. Since repetition and order does not matter in a set, these two sets are the same sets:

$$\{2, 1, 3, 2\} = \{3, 2, 1\}.$$

So, these sets both have three elements. □

Example 1.10. Let E be the set of all even integers and O be the set of all odd integers. Describe $E \cup O$ and $E \cap O$.

Solution. $E \cup O$ is the set of all integers that are odd or even. Since every integer is either odd or even, $E \cup O = \mathbb{Z}$.

By definition of intersection, $E \cap O$ is the set of all integers that are both even and odd. Since no integer is both even and odd, $E \cap O = \emptyset$. □

Example 1.11. Consider the function $f : \{1,2,3\} \to \{1,2,3,4\}$ defined by $f(1) = 2, f(2) = 2$, and $f(3) = 4$. Find the domain of f, the co-domain of f, the image of f, $f(\{1,2\})$, and $f^{-1}(\{2,3\})$.

Solution. The domain of f is $\{1,2,3\}$. The co-domain of f is $\{1,2,3,4\}$. The image of f is $\{f(1), f(2), f(3)\} = \{2,4\}$.

$$f(\{1,2\}) = \{f(1), f(2)\} = \{2,2\} = \{2\},$$

and

$$f^{-1}(\{2,3\}) = \{x \in \{1,2,3\} \mid f(x) \in \{2,3\}\}.$$

Thus, $f^{-1}(\{2,3\}) = \{1,2\}$. □

1.6 More Examples

Example 1.12. Given sets $A = \{1,2\}, B = \{0,1,-1\}$, write each of the following sets by listing all of its elements:

(a) $A \cup B$

(b) $A \cap B$

(c) $A \times B$

Solution. (a) $A \cup B$ consists of all elements that are in A or B. Thus, $A \cup B = \{1,2,0,-1\}$.

(b) $A \cap B$ consists of all elements that are in both A and B. Thus, $A \cap B = \{1\}$.

(c) $A \times B$ consists of all elements of the form (a,b), where $a \in A$ and $b \in B$. Thus,

$$A \times B = \{(1,0), (1,1), (1,-1), (2,0), (2,1), (2,-1)\}$$

□

Example 1.13. Prove that for all sets $A, B_1, B_2, \ldots$, we have $A \bigcap (\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} (A \cap B_n)$.

Solution. Suppose $x \in A \bigcap (\bigcup_{n=1}^{\infty} B_n)$. By definition of intersection, $x \in A$ and $x \in \bigcup_{n=1}^{\infty} B_n$. By definition of union, $x \in B_n$ for some n. This means $x \in A \cap B_n$ and thus $x \in \bigcup_{n=1}^{\infty} (A \cap B_n)$. Therefore,

$$A \bigcap (\bigcup_{n=1}^{\infty} B_n) \subseteq \bigcup_{n=1}^{\infty} (A \cap B_n). \qquad (*)$$

Suppose $x \in \bigcup_{n=1}^{\infty}(A \cap B_n)$. By definition of union, $x \in A \cap B_n$ for some n. Thus, by definition of intersection, $x \in A$ and $x \in B_n$. Therefore, $x \in A$ and $x \in \bigcup_{n=1}^{\infty} B_n$. This implies $x \in A \bigcap (\bigcup_{n=1}^{\infty} B_n)$. Therefore,

$$\bigcup_{n=1}^{\infty}(A \cap B_n) \subseteq A \bigcap (\bigcup_{n=1}^{\infty} B_n). \qquad (**)$$

Combining $(*)$ and $(**)$ we obtain the result. $\square$

Example 1.14. Describe each set as a subset of $\mathbb{R}^2$.

(a) $[0,1] \times \{1\}$.

(b) $[1,2] \times [0,1]$.

Solution. (a) This is the set of all (x, y), where $x \in [0,1]$ and $y = 1$. This is a horizontal segment connecting $(0,1)$ and $(1,1)$.

(b) This set consists of all points (x, y) for which $x \in [1,2]$ and $y \in [0,1]$. This is a filled square with vertices $(1,0), (2,0), (1,1)$, and $(2,1)$. $\square$

Example 1.15. Let C be the unit circle $x^2 + y^2 = 1$ in the xy-plane. Geometrically describe the set $C \times \mathbb{R}$.

Solution. $C \times \mathbb{R}$ is the set of all (x, y, z) for which $x^2 + y^2 = 1$. This means $C \times \mathbb{R}$ is the union of the translation of the unit circle C in the direction of the z-axis. This is a right circular cylinder. $\square$

Example 1.16. Suppose X and Y are finite nonempty sets of sizes m and n respectively. Let Y^X be the set of all function $f : X \to Y$. What is the size of Y^X? (This should tell you why we use the notation "Y^X".)

Solution. Let $f : X \to Y$ be a function. For each $x \in X$, $f(x)$ could be any element of Y. Thus, there are n possible values for $f(x)$. Since this is true for each element of X, there are n^m functions $f : X \to Y$. $\square$

Example 1.17. Define the Fibonacci sequence F_n by $F_0 = 0, F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$ for all $n \geq 0$. Prove that $F_n < 2^n$ for all $n \geq 0$.

Sketch. The fact that each term of the sequence depends on the previous terms reminds us of the method of Mathematical Induction. So, we will employ this method. However since each term depends on the previous two terms, we will have to start with proving the given statement for two values of n.

Solution. We will prove $F_n < 2^n$ by induction on n.

Basis step: $F_0 = 0 < 2^0 = 1$, and $F_1 = 1 < 2^1$.

Inductive step: Suppose for some $n \geq 1$, $F_k < 2^k$ for $k = 0, \ldots, n$. By assumption $F_{n+1} = F_n + F_{n-1} < 2^n + 2^{n-1} = 2^{n-1}(2+1) < 2^{n+1}$, as desired. This completes the solution. □

Example 1.18. Prove that if a real number x satisfies $|x| + x > 0$, then x is positive.

Solution. On the contrary assume x is not positive. Therefore, we have two cases:

Case I: $x = 0$. This means $|x| + x = 0$, which is a contradiction.

Case II: $x < 0$. This implies $|x| = -x$ and thus, $|x| + x = 0$, which is a contradiction.
Therefore x must be positive. □

Definition 1.9. Suppose D is a subset of $\mathbb{R}$. A function $f : D \to \mathbb{R}$ is said to be **periodic** if there is a positive real number T for which:

- For every $x \in \mathbb{R}$, the number x is in D if and only if $x + T \in D$, and
- $f(x + T) = f(x)$ for every $x \in D$.

Example 1.19. Prove $\sin x$ and $\sin(\pi x)$ are both periodic, but their sum is not.

Solution. First note that the domain of both functions is $\mathbb{R}$. Thus, $x \in \mathbb{R}$ iff $x + T \in \mathbb{R}$ is valid for every $x, T \in \mathbb{R}$.

By properties of $\sin x$ we know $\sin(x + 2\pi) = \sin x$ and $\sin(\pi(x + 2)) = \sin(\pi x)$. Thus, both functions $\sin x$ and $\sin(\pi x)$ are periodic.

To prove $\sin x + \sin(\pi x)$ is not periodic we will use proof by contradiction. Assume on the contrary there is a positive real number T for which for every $x \in \mathbb{R}$ we have

$$\sin(x + T) + \sin(\pi(x + T)) = \sin x + \sin(\pi x) \qquad (*)$$

Differentiating $(*)$ with respect to x twice, we obtain

$$-\sin(x + T) - \pi^2 \sin(\pi(x + T)) = -\sin x - \pi^2 \sin(\pi x) \qquad (**)$$

Adding $(*)$ and $(**)$ and dividing both sides by $1 - \pi^2$ we conclude $\sin(\pi(x+T)) = \sin(\pi x)$ for every $x \in \mathbb{R}$. Combining this with $(*)$ we conclude that $\sin(x + T) = \sin x$ for all $x \in \mathbb{R}$. This implies $T = 2k\pi$ for some positive integer k. Substituting x by x/π in $\sin(\pi(x + T)) = \sin(\pi x)$ we conclude $\sin(x + \pi T) = \sin(x)$. Therefore, $\pi T = 2n\pi$ for some positive integer n. Therefore, $\pi(2k\pi) = 2n\pi$. Hence $\pi = n/k$ is rational, which is a contradiction. □

Example 1.20. Prove that for every positive integer n, there is a polynomial $p_n(x)$ for which the n-th derivative of e^{x^2} at x is equal to $p_n(x)e^{x^2}$.

Solution. We will prove this by induction on n.
Basis step. By the Chain Rule, the first derivative of e^{x^2} is $2xe^{x^2}$. The fact that $p_1(x) = 2x$ is a polynomial proves the claim for $n = 1$.

Inductive step. Suppose the n-th derivative of e^{x^2} is $p_n(x)e^{x^2}$ for a polynomial p_n. Differentiating this using the Product Rule and the Chain Rule we conclude that the $(n + 1)$-th derivative of e^{x^2} is equal to $p_n'(x)e^{x^2} + p_n(x)2xe^{x^2} = (p_n'(x) + 2xp_n(x))e^{x^2}$. Since the derivative of a polynomial is a polynomial and the product and sum of polynomials are polynomials, $p_n'(x) + 2xp_n(x)$ is a polynomial. So, setting $p_{n+1}(x) = p_n'(x) + 2xp_n(x)$ we conclude the $(n + 1)$-th derivative of e^{x^2} is equal to $p_{n+1}(x)e^{x^2}$ for some polynomial p_{n+1}, as desired. $\square$

Example 1.21. Let $f : A \to B$ be a function, $S \subseteq A$, and $T \subseteq B$. Prove that:

(a) If f is one-to-one, then $S = f^{-1}(f(S))$.

(b) If f is onto, then $T = f(f^{-1}(T))$.

Solution. (a) By Theorem 1.2, $S \subseteq f^{-1}(f(S))$. It is enough to show $f^{-1}(f(S)) \subseteq S$. Suppose $x \in f^{-1}(f(S))$. By definition of pre-image, $f(x) \in f(S)$. By definition of $f(S)$ we conclude $f(x) = f(s)$ for some $s \in S$. Since f is one-to-one, $x = s$ and thus $x \in S$. This shows $f^{-1}(f(S)) \subseteq S$, as desired.

(b) By Theorem 1.2, $f(f^{-1}(T)) \subseteq T$. Thus, it is enough to prove $T \subseteq f(f^{-1}(T))$. Let $x \in T$. Since f is onto, there is $a \in A$ such that $f(a) = x$. Thus, by definition of pre-image $a \in f^{-1}(T)$. Therefore, by definition of image $f(a) \in f(f^{-1}(T))$. Since $f(a) = x$, we obtain $x \in f(f^{-1}(T))$. Therefore, $T \subseteq f(f^{-1}(T))$, as desired. $\square$

Example 1.22. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = x^2$. Find each of the following:

(a) $f([0, 1))$.

(b) $f^{-1}([-1, 0))$.

(c) $f^{-1}((0, 2))$.

Solution. (a) Note that if $x \in [0, 1)$, then $0 \le x^2 < 1$ and thus $f([0, 1)) \subseteq [0, 1)$. Furthermore, if $y \in [0, 1)$, then $\sqrt{y} \in [0, 1)$ and $f(\sqrt{y}) = y$. Therefore, $[0, 1) \subseteq f([0, 1))$. This shows $f([0, 1)) = [0, 1)$.

(b) By definition of pre-image, $x \in f^{-1}([-1, 0))$ if and only if $f(x) \in [-1, 0)$ if and only if $-1 \le x^2 < 0$, which is impossible. Therefore, $f^{-1}([-1, 0)) = \emptyset$.

(c) By definition of pre-image, $x \in f^{-1}(0,2)$ if and only if $f(x) \in (0,2)$, i.e. $0 < x^2 < 2$. This holds if and only if $0 < x < \sqrt{2}$ or $-\sqrt{2} < x < 0$. Therefore, $f^{-1}((0,2)) = (0,\sqrt{2}) \cup (-\sqrt{2},0)$. □

Example 1.23. Let $f : A \to B$ be a function. Find each of the following:

(a) $f(\emptyset)$.

(b) $f^{-1}(\emptyset)$.

(c) $f^{-1}(B)$.

Solution. (a) $f(\emptyset)$ consists of all elements of the form $f(x)$, where $x \in \emptyset$, but since $\emptyset$ contains no elements, $f(\emptyset) = \emptyset$.

(b) $f^{-1}(\emptyset)$ consists of all elements $a \in A$ for which $f(a) \in \emptyset$. Since $\emptyset$ contains no elements $f^{-1}(\emptyset) = \emptyset$.

(c) $f^{-1}(B)$ consists of all elements $a \in A$ for which $f(a) \in B$, but since B is the co-domain, $f(a)$ is always in B, and thus $f^{-1}(B) = A$. □

Example 1.24. Let $f : A \to B$ be a function, and S_i with $i \in I$ be a collection of subsets of A. Prove that

(a) $f\left(\bigcup\limits_{i\in I} S_i\right) = \bigcup\limits_{i\in I} f(S_i)$.

(b) $f\left(\bigcap\limits_{i\in I} S_i\right) \subseteq \bigcap\limits_{i\in I} f(S_i)$. By an example show that the equality does not always hold.

Solution. (a) Let $x \in f\left(\bigcup\limits_{i\in I} S_i\right)$. By definition of image, $x = f(s)$ for some $s \in \bigcup\limits_{i\in I} S_i$. By definition of union, $s \in S_j$ for some $j \in I$. Therefore, $x = f(s) \in f(S_j)$, which implies $x \in \bigcup\limits_{i\in I} f(S_i)$, by definition of union. The other inclusion is similar and is left as an exercise.

(b) Let $x \in f\left(\bigcap\limits_{i\in I} S_i\right)$. By definition of image, $x = f(s)$ for some $s \in \bigcap\limits_{i=1}^{n} S_i$. By definition of intersection, $s \in S_i$ for all $i \in I$, and thus $x = f(s) \in f(S_i)$ for all $i \in I$, by definition of image. Therefore, $x \in \bigcap\limits_{i\in I} f(S_i)$, by definition of intersection. This completes the proof.

Consider $f : \{1,2\} \to \{1\}$ given by $f(1) = f(2) = 1$. Let $S_1 = \{1\}$ and $S_2 = \{2\}$. Then, $S_1 \cap S_2 = \emptyset$ and thus $f(S_1 \cap S_2) = \emptyset$. On the other hand $f(S_1) = f(S_2) = \{1\}$ and thus $f(S_1) \cap f(S_2) \neq \emptyset$. □

Example 1.25. Determine if each function below is one-to-one, onto, both or neither.

(a) $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x,y) = (x+2y, x-y)$.

(b) $f : \mathbb{Z}^2 \to \mathbb{Q}^+$ given by $f(m,n) = 2^m \cdot 3^n$.

Solution. (a) Suppose $f(x,y) = f(a,b)$. This implies

$$\begin{cases} x + 2y = a + 2b \\ x - y = a - b \end{cases}$$

Subtracting the two equations above we obtain $3y = 3b$ and thus $y = b$. Substituting into the first equation we obtain $x = a$. Therefore, f is one-to-one.

Given $(a,b) \in \mathbb{R}^2$, we will need to determine if there is $(x,y) \in \mathbb{R}^2$ for which $f(x,y) = (a,b)$. Solving the system

$$\begin{cases} x + 2y = a \\ x - y = b \end{cases}$$

we obtain $x = (a+2b)/3$ and $y = (a-b)/3$. Therefore, this function is also onto.

(b) Suppose $f(m,n) = f(r,s)$ for some integers m,n,r,s. Therefore, $2^m \cdot 3^n = 2^r \cdot 3^s$. Without loss of generality assume $m \geq r$. We see that $2^{m-r} = 3^{s-n}$. If the exponent $m-r$ is positive, then the left side is even, while the right side is not. This contradiction shows $m = r$ and thus $n = s$. Therefore, f is one-to-one.

This function is not onto. For example $f(m,n) = 5$ has no solutions, because $2^m \cdot 3^n = 5$ is impossible by the uniqueness of prime factorization. □

Example 1.26. For a function $f : A \to B$ prove that the equality $f(S_1 \cap S_2) = f(S_1) \cap f(S_2)$ holds for all subsets S_1, S_2 of A if and only if f is one-to-one.

Solution. First, note that by Example 1.24, we know

$$f(S_1 \cap S_2) \subseteq f(S_1) \cap f(S_2).$$

($\Leftarrow$) Suppose f is one-to-one. Let $x \in f(S_1) \cap f(S_2)$. By definition, $x = f(s_1) = f(s_2)$ for some $s_1 \in S_1$, and some $s_2 \in S_2$. Since f is one-to-one, we have $s_1 = s_2$. Therefore, $s_1 \in S_1 \cap S_2$. This means $x \in f(S_1 \cap S_2)$. This shows $f(S_1) \cap f(S_2) = f(S_1 \cap S_2)$.

($\Rightarrow$) Assume $f(S_1 \cap S_2) = f(S_1) \cap f(S_2)$ for every two subsets S_1, S_2 of A. Suppose $f(a) = f(b)$, and let $S_1 = \{a\}, S_2 = \{b\}$. We know $f(S_1) = \{f(a)\}$ and $f(S_2) = \{f(b)\} = \{f(a)\}$. Therefore, $f(S_1) \cap f(S_2) = \{f(a)\}$. If $a \neq b$, then $S_1 \cap S_2 = \emptyset$, which means $f(S_1 \cap S_2) = \emptyset \neq \{f(a)\}$. Therefore, $a = b$. This shows f is one-to-one. □

Example 1.27. Suppose c is a real number and $\mathbf{v}$ is a vector in $\mathbb{R}^n$. Prove that if $c\mathbf{v} = \mathbf{0}$, then $c = 0$ or $\mathbf{v} = \mathbf{0}$.

Solution. Let $\mathbf{v} = (x_1, \ldots, x_n)$. On the contrary, assume neither c is zero, nor $\mathbf{v}$ is the zero vector. Therefore, x_i is not zero for some i. Since $c \neq 0$, we have $cx_i \neq 0$. Thus,

$$c\mathbf{v} = (cx_1, \ldots, cx_i, \ldots, cx_n) \neq \mathbf{0}.$$

This contradiction shows $c = 0$ or $\mathbf{v} = \mathbf{0}$. □

Further Reading: Click here for further reading on Sets, Maps, and Vector Spaces.

1.7 Exercises

Exercise 1.1. *Given the following sets A, B, C write down each of the sets $(A \times B) \cap C$, $A \cup (B \cap C)$ and $A \times B \times C$ by listing all of their elements in braces.*

$$A = \{1, -1\}, B = \{1, 0\}, C = \{(1, 1), (1, 0)\}.$$

Exercise 1.2. *For n sets $A_1, A_2, \ldots, A_n$, prove that $A_1 \times A_2 \times \cdots \times A_n = \emptyset$ if and only if $A_i = \emptyset$ for some i.*

Hint: Proof by contradiction might be useful.

Exercise 1.3. *Suppose for two nonempty sets A, B we know $A \times B = B \times A$. Prove that $A = B$.*

Exercise 1.4. *Prove or disprove:*

(a) For every three sets A, B, C we have $A - (B \cup C) = (A - B) \cap (A - C)$.

(b) For every three sets A, B, C we have $A - (B \cap C) = (A - B) \cup (A - C)$.

Exercise 1.5. *Determine (with full justification) which of the following statements are true.*

(a) $(\mathbb{Z} \times \mathbb{R}) \cup (\mathbb{R} \times \mathbb{Z}) = \mathbb{R} \times \mathbb{R}$.

(b) $(\mathbb{Z} \times \mathbb{R}) \cap (\mathbb{R} \times \mathbb{Z}) = \mathbb{Z} \times \mathbb{Z}$.

(c) $(\mathbb{R} - \mathbb{Z}) \times \mathbb{Z} = (\mathbb{R} \times \mathbb{Z}) - (\mathbb{Z} \times \mathbb{Z})$.

(d) $\{x \in \mathbb{R} \mid \sin x \in \mathbb{Z}\} = \{x \in \mathbb{R} \mid \cos x \in \mathbb{Z}\}$.

(e) $\{n \in \mathbb{Z} \mid n = 6m + 1 \text{ for some } m \in \mathbb{Z}\} = \{n \in \mathbb{Z} \mid n = 6m - 5 \text{ for some } m \in \mathbb{Z}\}$.

Definition 1.10. The **power set** of a set A, denoted by $\mathcal{P}(A)$ is the set consisting of all subsets of A.

Exercise 1.6. *Prove or disprove each of the following:*

(a) For every two sets A, B, we have $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$.

(b) For every two sets A, B, we have $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$.

(c) For every two sets A, B, we have $\mathcal{P}(A \times B) = \mathcal{P}(A) \times \mathcal{P}(B)$.

Exercise 1.7. *Suppose $J \subseteq I$ are two nonempty sets of indices, and each A_i with $i \in I$ is a set. Prove each of the following:*

(a) $\bigcup_{j \in J} A_i \subseteq \bigcup_{i \in I} A_i$.

(b) $\bigcap_{i \in I} A_i \subseteq \bigcap_{j \in J} A_i$.

Exercise 1.8. *Prove that*

$$\bigcup_{x \in [0,1]} ([x, 1] \times [0, x^2]) = \{(x, y) \mid 0 \leq x \leq 1 \text{ and } 0 \leq y \leq x^2\}$$

Exercise 1.9. *Prove that*

$$\bigcap_{x \in [0,1]} ([x, 1] \times [0, x^2]) = \{(1, 0)\}$$

Exercise 1.10. *Given a nonempty set A, what is the set $\bigcup_{a \in A} \{a\}$? How about $\bigcup_{a \in A} \{(a, 1)\}$? Prove your claims.*

Exercise 1.11. *Let X be a nonempty set with n elements. How many one-to-one functions $f : X \to X$ are there? How many onto functions $f : X \to X$ are there?*

Exercise 1.12. *The graph of a function $f : X \to Y$ is defined by $\Gamma_f = \{(x, f(x)) \mid x \in X\}$. Prove that two functions $f, g : X \to Y$ are equal if and only if $\Gamma_f = \Gamma_g$.*

Exercise 1.13. *Suppose f, g are two functions for which $R_g \subseteq D_f$. Prove or disprove each statement.*

(a) If both f and g are injective, then so is $f \circ g$.

(b) If both f and g are surjective, then so is $f \circ g$.

Exercise 1.14. *Determine if each function is injective, surjective or neither.*

(a) $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x, y) = (x + y, xy)$.

(b) $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x_1, \ldots, x_n) = x_1 + \cdots + x_n$.

(c) $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3 - 5x$.

Exercise 1.15. *Prove parts (b) and (c) of Theorem 1.2.*

Exercise 1.16. *Prove each of the following:*

(a) $\sum_{i=1}^{n} \frac{1}{i^2 + i} = \frac{n}{n+1}$ *for every $n \in \mathbb{Z}^+$.*

(b) $2^n > n$ for every $n \in \mathbb{Z}^+$.

(c) $2^n \geq n^2$ for every $n \in \mathbb{N}$ with $n \geq 4$.

Exercise 1.17. *Prove that if for a real number x, the number x^2 is irrational, then so is x.*

Exercise 1.18. *Let E, D be the set of all even and odd integers, respectively. Find a bijection $f : \mathbb{N} \to E$ and another bijection $g : D \to \mathbb{N}$.*

Exercise 1.19. *Suppose $p(x)$ is a polynomial. Prove that for every positive integer n, there is a polynomial $q(x)$ for which the n-th derivative of $e^{p(x)}$ is equal to $q(x)e^{p(x)}$ for every $x \in \mathbb{R}$.*

Exercise 1.20. *Prove that for every positive integer n, there is a polynomial p_n for which the n-th derivative of e^{-1/x^2} at $x \neq 0$ is equal to $p_n(1/x)e^{-1/x^2}$.*

Exercise 1.21. *Carefully prove all vector space properties I-VII of $\mathbb{R}^n$.*

Exercise 1.22. *Let $f : A \to B$ be a function, T be a subset of B. Prove that $f^{-1}(T^c) = (f^{-1}(T))^c$. (Note: For a subset S of A and a subset T of B we have $T^c = B - T$ and $S^c = A - S$.)*

Definition 1.11. A function $f : D \to \mathbb{R}$ is said to be **even** (resp. **odd**) if:

- D is a subset of $\mathbb{R}$ that satisfies $x \in D$ if and only if $-x \in D$, and

- $f(-x) = f(x)$ (resp. $f(-x) = -f(x)$) for every $x \in D$.

Exercise 1.23. *Prove that every function $f : \mathbb{R} \to \mathbb{R}$ can be written as sum of two functions $g, h : \mathbb{R} \to \mathbb{R}$, where g is even and h is odd. Prove the representation $f = g + h$ into sum of an even and an odd function is unique.*

Exercise 1.24. *Suppose functions $f, g : \mathbb{R} \to \mathbb{R}$ are n-times differentiable. Prove*

$$(fg)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x)g^{(n-k)}(x), \textit{ for all } x \in \mathbb{R}.$$

Note: $f^{(0)}(x) = f(x)$ and $g^{(0)}(x) = g(x)$.

1.8 Challenge Problems

Exercise 1.25. *Let $r \geq 2$ be a fixed positive integer, and let $\mathcal{F}$ be an infinite family of distinct sets, each of size r, no two of which are disjoint. Prove that there exists a set of size $r - 1$ that intersects each set in $\mathcal{F}$.*

Exercise 1.26. *Let A be a nonempty set. Suppose $f : \mathcal{P}(A) \to \mathcal{P}(A)$ is a bijection for which for every subsets X and Y of A:*

$$\textit{If } X \subseteq Y, \textit{ then } f(X) \subseteq f(Y).$$

(a) If A is finite, show that if $f(X) \subseteq f(Y)$, then $X \subseteq Y$.

(b) Show part (a) does not necessarily hold when A is infinite.

1.9 Summary

- To prove $A \subseteq B$, start with $x \in A$ and prove $x \in B$.
- To prove two sets A and B are equal we need to show if $x \in A$, then $x \in B$ and vice-versa.
- For a function $f : A \to B$, a subset S of A, and a subset T of B, we have the following:

$$x \in f(S) \text{ iff } x = f(s) \text{ for some } s \in S, \text{ and } y \in f^{-1}(T) \text{ iff } f(y) \in T.$$

- $f^{-1}(\bigcup_{i\in I} T_i) = \bigcup_{i\in I} f^{-1}(T_i)$ and $f^{-1}(\bigcap_{i\in I} T_i) = \bigcap_{i\in I} f^{-1}(T_i)$.
- $f(f^{-1}(T)) \subseteq T$ and $S \subseteq f^{-1}(f(S))$.
- To prove a statement by contradiction, assume the conclusion is false and after taking logical steps obtain a contradiction.
- To prove a statement depending on a positive integer n, first prove the statement for $n = 1$ (basis step), then prove that if the statement is true for n it must be true for $n + 1$ (inductive step).

Week 2

2.1 Subspaces

YouTube Video: https://youtu.be/3iWdoRqTpE0

Definition 2.1. A subset W of $\mathbb{R}^n$ is called a **subspace** of $\mathbb{R}^n$ if W along with the same operations of $\mathbb{R}^n$ satisfies all properties of a vector space, i.e. properties I-VII listed in the previous section.

Theorem 2.1 (Subspace Criterion)**.** *A subset W of $\mathbb{R}^n$ is a subspace if and only if it satisfies all of the following:*

- *W contains the zero vector, and*
- *for all $\mathbf{x}, \mathbf{y} \in W$ and $c \in \mathbb{R}$, we have $\mathbf{x} + \mathbf{y} \in W$ and $c\mathbf{x} \in W$. [We say W is closed under vector addition and scalar multiplication.]*

Example 2.1. Here are some examples of subspaces:

(a) The set of all points (x, y) on a given line $y = mx$ is a subspace of $\mathbb{R}^2$.

(b) The sets $\{\mathbf{0}\}$ and $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$.

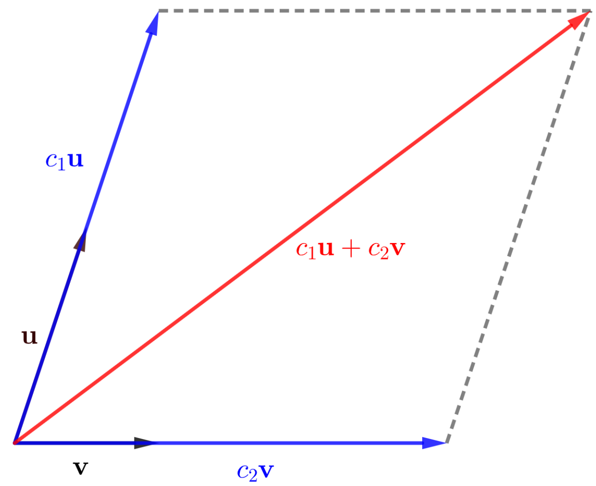
Example 2.2. If W and U are subspaces of $\mathbb{R}^n$, then so is $W \cap U$.

Solution. We will use the subspace criterion Theorem (i.e. Theorem 2.1). First note that $\mathbf{0}$ belongs to both U and W and thus it is in $U \cap W$.

Next, suppose $\mathbf{x}, \mathbf{y} \in U \cap W$ and $c \in \mathbb{R}$. By definition of intersection, $\mathbf{x}$, and $\mathbf{y}$ are in both U and W. Since U and W are both subspaces, by Theorem 2.1, we have $\mathbf{x} + \mathbf{y} \in U$, $\mathbf{x} + \mathbf{y} \in W$, $c\mathbf{x} \in U$ and $c\mathbf{x} \in W$. Therefore, by definition of intersection, $\mathbf{x} + \mathbf{y} \in U \cap W$, and $c\mathbf{x} \in U \cap W$, as desired. $\square$

2.2 Linear Dependence, Spanning, and Basis

Definition 2.2. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be a subset of the vector space $\mathbb{R}^n$, and $\mathbf{w}$ be a vector in $\mathbb{R}^n$. We say $\mathbf{w}$ is a **linear combination** of elements of S if $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$ for some $c_1, \ldots, c_m \in \mathbb{R}$. By definition, if S is the empty set, then the only linear combination of elements of S is $\mathbf{0}$, the zero vector.



We note that every vector $\mathbf{v} = (x, y, z)$ in $\mathbb{R}^3$ can be written as:

$$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1).$$

The vectors $(1, 0, 0), (0, 1, 0)$, and $(0, 0, 1)$ are in some way "independent" of one another. The next definition allows us to formalize this idea of "independence".

Definition 2.3. We say vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ are **linearly dependent** if one of these vectors can be written as a linear combination of the others. Otherwise, we say $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are **linearly independent**.

Example 2.3. Check if each of the following vectors are linearly dependent or linearly independent.

(a) $(1, 0, 0), (0, 1, 0)$, and $(0, 0, 1)$.

(b) $(1, 2, 4), (3, 1, 2)$, and $(4, 3, 6)$.

Theorem 2.2. *The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ are linearly dependent if and only if there are real numbers $c_1, \ldots, c_m$, not all zero, such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$.*
In other words, vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly independent if and only if the following statement is true

$$\text{If } c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0} \text{ for some scalars } c_1, c_2, \ldots, c_m, \text{ then } c_1 = c_2 = \cdots = c_m = 0.$$

Definition 2.4. Given a subspace V of $\mathbb{R}^n$, we say a subset $\mathcal{S}$ of V is **spanning** (or **generating**) if every $\mathbf{v} \in V$ is a linear combination of some vectors in $\mathcal{S}$.

Definition 2.5. We say a subset $\mathcal{B}$ of a subspace V of $\mathbb{R}^n$ is a **basis** if $\mathcal{B}$ is both linearly independent and spanning.

Example 2.4. Prove that $(1,0,0),(0,1,0),(0,0,1)$ form a basis for $\mathbb{R}^3$.

Theorem 2.3. *Let V be a subspace of $\mathbb{R}^n$. Vectors $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_m \in V$ form a basis for V if and only if every vector $\mathbf{w}\in V$ can be uniquely written as $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m$.*

2.3 Some Examples of Subspaces

YouTube Video: https://youtu.be/Qw4qSq0lgJ4

Example 2.5 (Span of vectors)**.** Let $\mathcal{A}$ be a set of vectors in $\mathbb{R}^n$, and let "span $\mathcal{A}$" be the set consisting of all vectors that are linear combinations of some vectors of $\mathcal{A}$. Then span $\mathcal{A}$ is a subspace of $\mathbb{R}^n$.

Definition 2.6. Let A be an $m\times n$ matrix. The **row space of** A denoted by $\mathrm{Row}(A)$ is the subspace of $\mathbb{R}^n$ spanned by the rows of A, and the **column space of** A denoted by $\mathrm{Col}(A)$ is the subspace of $\mathbb{R}^m$ spanned by the columns of A.

Example 2.6. Consider the matrix

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 & 2 \end{pmatrix}$$

Describe the row and column space of the matrix above.

Example 2.7 (Row space and column space)**.** Prove that row space and column space of every matrix are vector spaces.

2.4 Systems of Linear Equations

YouTube Video: https://youtu.be/a6rYNze3Zfw

Suppose we would like to solve the following system of equations:

$$\begin{cases} 3x + 2y - z = 4 \\ x + 3y - 2z = 1 \\ 5x + y - z = 4 \end{cases}$$

In high school algebra, we learn two methods for solving systems of linear equations: substitution and elimination. Substitution could typically get too computational, especially when the number of variables is too large. Elimination often works better, but we still need to keep track of too many things. Our objective is to keep track of all of the work in a more organized fashion. We will keep all coefficients and constants in a single matrix, separating the coefficients from constants by a vertical bar. This matrix is called the **augmented matrix** of the given system. For the example, the augmented matrix of the above system is as follows:

$$\left(\begin{array}{ccc|c} 3 & 2 & -1 & 4 \\ 1 & 3 & -2 & 1 \\ 5 & 1 & -1 & 4 \end{array}\right)$$

In the elimination method, we will add an appropriate multiple of one of the equations to another equation. This means we are doing the same thing to the rows of the augmented matrix. We note that each step is reversible and thus we are not inserting or eliminating any solutions. In this process, three operations are used. The operations (listed below) are called **elementary row operations**.

- **Row Addition:** Adding a scalar multiple of a row to another row.
- **Row Interchange:** Interchanging two rows.
- **Row Scale:** Multiplying a row by a nonzero number.

The objective is to obtain a matrix that satisfies all of the following.

- All zero rows are at the bottom.
- The entries below the first nonzero entry of each row are all zero.
- The leading nonzero entry of each row is to the left of the leading nonzero entry of all rows below it.

Such a matrix is called a matrix in **(row) echelon form**.

If in addition to the above, we also have the following two conditions:

- the first nonzero entry of each row is 1, and
- these 1's are the only nonzero entry of their column.

Then, we say the matrix is in **reduced (row) echelon form**.

To apply this method:

- Interchange rows so that the first entry of the first row is nonzero. (If the first column is all zero, apply this to the first nonzero column.)
- Using the first row and the row addition operation, make all entries below the first nonzero entry zero.
- If possible, by interchanging rows, make the second entry of the second row nonzero. If not, move on to the next entry.
- Repeat this process so that you obtain a matrix in echelon form.
- Scale all rows to obtain 1's as the leading nonzero entries.
- Turn the rest of the entries in columns of each leading 1 into zero to obtain a matrix in reduced echelon form.

Theorem 2.4. *Every matrix can be turned into a matrix in reduced echelon form by applying the three elementary row operations. Furthermore, the reduced echelon form for any matrix is unique.*

Definition 2.7. The leading nonzero entries in a matrix in echelon form are called **pivot** entries. Each column that contains a pivot entry is called a **pivot column**.

Definition 2.8. A system of linear equations is called **homogeneous** if the right hand side of the system is all zeros. In other words, any homogeneous system is of the following form:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k = 0 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nk}x_k = 0 \end{cases}$$

Here, all a_{ij}'s are constants. Note that every homogeneous system has a **trivial** solution

$$x_1 = x_2 = \cdots = x_k = 0.$$

Intuitively, in a homogeneous system if the number of equations is less than the number of variables, we must have infinitely many solutions. Let's test this hypothesis with an example.

Example 2.8. Find all solutions of the system:

$$\begin{cases} 2x_1 - x_2 + 3x_3 + x_4 = 0 \\ x_1 - 3x_2 + x_4 = 0 \\ x_2 - x_3 + 4x_4 = 0 \end{cases}$$

With the method used in the solution of the above example we can prove the following theorem:

Theorem 2.5. *Any homogeneous system that has less equations than variables has a nontrivial solution.*

Corollary 2.1. Every $n+1$ vectors in $\mathbb{R}^n$ are linearly dependent.

2.5 More Examples

Example 2.9. Determine if each of the following is a subspace of $\mathbb{R}^2$.

(a) The set of points on the line $3x + 2y = 1$.

(b) The set of points on the line $4x - 3y = 0$.

(c) The set of points on the unit circle $x^2 + y^2 = 1$.

Solution. (a) This is not a subspace of $\mathbb{R}^2$ since $(0,0)$ does not lie on this line, but the origin lies on every subspace.

(b) This is a subspace. To prove that we will use the Subspace Criterion. First, note that $(0,0)$ is on this line. Suppose (a,b) and (c,d) lie on this line and $r \in \mathbb{R}$. By assumption,

$$4a - 3b = 0, \text{ and } 4c - 3d = 0.$$

We have

$$4(a+c)-3(b+d)=(4a-3b)+(4c-3d)=0+0=0, \text{ and } 4(ra)-3(rb)=r(4a-3b)=r0=0.$$

Therefore, $(a+c,b+d)$ and (ra,rb) both belong to the same line. Thus, this line is a subspace of $\mathbb{R}^2$.

(c) This is not a subspace since it does not contain $(0,0)$. □

Example 2.10. Prove that every set of vectors that contains the zero vector is linearly dependent.

Solution. Let $\mathcal{S}$ be a set of vectors containing $\mathbf{0}$. We see that $1\mathbf{0}=\mathbf{0}$ and the coefficient 1 is nonzero. Therefore, by Theorem 2.2, the set $\mathcal{S}$ is linearly dependent. □

Example 2.11. Prove the vectors $\mathbf{x}=(1,2)$, and $\mathbf{y}=(-1,2)$ form a basis for $\mathbb{R}^2$.

Solution. We need to show $\mathbf{x}$ and $\mathbf{y}$ are linearly independent and spanning.
For linear independence, suppose $c_1\mathbf{x}+c_2\mathbf{y}=\mathbf{0}$, for some real numbers c_1,c_2. Thus, $(c_1-c_2,2c_1+2c_2)=(0,0)$, which implies $c_1=c_2$ and $c_1=-c_2$. This yields $c_1=c_2=0$. Therefore, $\mathbf{x}$ and $\mathbf{y}$ are linearly independent.
For spanning, suppose $(a,b)\in\mathbb{R}^2$. We will have to find $c_1,c_2\in\mathbb{R}$ for which $c_1\mathbf{x}+c_2\mathbf{y}=(a,b)$. This means we need to solve the system:

$$\begin{aligned} c_1-c_2&=a\\ 2c_1+2c_2&=b \end{aligned}$$

Now solve this and find c_1 and c_2 in terms of a and b, and your solution would be complete. □

Example 2.12. Let S and T be two subsets of $\mathbb{R}^n$. Then span $S=$ span T if and only if $S\subseteq$ span T and $T\subseteq$ span S.

Solution. $\Rightarrow$: Suppose span $S=$ span T. By definition of span, $S\subseteq$ span $S=$ span T. Similarly $T\subseteq$ span $T=$ span S, as desired.

$\Leftarrow$: Now, suppose $S\subseteq$ span T, and $T\subseteq$ span S. By definition of span, every element $\mathbf{v}\in$ span T is of the form $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ for some $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_m\in T$. Since $T\subseteq$ span S, we have $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_m\in$ span S. Since span S is a subspace, $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n\in$ span S. Therefore, span $T\subseteq$ span S. Similarly span $S\subseteq$ span T. This implies span $S=$ span T, as desired. □

Example 2.13. Let $\mathbf{v}_1,\ldots,\mathbf{v}_m\in\mathbb{R}^n$ be linearly independent. Consider arbitrary vectors $\mathbf{w}_1,\ldots,\mathbf{w}_m\in\mathbb{R}^k$, and let $\mathbf{x}_1=(\mathbf{v}_1,\mathbf{w}_1),\ldots,\mathbf{x}_m=(\mathbf{v}_m,\mathbf{w}_m)\in\mathbb{R}^{n+k}$ be vectors created by placing components of $\mathbf{v}_j$ followed by components of $\mathbf{w}_j$. Prove that $\mathbf{x}_1,\ldots,\mathbf{x}_m$ are linearly independent.

Solution. Let $c_1,\ldots,c_m\in\mathbb{R}$ be scalars for which

$$c_1\mathbf{x}_1+\cdots+c_m\mathbf{x}_m=\mathbf{0}.$$

Using the way $\mathbf{x}_j$'s are created we have

$$c_1(\mathbf{v}_1, \mathbf{w}_1)+\cdots+c_m(\mathbf{v}_m, \mathbf{w}_m) = \mathbf{0} \Rightarrow (c_1\mathbf{v}_1+\cdots+c_m\mathbf{v}_m, c_1\mathbf{w}_1+\cdots+c_m\mathbf{w}_m) = \mathbf{0} \Rightarrow c_1\mathbf{v}_1+\cdots+c_m\mathbf{v}_m = \mathbf{0}.$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent we obtain $c_1 = \cdots = c_m = 0$, and hence $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are linearly independent. □

Example 2.14. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent vectors of $\mathbb{R}^k$ and $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ are also linearly independent vectors of $\mathbb{R}^k$. Prove that $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_m$ are linearly independent if and only if

$$\text{span}\ \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \cap \text{span}\ \{\mathbf{w}_1, \ldots, \mathbf{w}_m\} = \{\mathbf{0}\}.$$

Solution. For simplicity, let $V = \text{span}\ \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, and $W = \text{span}\ \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$.

$\Rightarrow$: Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_m$ are linearly independent and $\mathbf{x} \in V \cap W$. Thus $\mathbf{x} = \sum\limits_{i=1}^{n} a_i\mathbf{v}_i = \sum\limits_{j=1}^{m} b_j\mathbf{w}_j$, for some $a_i, b_j \in \mathbb{R}$. Therefore, $\sum\limits_{i=1}^{n} a_i\mathbf{v}_i - \sum\limits_{j=1}^{m} b_j\mathbf{w}_j = \mathbf{0}$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_m$ are linearly independent, we must have $a_i = b_j = 0$ and thus $\mathbf{x} = \mathbf{0}$. So, we proved every element of $V \cap W$ is the zero vector. On the other hand $\mathbf{0} \in V \cap W$. Therefore, $V \cap W = \{\mathbf{0}\}$.

$\Leftarrow$: Now assume $V \cap W = \{\mathbf{0}\}$. Suppose $\sum\limits_{i=1}^{n} a_i\mathbf{v}_i + \sum\limits_{j=1}^{m} b_j\mathbf{w}_j = \mathbf{0}$. This implies $\sum\limits_{i=1}^{n} a_i\mathbf{v}_i = -\sum\limits_{j=1}^{m} b_j\mathbf{w}_j \in V \cap W$, which implies $\sum\limits_{i=1}^{n} a_i\mathbf{v}_i = -\sum\limits_{j=1}^{m} b_j\mathbf{w}_j = \mathbf{0}$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{w}_1, \ldots, \mathbf{w}_m$ are linear independent we must have $a_i = b_j = 0$ for all i, j. Therefore, $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_m$ are linearly independent. □

Example 2.15. Determine if each of the following matrices is in echelon form, reduced echelon form or neither. If the matrix is not in reduced echelon form, turn it into reduced echelon form by appropriate elementary row operations. In each step make sure you specify which row operation is used.

(a) $\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 2 & 3 \end{pmatrix}$

(b) $\begin{pmatrix} -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$

Solution. (a) This is not in echelon form. Applying $R_3 + 2R_2$, then, $-R_2$ we obtain the following:

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 2 & 3 \end{pmatrix} \xrightarrow{R_3+2R_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \xrightarrow{-R_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Now, if we apply $R_3/3$ followed by $R_1 + R_3$ we obtain a matrix in reduced echelon form as shown below:

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \xrightarrow{R_3/3} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_1+R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(b) This is in echelon form but is not in reduced echelon form. Applying $-R_1$ and $R_3/5$ yields a matrix in reduced echelon form.

$$\begin{pmatrix} -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} \xrightarrow{-R_1} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} \xrightarrow{R_3/5} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

□

Example 2.16. Find all values of h for which the following system has a solution.

$$\begin{cases} x_1 + 2x_2 - x_3 = 7 + h \\ x_2 - 2x_3 = 3 \\ 2x_1 + 5x_2 - 4x_3 = h \end{cases}$$

Solution. We will row reduce the augmented matrix associated with the above system:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 7+h \\ 0 & 1 & -2 & 3 \\ 2 & 5 & -4 & h \end{array}\right) \xrightarrow{R_3-2R_1} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 7+h \\ 0 & 1 & -2 & 3 \\ 0 & 1 & -2 & -14-h \end{array}\right) \xrightarrow{R_3-R_2} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 7+h \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & -17-h \end{array}\right)$$

Note that the $(1,2)$ entry can be easily made zero by applying $R_1 - 2R_2$. This means, from the first two equations, we can find x_1, x_2 in terms of x_3. For this system to have a solution we need $0 = -17 - h$, which is obtained from the last equation. Therefore, $\boxed{h = -17}$. □

Further Reading: Click here and here for further reading on systems of linear equations and echelon forms.

2.6 Exercises

Exercise 2.1. *Determine if each of the following is a subspace of $\mathbb{R}^n$ once by checking if they satisfy all vector spaces properties I-VII, and once by using the subspace criterion.*

(a) The set of all vectors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ satisfying $x_1 + 2x_2 + \cdots + nx_n = 0$.

(b) The empty set.

(c) The set of all vectors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ satisfying $x_1^2 + x_2^2 + \cdots + x_n^2 = 0$.

(d) The set of all vectors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ satisfying $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$.

Exercise 2.2. *Determine if the following set is a subspace of* $\mathbb{R}^3$*:*

$$\{(x, y, z) \in \mathbb{R}^3 \mid (x^2 + 2y^2 + 3z^2)(z - 3y) = 0\}.$$

Exercise 2.3. *Suppose* U *and* W *are subspaces of* $\mathbb{R}^n$ *for which* $U \cup W$ *is also a subspace. Prove that* $U \subseteq W$ *or* $W \subseteq U$*.*

Hint: Use proof by contradiction.

Exercise 2.4. *Consider the homogeneous system*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nk}x_k = 0 \end{cases}$$

Prove that the set of vectors $(x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$ *satisfying the system above is a subspace of* $\mathbb{R}^k$*.*

Exercise 2.5. *Suppose* V *and* W *are subspaces of* $\mathbb{R}^n$*. Define*

$$V + W = \{\mathbf{v} + \mathbf{w} \mid \mathbf{v} \in V, \text{ and } \mathbf{w} \in W\}.$$

Prove that $V + W$ *is a subspace of* $\mathbb{R}^n$*.*

Exercise 2.6. *Suppose* V *is the subset of* $\mathbb{R}^3$ *consisting of all points* (x, y, z) *for which*

$$x + 2y - z = 0, \text{ and } 2x - 4y + 7z = 0.$$

Prove that V *is a subspace of* $\mathbb{R}^3$*.*

Exercise 2.7. *Suppose* $A = (x_1, y_1)$*, and* $B = (x_2, y_2)$ *are two distinct points on the plane. Let* S *be the set of all points that are equidistant from* A *and* B*. Find the necessary and sufficient condition on points* A, B *for which* S *is a subspace of* $\mathbb{R}^2$*.*

Exercise 2.8. *Prove that the only finite subspace of* $\mathbb{R}^n$ *is the trivial subspace* $\{\mathbf{0}\}$ *containing only the zero vector.*

Exercise 2.9. *Suppose* V, W *are two subspaces of* $\mathbb{R}^n$ *for which* $V \cap W$ *contains at least one nonzero vector. Prove that* $V \cap W$ *is an infinite set.*

Exercise 2.10. *Show the only proper subspace of* $\mathbb{R}$ *is* $\{0\}$*.*

Exercise 2.11. *Prove that if* $n > 1$*, then* $\mathbb{R}^n$ *can be written as the union of all of its proper subspaces.*

Exercise 2.12. *Prove the following set is a subspace of* $\mathbb{R}^3$*, once by showing it satisfies all vector space properties I-VII, and once by applying the Subspace Criterion.*

$$\{(x, y, z) \in \mathbb{R}^3 \mid x + y + 2z = 0, \text{ and } z - 2y + 3x = 0\}$$

Exercise 2.13. *Determine if each of the following matrices is in echelon form, reduced echelon form or neither. If a matrix is not in reduced echelon form, turn it into reduced echelon form by appropriate elementary row operations. In each step make sure you specify which row operation is used.*

(a) $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 5 \end{pmatrix}$

Exercise 2.14. *Using elementary row operations, find all solutions of each system or show the system has no solutions.*

(a) $\begin{cases} x_1 + 3x_2 + x_4 = 5 \\ x_2 - x_3 + 5x_4 = 1 \\ 2x_1 - x_3 + x_4 = 0 \end{cases}$

(b) $\begin{cases} x_1 + x_2 + 3x_3 - x_4 = 5 \\ x_2 - x_3 + 5x_4 = -2 \\ 2x_1 + 3x_2 + 5x_3 + 3x_4 = 0 \end{cases}$

Exercise 2.15. *Show that if a matrix B is obtained by applying an elementary row operation to a matrix A, then $\mathrm{Row}(A) = \mathrm{Row}(B)$. (Hint: Check each of the three row operations separately. You could use Example 2.12.) By an example show that $\mathrm{Col}(A) = \mathrm{Col}(B)$ does not always hold.*

Exercise 2.16. *Describe all 2×2 matrices that are in reduced echelon form.*

Exercise 2.17. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a subspace V of $\mathbb{R}^n$ for which some of them are linearly dependent. Prove that all of them are linearly dependent.*

Exercise 2.18. *Prove that if two vectors in $\mathbb{R}^n$ are linearly dependent, then one of them is a scalar multiple of the other. By an example show that it is not necessarily true that both must be scalar multiples of the other.*

Exercise 2.19. *Find three vectors in $\mathbb{R}^3$ that are linearly dependent but each pair of them are linearly independent.*

Exercise 2.20. *Find all values of real number h for which each equation has a solution or show no such h exists.*

(a) $\begin{cases} x_1 + 3x_2 - x_3 = h + 2 \\ 2x_1 + x_2 - x_3 = h \\ -3x_1 + x_2 + x_3 = h + 1 \end{cases}$

(b) $\begin{cases} x_1 + x_2 - 2x_3 = h + 2 \\ x_1 + x_3 = 5 \\ -3x_1 + x_2 = 3h \end{cases}$

(c) $\begin{cases} x_1 + x_2 - 2x_3 + x_4 = h \\ x_1 + x_3 - 2x_4 = 5 \\ 3x_1 + 2x_2 - 3x_3 = 2h + 9 \end{cases}$

Exercise 2.21. *For a subspace V of $\mathbb{R}^n$ and a vector $\mathbf{x} \in \mathbb{R}^n$, define the set $\mathbf{x} + V$ by*

$$\mathbf{x} + V = \{\mathbf{x} + \mathbf{v} \mid \mathbf{v} \in V\}.$$

Prove that $\mathbf{x} + V$ is a subspace of $\mathbb{R}^n$ if and only if $\mathbf{x} \in V$.

2.7 Challenge Problems

Exercise 2.22. *Let $0 < k < m \leq n$ be integers. Prove that there are m linearly dependent vectors in $\mathbb{R}^n$, every k of which are linearly independent.*

2.8 Summary

- To prove $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent vectors, start with $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ and prove $c_1 = \cdots = c_n = 0$.
- To prove $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are generating, start with an arbitrary vector in the vector space and show it can be written as a linear combination of $\mathbf{v}_j$'s.
- A basis is a set of vectors that are linearly independent and generating.
- Every matrix can be turned into a matrix in echelon form by using three row operations: row addition, row interchange, and row scale.
- W is a subspace of $\mathbb{R}^n$ if W along with the operations of $\mathbb{R}^n$ itself satisfies all properties I-VII of a vector space.
- To prove W is a subspace of $\mathbb{R}^n$ we use the Subspace Criterion: W contains the zero vector, and W is closed under addition and scalar multiplication.

Week 3

3.1 Dimension of a Vector Space

YouTube Video: https://youtu.be/r_HB3Mop058

Theorem 3.1. *Assume V is a subspace of $\mathbb{R}^n$. Then, there is an integer $m \leq n$ for which V has a basis consisting of m distinct vectors. Furthermore, every basis of V contains precisely m vectors.*

Definition 3.1. A subspace V of $\mathbb{R}^n$ is said to have **dimension** m, written as $\dim V = m$, if it has a basis of size m.

Example 3.1. Find the dimension of each of the following vector spaces.

(a) $\mathbb{R}^n$

(b) $\{\mathbf{0}\}$ as a subspace of $\mathbb{R}^n$.

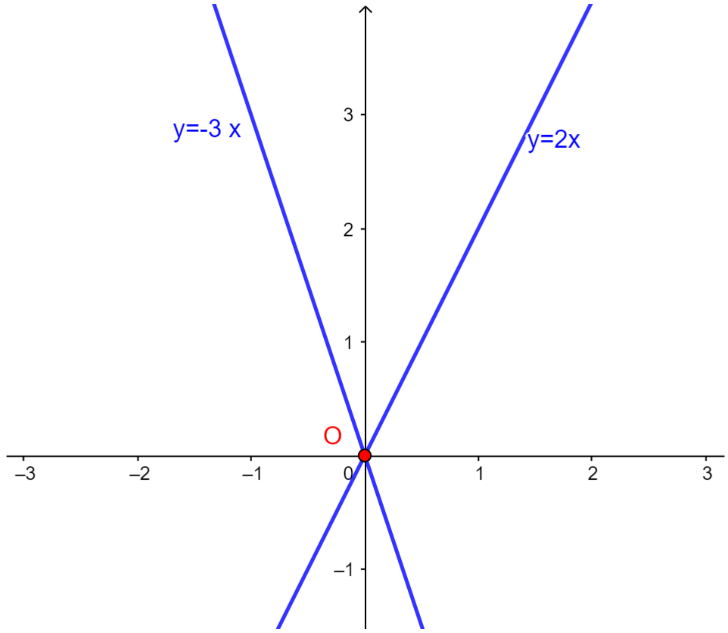
(c) The set of points on the line $y = 3x$ in the xy-plane.

Theorem 3.2. *Let V be a subspace of $\mathbb{R}^n$ of dimension m. Then,*

(a) Every m linearly independent vectors in V form a basis for V.

(b) Every m spanning vectors in V form a basis for V.

Example 3.2. All subspaces of $\mathbb{R}^2$ are either $\{\mathbf{0}\}$, lines through the origin or $\mathbb{R}^2$ itself.



Subspaces of $\mathbb{R}^2$: The origin; All lines through the origin; and $\mathbb{R}^2$ itself.

3.2 Analyzing Pivots

YouTube Video: https://youtu.be/3HKq1MtAUVk

Theorem 3.3. *Let A be a matrix.*

- *The dimension of* $\mathrm{Row}(A)$ *is equal to the number of pivot entries of the echelon form of A. Furthermore, the nonzero rows of the echelon form of A form a basis for* $\mathrm{Row}(A)$.
- *The dimension of* $\mathrm{Col}(A)$ *is equal to the number of pivot entries of the echelon form of A. Furthermore, the pivot columns of A form a basis for* $\mathrm{Col}(A)$.

Example 3.3. Find a basis for $\mathrm{Row}(A)$ and $\mathrm{Col}(A)$, where

$$A = \begin{pmatrix} 0 & 1 & 3 & 0 \\ -1 & -1 & 3 & -1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Remark. Note that to find a basis for $\mathrm{Col}(A)$ we **must** look at pivot columns of A, and not those of its echelon form, but to find a basis for $\mathrm{Row}(A)$ we **must** look at rows of its echelon form with pivot entries, and not those of A.

Definition 3.2. The **rank** of a matrix A, denoted by $\mathrm{rank}\, A$, is the dimension of $\mathrm{Row}(A)$ (which is the same as the dimension of $\mathrm{Col}(A)$).

Definition 3.3. The **transpose** of an $m \times n$ matrix A is an $n \times m$ matrix denoted by A^T whose every (i, j) entry is the (j, i) entry of A.

Theorem 3.4. *For every matrix A, we have* rank $A =$ rank A^T.

Example 3.4. Find a basis for the subspace of $\mathbb{R}^4$ generated by $(1, 2, 0, 1), (-1, 1, 2, 1), (1, 5, 2, 3), (1, 1, -2, 0)$.

Example 3.5 (Null space)**.** Given an $m \times n$ matrix A whose columns are $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{R}^m$, the set of all vectors $\mathbf{v} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ for which

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \mathbf{0}$$

is a subspace of $\mathbb{R}^n$.

Definition 3.4. The subspace defined in the previous example is called the **null space** or the **kernel** of A.

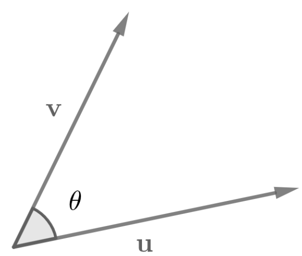
3.3 Inner Products and Angles

YouTube Video: https://youtu.be/ysLQ7RmvSsU

To better understand the geometry of $\mathbb{R}^n$, we need to define the notion of "angles" between vectors.

Example 3.6. Consider the vectors $\mathbf{u} = (x_1, y_1)$ and $\mathbf{v} = (x_2, y_2)$ in $\mathbb{R}^2$. Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$. Using the law of cosines, prove that

$$x_1x_2 + y_1y_2 = \sqrt{x_1^2 + y_1^2}\,\sqrt{x_2^2 + y_2^2}\cos\theta.$$



Definition 3.5. An **inner product** (or **scalar product**) on $\mathbb{R}^n$ is a function that assigns a real number $\langle \mathbf{x}, \mathbf{y} \rangle$ to every pair of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ that satisfies the following for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and all $a, b \in \mathbb{R}$:

(a) $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ if $\mathbf{x} \neq \mathbf{0}$ (Positivity),

(b) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (Symmetry),

(c) $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle$ (Linearity).

Note that by symmetry and linearity with respect to the first vector we can obtain linearity with respect to the second vector:

$$\langle \mathbf{z}, a\mathbf{x} + b\mathbf{y} \rangle = a\langle \mathbf{z}, \mathbf{x} \rangle + b\langle \mathbf{z}, \mathbf{y} \rangle$$

Example 3.7. The following are two examples of inner products in $\mathbb{R}^n$.

(a) $\langle (x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n$.

(b) $\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \rangle = x_1y_1 + 2x_2y_2 + \cdots + nx_ny_n$.

Remark: The first inner product of $\mathbb{R}^n$ in the example above is called the **standard inner product of** $\mathbb{R}^n$. It is also sometimes called the **dot product of** $\mathbb{R}^n$, and is denoted by "$\cdot$". In other words:

$$(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = \sum_{j=1}^{n} x_jy_j.$$

When a particular inner product for $\mathbb{R}^n$ is not specified we will use the dot product above.

The **length** of a vector $\mathbf{v} \in \mathbb{R}^n$ relative to an arbitrary inner product is given by $||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$. Therefore, the length of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ relative to the standard inner product is given by

$$||\mathbf{x}|| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$

which matches the familiar Euclidean distance formulas in $\mathbb{R}^2$ and $\mathbb{R}^3$ (See Theorem 1.3).

By Example 3.6 we notice that, in $\mathbb{R}^2$, when $\theta = \dfrac{\pi}{2}$ we have $\mathbf{v} \cdot \mathbf{w} = 0$. This suggests the following definition:

Definition 3.6. Given an inner product on $\mathbb{R}^n$, we say two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are **orthogonal** (or **perpendicular**) iff $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. We say nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **orthogonal** iff $\mathbf{v}_i$, and $\mathbf{v}_j$ are orthogonal for every $i \neq j$.

Example 3.8. Show that $(1, 3, -1)$ and $(-2, 1, 1)$ are orthogonal vectors of $\mathbb{R}^3$ using the standard inner product of $\mathbb{R}^3$, but not relative to the second inner product stated in Example 3.7.

Example 3.9. Let $\mathbf{e}_i \in \mathbb{R}^n$ be the vector whose i-th component is 1 and whose all other components are zero. Then, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an orthogonal basis for $\mathbb{R}^n$ relative to the dot product of $\mathbb{R}^n$.

YouTube Video: https://youtu.be/8zbPkCAyDNY

Theorem 3.5 (Pythagorean Theorem)**.** *If vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are orthogonal relative to an inner product of $\mathbb{R}^n$, then*

$$||\mathbf{v}||^2 + ||\mathbf{w}||^2 = ||\mathbf{v} + \mathbf{w}||^2.$$

Example 3.6 suggests we should define the angle θ between two vectors $\mathbf{v}, \mathbf{w}$ by $\cos\theta = \dfrac{\langle \mathbf{v}, \mathbf{w} \rangle}{||\mathbf{v}||\ ||\mathbf{w}||}$. In order for us to be able to define the angle between two vectors by $\cos\theta = \dfrac{\langle \mathbf{v}, \mathbf{w} \rangle}{||\mathbf{v}||\ ||\mathbf{w}||}$ we need the following:

Theorem 3.6 (Cauchy-Schwarz Inequality)**.** *Given an inner product $\langle , \rangle$ of $\mathbb{R}^n$, we have*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq ||\mathbf{v}||\ ||\mathbf{w}||.$$

Definition 3.7. The **angle** between two nonzero vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ relative to a given inner product $\langle , \rangle$ is defined by

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{||\mathbf{v}||\ ||\mathbf{w}||}\right).$$

Example 3.10. Find the angle between $(1, 2, -1)$ and $(1, 1, 3)$, once relative to the standard inner product and once relative to the inner product given by

$$\langle (x_1, y_1, z_1), (x_2, y_2, z_2) \rangle = x_1x_2 + 2y_1y_2 + 3z_1z_2.$$

Theorem 3.7. *Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ are orthogonal (nonzero) vectors with respect to some inner product of $\mathbb{R}^n$. Then, they are linearly independent.*

YouTube Video: https://youtu.be/Rtkvgp3EdyU

The following theorem allows us to find an orthogonal basis for any subspace of $\mathbb{R}^n$.

Theorem 3.8 (Gram-Schmidt Orthogonalization Process)**.** *Let $\langle , \rangle$ be an inner product on $\mathbb{R}^n$, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be linearly independent vectors in $\mathbb{R}^n$. Define vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ recursively as follows:*

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 \\
&\vdots \\
\mathbf{w}_m &= \mathbf{v}_m - \frac{\langle \mathbf{v}_m, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_m, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_m, \mathbf{w}_{m-1} \rangle}{\langle \mathbf{w}_{m-1}, \mathbf{w}_{m-1} \rangle}\mathbf{w}_{m-1}
\end{aligned}$$

Then $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ form an orthogonal basis for the subspace of $\mathbb{R}^n$ spanned by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$.

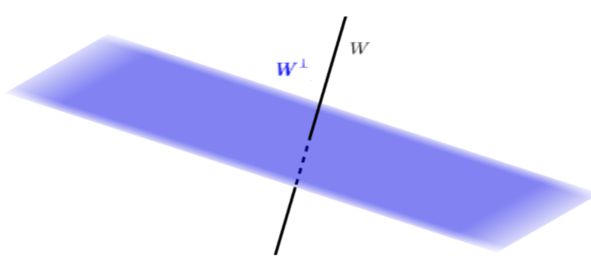
Corollary 3.1. Every subspace of $\mathbb{R}^n$ has an orthogonal basis.

Definition 3.8. We say vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **orthonormal** relative to an inner product $\langle , \rangle$ if they are orthogonal and $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ for every i. (i.e. all of them have length 1.)

Example 3.11. Find an orthogonal basis for the subspace of $\mathbb{R}^4$ generated by $(1, 2, 0, -1), (0, 1, 0, 2), (0, 0, 2, 1)$.

Definition 3.9. Let W be a subspace of $\mathbb{R}^n$. The **orthogonal complement of** W relative to an inner product $\langle , \rangle$, denoted by $W^\perp$, is defined as

$$W^\perp = \{\mathbf{v} \in \mathbb{R}^n \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}.$$



Theorem 3.9. *Let W be a subspace of $\mathbb{R}^n$. Then $W^\perp$ is a subspace of $\mathbb{R}^n$ and*

$$\dim W + \dim W^\perp = n.$$

Proof. The fact that $W^{\perp}$ is a subspace is left as an exercise.

Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ be an orthogonal basis for W, and $\{\mathbf{w}_{k+1}, \ldots, \mathbf{w}_m\}$ be an orthogonal basis for $W^{\perp}$. Since each element of W is orthogonal to each element of $W^{\perp}$, $\mathcal{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is an orthogonal set (of nonzero vectors) and thus it is linearly independent. It is left to prove $\mathcal{B}$ is generating. Let $\mathbf{v} \in \mathbb{R}^n$. Using a method similar to Gram-Schmidt process, we see that $\mathbf{x} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \cdots - \frac{\langle \mathbf{v}, \mathbf{w}_k \rangle}{\langle \mathbf{w}_k, \mathbf{w}_k \rangle}\mathbf{w}_k$ is orthogonal to $\mathbf{w}_1, \ldots, \mathbf{w}_k$ and thus all elements of W, by linearity of inner product. Therefore, $\mathbf{x} \in W^{\perp}$. This implies there are scalars $c_{k+1}, \ldots, c_m$ for which

$$\mathbf{x} = c_{k+1}\mathbf{w}_{k+1} + \cdots + c_m \mathbf{w}_m,$$

This means

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_k \rangle}{\langle \mathbf{w}_k, \mathbf{w}_k \rangle}\mathbf{w}_k + c_{k+1}\mathbf{w}_{k+1} + \cdots + c_m \mathbf{w}_m \in \operatorname{span} \mathcal{B}.$$

We showed $\mathcal{B}$ is a basis for $\mathbb{R}^n$. Since $\dim \mathbb{R}^n = n$ we conclude that $m = n$, as desired. □

3.4 Warm-ups

Example 3.12. Find the angle between vectors $(1, 0, -1)$ and $(2, 1, 2)$ in $\mathbb{R}^3$ using the standard inner product.

Solution. If the angle between these two vectors is θ, then we have

$$\cos\theta = \frac{(1, 0, -1) \cdot (2, 1, 2)}{||(1, 0, -1)||\ ||(2, 1, 2)||} = \frac{2 + 0 - 2}{\sqrt{1 + 0 + 1}\sqrt{4 + 1 + 4}} = 0 \Rightarrow \theta = \frac{\pi}{2}.$$

Therefore, the angle is $\pi/2$. □

3.5 More Examples

Example 3.13. Find a vector in $\mathbb{R}^3$ in the direction of $(1, -2, 2)$ that has length 4 with the Euclidean length.

Solution. The vector must be of the form $c(1, -2, 2)$ where c is a positive constant. The length must be four and thus $c^2 + 4c^2 + 4c^2 = 4^2$, which means $c = \frac{4}{3}$. The answer is $\left(\frac{4}{3}, \frac{-8}{3}, \frac{8}{3}\right)$.

Another method would be to notice that $||(1, -2, 2)|| = \sqrt{1 + 4 + 4} = 3$. Thus, by properties of length we have

$$\left|\left|\frac{4}{3}(1, -2, 2)\right|\right| = \frac{4}{3} \cdot 3 = 4.$$

This, of course, yields the same answer. □

Example 3.14. Suppose $\{\mathbf{v}, \mathbf{w}\}$ is a basis for a 2-dimensional subspace V of $\mathbb{R}^n$. Let a, b be two real numbers. Prove that $\{\mathbf{v} + a\mathbf{w}, \mathbf{v} + b\mathbf{w}\}$ is a basis for V if and only if $a \neq b$.

Solution. $\Rightarrow$: Suppose $\{\mathbf{v} + a\mathbf{w}, \mathbf{v} + b\mathbf{w}\}$ is a basis for V. Thus, $\mathbf{v} + a\mathbf{w}$ and $\mathbf{v} + b\mathbf{w}$ cannot be scalar multiples and thus $\mathbf{v}+a\mathbf{w} \neq \mathbf{v}+b\mathbf{w}$, which means $a \neq b$, as desired. (Since $a = b$ implies $\mathbf{v}+a\mathbf{w} = \mathbf{v}+b\mathbf{w}$.)

$\Leftarrow$: Now, assume $a \neq b$. We will show $\mathbf{v} + a\mathbf{w}, \mathbf{v} + b\mathbf{w}$ are linearly independent. Suppose $c_1(\mathbf{v} + a\mathbf{w}) + c_2(\mathbf{v} + b\mathbf{w}) = \mathbf{0}$, for some scalars c_1, c_2. This means $(c_1 + c_2)\mathbf{v}_1 + (ac_1 + bc_2)\mathbf{w} = \mathbf{0}$. Since $\mathbf{v}, \mathbf{w}$ are linearly independent we must have $c_1 + c_2 = ac_1 + bc_2 = 0$. Eliminating c_1 from the two equations we obtain $(b - a)c_2 = 0$, which implies $c_2 = 0$ since $a \neq b$. This implies $c_1 = 0$ from the equation $c_1 + c_2 = 0$. This means $\mathbf{v} + a\mathbf{w}, \mathbf{v} + b\mathbf{w}$ are linearly independent. Since the dimension of V is 2, $\{\mathbf{v} + a\mathbf{w}, \mathbf{v} + b\mathbf{w}\}$ is a basis for V. $\square$

Example 3.15. Prove that if $||\cdot||$ is a length relative to an inner product of $\mathbb{R}^n$ and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, then

$$||\mathbf{v} + \mathbf{w}||^2 + ||\mathbf{v} - \mathbf{w}||^2 = 2(||\mathbf{v}||^2 + ||\mathbf{w}||^2).$$

Solution. By definition we have $||\mathbf{v} \pm \mathbf{w}||^2 = \langle \mathbf{v} \pm \mathbf{w}, \mathbf{v} \pm \mathbf{w}\rangle$. By linearity and symmetry this simplifies to

$$\langle \mathbf{v} \pm \mathbf{w}, \mathbf{v} \pm \mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle \pm \langle \mathbf{v}, \mathbf{w}\rangle \pm \langle \mathbf{w}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle \pm 2\langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle.$$

Summing the two together and using the fact that $\langle \mathbf{v}, \mathbf{v}\rangle = ||\mathbf{v}||^2$ and $\langle \mathbf{w}, \mathbf{w}\rangle = ||\mathbf{w}||^2$ we obtain the result. $\square$

Example 3.16. Consider the subspace V of $\mathbb{R}^4$ spanned by $\mathbf{v} = (1, 1, 1, 0)$ and $\mathbf{w} = (0, 1, 2, 0)$. Find a basis for the orthogonal complement of V relative to the dot product.

Solution. Note that since $\mathbf{v}$ and $\mathbf{w}$ are not multiples of each other, $\dim V = 2$. By Theorem 3.9, we have $\dim V^\perp = 4 - 2 = 2$.

We will find a basis for $\mathbb{R}^4$ containing $\mathbf{v}$ and $\mathbf{w}$. To do that, we will place these vectors in rows of a matrix, and row reduce the matrix as below:

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{pmatrix}$$

This matrix is already in echelon form. By adding $\mathbf{e}_3$ and $\mathbf{e}_4$ to the rows of this matrix, we obtain a matrix in echelon form. Thus, $\mathbf{v}, \mathbf{w}, \mathbf{e}_3, \mathbf{e}_4$ form a basis for $\mathbb{R}^4$. Now, can apply the Gram-Schmidt process to $\mathbf{v}, \mathbf{w}, \mathbf{e}_3, \mathbf{e}_4$.

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v} = (1, 1, 1, 0) \\
\mathbf{w}_2 &= \mathbf{w} - \frac{\mathbf{w} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 = (-1, 0, 1, 0) \\
\mathbf{w}_3 &= \mathbf{e}_3 - \frac{\langle \mathbf{e}_3, \mathbf{w}_1\rangle}{\langle \mathbf{w}_1, \mathbf{w}_1\rangle}\mathbf{w}_1 - \frac{\langle \mathbf{e}_3, \mathbf{w}_2\rangle}{\langle \mathbf{w}_2, \mathbf{w}_2\rangle}\mathbf{w}_2 = (\frac{1}{6}, -\frac{1}{3}, \frac{1}{6}, 0) \\
\mathbf{w}_4 &= \mathbf{e}_4 - \frac{\langle \mathbf{e}_4, \mathbf{w}_1\rangle}{\langle \mathbf{w}_1, \mathbf{w}_1\rangle}\mathbf{w}_1 - \frac{\langle \mathbf{e}_4, \mathbf{w}_2\rangle}{\langle \mathbf{w}_2, \mathbf{w}_2\rangle}\mathbf{w}_2 - \frac{\langle \mathbf{e}_4, \mathbf{w}_3\rangle}{\langle \mathbf{w}_3, \mathbf{w}_3\rangle}\mathbf{w}_3 = \mathbf{e}_4
\end{aligned}$$

The vectors $\mathbf{w}_3, \mathbf{w}_4$ are linearly independent and are in $V^\perp$. Since $\dim V^\perp = 2$, the two vectors $\mathbf{w}_3$ and $\mathbf{w}_4$ form a basis for $V^\perp$. $\square$

Example 3.17. Let $c \in \mathbb{R}$ be a constant. For which constants c does the function below,

$$\langle (x_1, y_1), (x_2, y_2) \rangle = x_1x_2 + cy_1y_2$$

defines an inner product on $\mathbb{R}^2$?

Scratch: Positivity means, from $x^2 + cy^2 = 0$ we need to be able to imply $x = y = 0$. This means c cannot be nonpositive.

Solution. We claim that the given expression is an inner product if and only if c is positive.
If $c \leq 0$, then $\langle (0,1), (0,1) \rangle = c \leq 0$, violating the positivity property. Thus, it is not an inner product.

Now assume $c > 0$ and let $\mathbf{x} = (x_1, y_1), \mathbf{y} = (x_2, y_2), \mathbf{z} = (x_3, y_3) \in \mathbb{R}^2$, $a, b \in \mathbb{R}$. If $(x_1, y_1) \neq \mathbf{0}$, then $x_1^2 + cy_1^2 > 0$, since x_1^2 and y_1^2 are both nonnegative, c is positive and not both x_1 and y_1 are zero. This means we obtain the positivity.
$\langle \mathbf{x}, \mathbf{y} \rangle = x_1x_2 + cy_1y_2 = x_2x_1 + cy_2y_1 = \langle \mathbf{y}, \mathbf{x} \rangle$. This proves the symmetry.
$\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = (ax_1 + bx_2)x_3 + c(ay_1 + by_2)y_3 = a(x_1x_3 + cy_1y_3) + b(x_2x_3 + cy_2y_3) = a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle$. This proves the linearity. Therefore, when c is positive, the given function defines an inner product. $\square$

Example 3.18. Suppose $\mathbf{v}_1, \mathbf{v}_2$ form an orthogonal basis for $\mathbb{R}^2$ with respect to some inner product. Prove that if $\mathbf{w}$ is orthogonal to $\mathbf{v}_1$, then $\mathbf{w} = c\mathbf{v}_2$ for some scalar c.

Solution. Since $\mathbf{v}_1, \mathbf{v}_2$ is a basis for $\mathbb{R}^2$, there are scalars c_1, c_2 for which $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. By assumption and linearity of inner product we obtain the following:

$$\langle \mathbf{w}, \mathbf{v}_1 \rangle = 0 \Rightarrow c_1\langle \mathbf{v}_1, \mathbf{v}_1 \rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_1 \rangle = 0.$$

Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal we obtain $c_1\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 0$. Since $\mathbf{v}_1$ is an element of a basis, we know $\mathbf{v}_1 \neq \mathbf{0}$, and by positivity of inner products we conclude that $c_1 = 0$, which means $\mathbf{w} = c_2\mathbf{v}_2$, as desired. $\square$

Example 3.19. Prove that for all real numbers $x_1, \ldots, x_n, y_1, \ldots, y_n$ we have

$$(x_1y_1 + \cdots + x_ny_n)^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

Solution. We will use the Cauchy-Schwarz Inequality for the standard inner product of $\mathbb{R}^n$. Consider the two vectors

$$\mathbf{v} = (x_1, \ldots, x_n), \text{ and } \mathbf{w} = (y_1, \ldots, y_n) \text{ in } \mathbb{R}^n.$$

We have

$$\mathbf{v} \cdot \mathbf{w} = x_1y_1 + \cdots + x_ny_n, ||\mathbf{v}|| = \sqrt{x_1^2 + \cdots + x_n^2}, \text{ and } ||\mathbf{w}|| = \sqrt{y_1^2 + \cdots + y_n^2}.$$

Applying the Cauchy-Schwarz Inequality we obtain:

$$|x_1y_1 + \cdots + x_ny_n| \leq \sqrt{x_1^2 + \cdots + x_n^2} \cdot \sqrt{y_1^2 + \cdots + y_n^2}.$$

Squaring both sides we obtain the result. $\square$

Example 3.20. Find a basis for the orthogonal complement of $V = \text{span}\,\{(1, 2, -1), (0, 1, 1)\}$ under the standard inner product.

Solution. Placing these vectors into rows of a matrix we obtain the following matrix:

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

This matrix is in echelon form and adding $(0, 0, 1)$ gives us another matrix in echelon form:

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

This means the vectors $\mathbf{v}_1 = (1, 2, -1), \mathbf{v}_2 = (0, 1, 1), \mathbf{v}_3 = (0, 0, 1)$ are linearly independent and thus they form a basis for $\mathbb{R}^3$. Since the dimension of V is 2, by Theorem 3.9 the dimension of its orthogonal complement is 1. Using the Gram-Schmidt process we will find the following vectors:

$$\mathbf{w}_1 = \mathbf{v}_1 = (1, 2, -1),$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 = \left(-\frac{1}{6}, \frac{2}{3}, \frac{7}{6}\right) = \frac{1}{6}(-1, 4, 7),$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2}\mathbf{w}_2 = \left(\frac{3}{11}, -\frac{1}{11}, \frac{1}{11}\right) = \frac{1}{11}(3, -1, 1)$$

Thus, $\mathbf{w}_3$ is orthogonal to every element of V. Therefore, $\mathbf{w}_3$ forms a basis for the orthoghonal complement of V. $\square$

3.6 Exercises

Exercise 3.1. *Determine which of the following vectors form a basis for the appropriate* $\mathbb{R}^n$.

(a) $(1, 2), (-2, 1)$.

(b) $(1, 0, 1), (1, 1, 2), (-1, -2, -3)$.

(c) $(1, 0), (2, 3), (1, 1)$.

(d) $(1, 0, 0), (0, 1, 1), (0, 1, 2)$.

Exercise 3.2. *Suppose the homogeneous system*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases}$$

has only the trivial solution. Prove that for every $b_1, b_2, \ldots, b_n \in \mathbb{R}$*, the system*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

has a unique solution.

Hint: Use $\dim \mathbb{R}^n = n$, and consider the vectors $(a_{11}, a_{21}, \ldots, a_{n1}), \ldots, (a_{1n}, a_{2n}, \ldots, a_{nn})$.

Exercise 3.3. *Determine the dimension of each vector space.*

(a) The subspace of $\mathbb{R}^3$ *generated by vectors* $(1, 2, -1), (2, 3, 4)$*, and* $(4, 10, 2)$.

(b) The subspace of $\mathbb{R}^3$ *generated by* $(1, 2, 0), (-1, 1, 1)$*, and* $(1, 5, 1)$.

Exercise 3.4. *Let* V *be a subspace of* $\mathbb{R}^n$*. Prove that if* $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is a linearly independent set of vectors in* V*, then there is a basis for* V *that contains* $\mathcal{A}$.

Hint: Consider the subspace generated by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. If this subspace is not V, and then add an element $\mathbf{v}_{k+1}$ from V, but outside of span $\mathcal{A}$ to the set $\mathcal{A}$. Show this new larger set is linearly independent. Repeat this until you get a basis. You must show this process ends. This is where you should use the fact that every $n + 1$ vectors in $\mathbb{R}^n$ are linearly dependent.

Exercise 3.5. *Suppose* W *and* V *are subspaces of* $\mathbb{R}^n$ *for which* $W \subseteq V$*. Prove that if* $\dim W = \dim V$*, then* $W = V$.

Exercise 3.6. *Let* V *be a subspace of* $\mathbb{R}^n$*. Prove that if* $\mathcal{A}$ *is a spanning subset of* V*, then there is a basis for* V *that is a subset of* $\mathcal{A}$.

Exercise 3.7. *Find the dimension of the vector space spanned by* $(0, 0, 1, 1), (1, 1, 0, 0)$*, and* $(1, 1, 0, 1)$.

Exercise 3.8. *Find the angle between:*

(a) $(1, 2, -1)$ *and* $(0, 2, -1)$ *in* $\mathbb{R}^3$ *with the standard inner product.*

(b) $(1, 1, 5)$ *and* $(1, -1, 0)$ *with the inner product given by* $\langle (x_1, y_1, z_1), (x_2, y_2, z_2) \rangle = x_1x_2 + 2y_1y_2 + 3z_1z_2$.

Exercise 3.9. *Determine if the triangle whose vertices are* $A = (1, 2, 2), B = (-1, 1, 0), C = (2, -2, 1)$ *is a right triangle.*

Exercise 3.10. *Consider $\mathbb{R}^3$ with the standard inner product. Find an orthogonal basis for $\mathbb{R}^3$ for which one of the elements of this basis is $(1, 2, -1)$.*

Hint: Use the idea of echelon form to extend this vector to a basis. Then apply Gram-Schmidt. See Example 3.20.

Exercise 3.11. *Find all real numbers c for which the vectors $(1, 2, c)$ and $(-1, -c, c+1)$ are orthogonal with respect to the standard inner product. For this value of c, give an example of an inner product where these two vectors are not orthogonal.*

Exercise 3.12. *Find all inner products of $\mathbb{R}^2$, if any, for which $||\mathbf{e}_1|| = 4$, $||\mathbf{e}_2|| = 3$ and the angle between $\mathbf{e}_1$ and $\mathbf{e}_2$ is $\frac{\pi}{3}$.*

Hint: First, use the given assumptions to find $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$. Next, write (x_1, x_2) and (y_1, y_2) as linear combinations of $\mathbf{e}_1, \mathbf{e}_2$. Then use linearity, symmetry and the given assumptions to evaluate $\langle (x_1, x_2), (y_1, y_2) \rangle$. Finally, show the result is in fact an inner product.

Exercise 3.13. *Suppose $c_1, \ldots, c_n$ are real constants. Define a function $\langle , \rangle$ by*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{j=1}^{n} c_j x_j y_j.$$

(a) Show $\langle , \rangle$ is linear and symmertic.

(b) Show $\langle , \rangle$ is an inner product iff $c_1, \ldots, c_n$ are all positive.

Exercise 3.14. *Suppose $\langle , \rangle$ is an inner product of $\mathbb{R}^n$. Let $\mathbf{v} = (x_1, \ldots, x_n)$ and $\mathbf{w} = (y_1, \ldots, y_n)$ be two vectors in $\mathbb{R}^n$.*

(a) By writing $\mathbf{v}$ and $\mathbf{w}$ as linear combinations of $\mathbf{e}_1, \ldots, \mathbf{e}_n$, and applying linearity prove that $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{k=1}^{n} \sum_{j=1}^{n} \langle \mathbf{e}_k, \mathbf{e}_j \rangle x_k y_j$.

(b) Using the previous part and Exercise 3.13, deduce that every inner product of $\mathbb{R}^n$ for which $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are orthogonal is of the form $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^{n} c_j x_j y_j$ for some positive real numbers $c_1, \ldots, c_n$, and every such function is an inner product.

Exercise 3.15. *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ be orthogonal vectors with respect to some inner product of $\mathbb{R}^n$. Prove that*

$$||\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n||^2 = ||\mathbf{v}_1||^2 + ||\mathbf{v}_2||^2 + \cdots + ||\mathbf{v}_n||^2.$$

Hint: Use the Pythagorean Theorem and proof by induction.

Exercise 3.16. *Suppose $\langle , \rangle$ is an inner product of $\mathbb{R}^n$. Using linearity prove that for every $\mathbf{w} \in \mathbb{R}^n$ we have $\langle \mathbf{0}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{0} \rangle = 0$. Deduce the Cauchy-Schwarz inequality in the case when $\mathbf{v} = \mathbf{0}$. (In class we assumed $\mathbf{v} \neq \mathbf{0}$.)*

Exercise 3.17. *Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for $\mathbb{R}^n$. For every two vectors*

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n, \text{ and } \mathbf{w} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n \text{ in } \mathbb{R}^n,$$

define $\langle \mathbf{v}, \mathbf{w} \rangle = a_1b_1 + a_2b_2 + \cdots + a_nb_n$. Prove that this defines an inner product on $\mathbb{R}^n$.

Exercise 3.18. *Prove the converse of the Pythagorean Theorem stated below:*

Given an inner product $\langle , \rangle$ of $\mathbb{R}^n$ and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, if $||\mathbf{v} + \mathbf{w}||^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2$, then $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Exercise 3.19. *Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, and $\langle , \rangle$ be an inner product of $\mathbb{R}^n$. Prove that $|\langle \mathbf{v}, \mathbf{w} \rangle| = ||\mathbf{v}||\ ||\mathbf{w}||$ if and only if $\mathbf{w}$ is a scalar multiple of $\mathbf{v}$ or $\mathbf{v} = \mathbf{0}$.*

Hint: Follow the proof of Cauchy-Schwarz inequality and see when the equality holds.

Exercise 3.20. *Let θ be the angle between two nonzero vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Prove that*

(a) if $\theta = 0$, then $\mathbf{v} = c\mathbf{w}$ for some positive real number c.

(b) if $\theta = \pi$, then $\mathbf{v} = c\mathbf{w}$ for some negative real number c.

Exercise 3.21. *Let A be an $m \times n$ matrix with real entries. We have shown that* $\mathrm{Row}(A)$ *and* $\mathrm{Ker}\ A$ *are both subspaces of $\mathbb{R}^n$. What is the relationship between* $\mathrm{Ker}\ A$ *and* $(\mathrm{Row}(A))^\perp$*? Justify your answer.*

Hint: Show that a vector is in $(\mathrm{Row}(A))^\perp$ if and only if it is orthogonal to all rows of A.

Exercise 3.22. *Let S be a nonempty subset of $\mathbb{R}^n$, and $\langle , \rangle$ be an inner product of $\mathbb{R}^n$. Prove that $S^\perp$ defined by*

$$S^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{s} \rangle = 0 \text{ for all } \mathbf{s} \in S\}$$

is a subspace of $\mathbb{R}^n$.

Exercise 3.23. *Suppose W is a subspace of $\mathbb{R}^n$. Prove that $(W^\perp)^\perp = W$.*

Hint: Show the dimension of both sides are the same, and the right hand side is a subset of the left hand side.

Exercise 3.24. *Suppose $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ satisfy $\mathbf{v} \cdot \mathbf{w} = 0$. Prove that $||\mathbf{v} + \mathbf{w}|| = ||\mathbf{v} - \mathbf{w}||$ once using properties of inner product, once using the definition of dot product, and once using geometry. (For the geometric proof, you may use known facts from Euclidean geometry.)*

Exercise 3.25. *For every inner product $\langle , \rangle$ on $\mathbb{R}^n$, its corresponding length, and every two vectors $\mathbf{u}, \mathbf{v}$ prove the polarization identity stated below:*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{2}\left(||\mathbf{u} + \mathbf{v}||^2 - ||\mathbf{u} - \mathbf{v}||^2\right).$$

Exercise 3.26. *Suppose $\mathbf{u}, \mathbf{v}$ are nonzero vectors in $\mathbb{R}^n$ and c is a nonzero scalar. Suppose the angle between $\mathbf{u}$ and $\mathbf{v}$ is α and the angle between $c\mathbf{u}$ and $\mathbf{v}$ is β. Prove that:*

(a) if $c > 0$, then $\alpha = \beta$.

(b) if $c < 0$, then $\alpha + \beta = \pi$.

Exercise 3.27. *Suppose V, W are subspaces of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Prove that $V \times W$ is a subspace of $\mathbb{R}^{m+n}$ and that*

$$\dim(V \times W) = \dim V + \dim W.$$

Note that vectors in $V \times W$ are obtained by placing coordinates of each vector of V followed by coordinates of each vector of W. In other words, in order to obtain a vector in R^{m+n} we drop the brackets for vectors in V and W.

Definition 3.10. Given two points $A, B \in \mathbb{R}^n$, the **vector** $\overrightarrow{AB}$ is defined as $B - A$. The **line segment** AB is the set of all points $X \in \mathbb{R}^n$ for which $X = tA + (1-t)B$ for some $t \in [0,1]$. The set $AB \cup BC \cup CA$, denoted by ABC, is called a **triangle** if $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are linearly independent vectors. The three segments AB, BC and CA are called sides of triangle ABC. A **side length** in triangle ABC is the length of each of the three vectors $\overrightarrow{AB}$, $\overrightarrow{AC}$ and $\overrightarrow{BC}$. The **angle** between vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$ is denoted by $\angle BAC$ or $\angle CAB$. Similarly $\angle ABC, \angle CBA, \angle ACB, \angle BCA$ are all defined. Note that to define this angle $\mathbb{R}^n$ must be equipped with an inner product. Each of the three points A, B, C is called a **vertex** of triangle ABC. A triangle ABC is called a **right triangle** if one is its angles is $\dfrac{\pi}{2}$. A triangle is called **isosceles** if two of its side lengths are equal. A triangle is called **equilateral** if all three side lengths are equal.

Exercise 3.28. *Prove that for every three points $A, B, C \in \mathbb{R}^n$ we have $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.*

Exercise 3.29. *Prove that if ABC is a triangle, then vectors $\overrightarrow{AB}, \overrightarrow{BC}, \overrightarrow{AC}$ are pairwise linearly independent, but all three of them are linearly dependent.*

Exercise 3.30. *Prove the Law of Cosines:*

> *Consider $\mathbb{R}^n$ equipped with an inner product. Suppose ABC is a triangle in $\mathbb{R}^n$. Then,*
> $$||\overrightarrow{BC}||^2 = ||\overrightarrow{AB}||^2 + ||\overrightarrow{AC}||^2 - 2||\overrightarrow{AB}||\ ||\overrightarrow{AC}|| \cos(\angle BAC).$$

Exercise 3.31. *Prove the Law of Sines:*

> *Consider $\mathbb{R}^n$ equipped with an inner product. Suppose ABC is a triangle in $\mathbb{R}^n$. Then,*
> $$\frac{||\overrightarrow{AB}||}{\sin(\angle ACB)} = \frac{||\overrightarrow{AC}||}{\sin(\angle ABC)}.$$

Exercise 3.32. *Assume $\mathbb{R}^n$ is equipped with an inner product. Show that a triangle ABC in $\mathbb{R}^n$ is isosceles if and only if two of its angles are congruent.*

Exercise 3.33. *Assume $\mathbb{R}^n$ is equipped with an inner product. Show that a triangle ABC in $\mathbb{R}^n$ is equilateral if and only if all of its angles are $\dfrac{\pi}{3}$.*

3.7 Challenge Problems

Exercise 3.34. *Let k be a positive integer. Find the smallest positive integer n for which there are k nonzero vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ for which v_i and v_j are orthogonal for every i and j for which $i > j+1$.*

Exercise 3.35. *Prove the Angle Bisector Theorem:*

Consider $\mathbb{R}^n$ equipped with an inner product. Suppose ABC is a triangle in $\mathbb{R}^n$, and D is a point on side BC for which $\angle BAD = \angle DAC$. Then, $\dfrac{||\overrightarrow{AB}||}{||\overrightarrow{AC}||} = \dfrac{||\overrightarrow{BD}||}{||\overrightarrow{DC}||}$.

Exercise 3.36. *Consider $\mathbb{R}^n$ equipped with an inner product. Prove the altitudes of ABC are concurrent. In other words, prove there is a unique point $H \in \mathbb{R}^n$ for which:*

$$\langle \overrightarrow{AH}, \overrightarrow{BC} \rangle = \langle \overrightarrow{BH}, \overrightarrow{AC} \rangle = \langle \overrightarrow{CH}, \overrightarrow{AB} \rangle = 0.$$

Exercise 3.37. *Consider $\mathbb{R}^n$ equipped with an inner product. Assume D is a point on the segment BC of triangle ABC. Prove $\angle BAC = \angle BAD + \angle DAC$.*

Exercise 3.38. *Consider $\mathbb{R}^n$ equipped with an inner product. Prove that the sum of the three angles of every triangle is π.*

3.8 Summary

- The number of pivot entries is the same as both the dimension of row space and the dimension of column space.
- To find a basis for a space spanned by a set of vectors in $\mathbb{R}^n$:
  - Place these vectors in rows of a matrix.
  - Row reduce this matrix.
  - The nonzero rows of the echelon form, create a basis for the desired space.
- The dimension of a vector space is the number of vectors in a basis of that vector space.
- In a vector space of dimension n every $n+1$ (or more) vectors are linearly dependent.
- Rank of a matrix is the dimension of its column space.
- To show $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis for a vector space V we can do one of the following:
  - $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent and spanning.
  - $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent and $\dim V = n$.
  - $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are spanning and $\dim V = n$.

- To check if $\langle \mathbf{v}, \mathbf{w} \rangle$ is an inner product we need to check if it satisfies three properties: Positivity, Symmetry, and Linearity.
- The angle θ between two vectors $\mathbf{v}, \mathbf{w}$ is given by $\cos\theta = \dfrac{\langle \mathbf{v}, \mathbf{w} \rangle}{||\mathbf{v}||\ ||\mathbf{w}||}$.
- If the angle between two vectors is $\pi/2$ we say the two vectors are orthogonal.
- Pythagorean Theorem: If $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, then $||\mathbf{v}+\mathbf{w}||^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2$.
- Cauchy-Schwarz Inequality: In any inner product space $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq ||\mathbf{v}||\ ||\mathbf{w}||$.
- Given linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in an inner product space, we can use the Gram-Schmidt process to find an orthogonal basis for the subspace spanned by $\mathbf{v}_i$'s.
- The orthogonal complement of a subspace W is the subspace consisting of all vectors that are orthogonal to all vectors in W.
- To find a basis for $W^\perp$:
  - First find a basis for W.
  - Extend that basis to a basis of $\mathbb{R}^n$ using echelon form.
  - Start from vectors in W and apply the Gram-Schmidt process. This produces an orthogonal basis for W followed by an orthogonal basis for $W^\perp$.

Week 4

4.1 Norms

YouTube Video: https://youtu.be/UgwMfs9w91s

The definition of length, $||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$, in the previous section relied on an inner product of $\mathbb{R}^n$, however the concept of "length" can be defined independently. We will define a "norm" (or length) to be an assignment of nonnegative real numbers to vectors that satisfies certain properties that we expect from a geometric length.

Definition 4.1. A **norm** on $\mathbb{R}^n$ is a function that assigns to any vector $\mathbf{v} \in \mathbb{R}^n$ a nonnegative real number $||\mathbf{v}||$ that satisfies all of the following:

(a) $||\mathbf{v}|| > 0$ for every nonzero $\mathbf{v} \in \mathbb{R}^n$ (Positivity),

(b) $||\mathbf{v} + \mathbf{w}|| \leq ||\mathbf{v}|| + ||\mathbf{w}||$ for every $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ (Triangle Inequality), and

(c) $||c\mathbf{v}|| = |c| \; ||\mathbf{v}||$ for every $\mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ (Homogeneity).

The following theorem connects the two notions of inner product and norm.

Theorem 4.1. *If $\langle , \rangle$ is an inner product on $\mathbb{R}^n$, then the function defined by $||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ is a norm.*

Example 4.1. The following are examples of norms on $\mathbb{R}^n$.

(a) $||(x_1, x_2, \ldots, x_n)|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

(b) $||(x_1, x_2, \ldots, x_n)|| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$.

Remark. When a particular norm of $\mathbb{R}^n$ is not specified we use the Euclidean norm, stated in the example above.

4.2 Linear Transformations and Matrices

YouTube Video: https://youtu.be/WA12KFfJ-Uk

Remark. All vector spaces are subspaces of $\mathbb{R}^n$ for some positive integer n.

Definition 4.2. Let V, W be two vector spaces. (i.e. V is a subspace of $\mathbb{R}^m$ and W is a subspace of $\mathbb{R}^n$ for some positive integers m, n.) A function $L : V \to W$ is said to be **linear** (or a **linear transformation**) if for all $\mathbf{v}, \mathbf{w} \in V$ and $c \in \mathbb{R}$,

- $L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$ (Additivity), and
- $L(c\mathbf{v}) = cL(\mathbf{v})$ (Homogeneity)

Example 4.2. Determine which of the following functions are linear:

(a) $f : \mathbb{R} \to \mathbb{R}, f(x) = cx$, where c is a constant.

(b) $g : \mathbb{R}^2 \to \mathbb{R}, g(x, y) = 2x + 3y$.

(c) $h : \mathbb{R} \to \mathbb{R}, h(x) = 2x + 3$.

(d) $k : \mathbb{R}^n \to \mathbb{R}$, $k(\mathbf{v}) = \langle \mathbf{w}, \mathbf{v} \rangle$, where $\mathbf{w}$ is a fixed vector and $\langle , \rangle$ in an inner product of $\mathbb{R}^n$.

Theorem 4.2. *Let $L : V \to W$ be a function between vector spaces. Then, the following are equivalent.*

(a) L is linear.

(b) $L(\mathbf{u} + c\mathbf{v}) = L(\mathbf{u}) + cL(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ and all $c \in \mathbb{R}$.

(c) $L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ and all $a, b \in \mathbb{R}$.

Example 4.3. Prove that all linear transformations $f : \mathbb{R}^n \to \mathbb{R}$ are given by $f(\mathbf{v}) = \mathbf{w} \cdot \mathbf{v}$, where $\mathbf{w}$ is a fixed vector in $\mathbb{R}^n$.

Example 4.4. Identify all linear transformations $f : \mathbb{R}^3 \to \mathbb{R}^2$.

Example 4.5. Prove that all linear transformations $f : \mathbb{R}^n \to \mathbb{R}^m$ are given by

$$f(\mathbf{v}) = \begin{pmatrix} \mathbf{w}_1 \cdot \mathbf{v} \\ \vdots \\ \mathbf{w}_m \cdot \mathbf{v} \end{pmatrix},$$

where $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \mathbb{R}^n$ are fixed vectors.

Definition 4.3. Given an $m \times n$ matrix

$$A = \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{pmatrix},$$

where $\mathbf{w}_j$'s are rows of A, and given a column vector $\mathbf{v} \in \mathbb{R}^n$. The product of A and $\mathbf{v}$, denoted by $A\mathbf{v}$, is defined by:

$$A\mathbf{v} = \begin{pmatrix} \mathbf{w}_1 \cdot \mathbf{v} \\ \vdots \\ \mathbf{w}_m \cdot \mathbf{v} \end{pmatrix}.$$

YouTube Video: https://youtu.be/SIClX9f9Bp0

Theorem 4.3. *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if there is an $m \times n$ matrix A for which $f(\mathbf{v}) = A\mathbf{v}$. Furthermore, for every given linear transformation f the matrix A is unique, and the columns of A are given by $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$. In other words,*

$$A = (f(\mathbf{e}_1) \cdots f(\mathbf{e}_n)).$$

Definition 4.4. The matrix A of the linear transformation f in theorem above is called the matrix of f with respect to the standard basis and is denoted by M_f.

Example 4.6. Let α be an angle. Suppose $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation with angle α about the origin. From geometry we know R_α is linear. Find the matrix of R_α with respect to the standard basis.

Remark. The set of $m \times n$ matrices with real entries is denoted by $M_{m\times n}(\mathbb{R})$. The set of (square) $n \times n$ matrices with real entries is denoted by $M_n(\mathbb{R})$.

Definition 4.5. Let $A \in M_{m\times n}(\mathbb{R})$ and $B \in M_{n\times k}(\mathbb{R})$ matrix. The matrix AB is an $m \times k$ matrix whose j-th column is obtained from multiplying A by the j-th column of B. In other words, the (i, j) entry of AB is obtained by finding the dot product of the i-th row of A with the j-th column of B.

Remark. Note that to be able to evaluate the multiplication AB of two matrices A and B, we need the number of columns of A to be the same as the number of rows of B.

Example 4.7. Evaluate the matrices AB and BA, where

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 0 & -1 \\ 2 & 3 \\ 5 & -1 \end{pmatrix}.$$

Example 4.8. Consider a 2×3 matrix A and a vector $\mathbf{v}$ as follows:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \text{ and } \mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Show that $A\mathbf{v}$ is the following linear combination of columns of A:

$$A\mathbf{v} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}.$$

Remark. For every $m \times n$ matrix A and every column vector $\mathbf{v} \in \mathbb{R}^n$ the vector $A\mathbf{v}$ is a linear combination of columns of A with coefficients from entries of $\mathbf{v}$.

Theorem 4.4. *If the functions $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^k$ are linear, then $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ is also linear and $M_{g\circ f} = M_g M_f$.*

Proof. The part that $g \circ f$ is linear is left as an exercise. We know the j-th column of the matrix of $g \circ f$ is $g \circ f(\mathbf{e}_j)$. This equals $g(f(\mathbf{e}_j)) = M_g f(\mathbf{e}_j)$. Since the j-th column of M_f is $f(\mathbf{e}_j)$, the j-th column of $M_g M_f$ is precisely $M_g f(\mathbf{e}_j)$. Therefore, the j-th column of $M_g M_f$ is precisely $g \circ f(\mathbf{e}_j)$. Therefore, the matrix of $M_{g \circ f}$ in standard basis is $M_g M_f$, as desired. $\square$

Example 4.9. The matrix of the identity function $I : \mathbb{R}^n \to \mathbb{R}^n$ defined by $I(\mathbf{x}) = \mathbf{x}$ is given by

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

The matrix above is called the **identity** matrix of size n and is denoted by I_n.

Definition 4.6. The **sum** of two matrices A, B is defined precisely when they have the same size. In which case, the addition is entry-wise. In other words, the (i, j) entry of $A + B$ is the sum of (i, j) entries of A and B. Any $m \times n$ matrix A can be multiplied by any scalar c. The result is also an $m \times n$ matrix, whose (i, j) entry is c times the (i, j) entry of A for every i, j.

Theorem 4.5. *For matrices A, B, C and a real number r we have the following:*

(a) $A(BC) = (AB)C$. (Associativity).

(b) $A(B + C) = AB + AC$, and $(B + C)A = BA + CA$. (Distributivity).

(c) $AI_n = A$, and $I_m A = A$. (Multiplicative Identity)

(d) $r(AB) = (rA)B = A(rB)$.

Provided that in each case the appropriate multiplication or addition is defined.

Note that in general for two matrices A, B the products AB and BA are not equal, even if both of them are defined.

Definition 4.7. We say two matrices A, B **commute** if $AB = BA$.

4.3 Kernel and Image

YouTube Video: https://youtu.be/pmU5t_1G1g8

Definition 4.8. Given a linear transformation $L : V \to W$, the **kernel** of L is defined as Ker $L = L^{-1}(\{\mathbf{0}\})$. In other words,

$$\text{Ker } L = \{\mathbf{v} \in V \mid L(\mathbf{v}) = \mathbf{0}\}.$$

The **image** (or range) of L is defined as

$$\text{Im } L = \{\mathbf{w} \in W \mid \mathbf{w} = L(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}.$$

Theorem 4.6. *Let $L : V \to W$ be a linear transformation of vector spaces. Then* Ker L *is a subspace of* V *and* Im L *is a subspace of* W.

Proof. We will use the subspace criterion for both.

First, note that $L(\mathbf{0}) = L(0\mathbf{0}) = 0L(\mathbf{0}) = \mathbf{0}$, by homogeneity. Thus, $L(\mathbf{0}) = \mathbf{0}$. Therefore, $\mathbf{0} \in \text{Ker } L$. Now, assume $\mathbf{x}, \mathbf{y} \in \text{Ker } L$, and $c \in \mathbb{R}$. By definition, $L(\mathbf{x}) = L(\mathbf{y}) = \mathbf{0}$. By linearity we have

$$L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}) = \mathbf{0} + \mathbf{0} = \mathbf{0}, \text{ and } L(c\mathbf{x}) = cL(\mathbf{x}) = c\mathbf{0} = \mathbf{0} \Rightarrow \mathbf{x} + \mathbf{y}, c\mathbf{x} \in \text{Ker } L.$$

Therefore, Ker L is a subspace of V.

Since $L(\mathbf{0}) = \mathbf{0}$, the zero vector of W is in Im L. Assume $\mathbf{v}, \mathbf{w} \in \text{Im } L$. Thus, there are vectors $\mathbf{x}, \mathbf{y} \in V$ for which $\mathbf{v} = L(\mathbf{x})$ and $\mathbf{w} = L(\mathbf{y})$. Given a scalar $c \in \mathbb{R}$ we have

$$\mathbf{v} + \mathbf{w} = L(\mathbf{x}) + L(\mathbf{y}) = L(\mathbf{x} + \mathbf{y}) \in \text{Im } L, \text{ and } c\mathbf{v} = cL(\mathbf{x}) = L(c\mathbf{x}) \in \text{Im } L.$$

Therefore, Im L is a subspace of W. □

The last part of the theorem above can be generalized as follows:

Example 4.10. Find the kernel and image of the linear transformation $L : \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$L(x, y, z) = (x + 2y + z, 2x - y - z).$$

Theorem 4.7. *Suppose $L : V \to W$ is a linear transformation of vector spaces. Then,* Ker $L = \{\mathbf{0}\}$ *if and only if L is one-to-one. If L is one-to-one, then* $\dim \text{Im } L = \dim V$.

Theorem 4.8. *Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation whose matrix in the standard basis is A. Then,*

(a) Im $L = \text{Col}(A)$.

(b) Ker $L = (\text{Row}(A))^\perp$.

(c) $\dim \text{Ker } L + \dim \text{Im } L = n$.

Theorem 4.9 (Rank-Nullity Theorem)**.** *Let V and W be vector spaces, and $L : V \to W$ be a linear transformation. Then,*

$$\dim \text{Ker } L + \dim \text{Im } L = \dim V.$$

4.4 More Examples

Example 4.11. Prove that if $||\cdot||$ is a norm on $\mathbb{R}^n$, then $||\mathbf{0}|| = 0$.

Solution. By homogeneity $||0\mathbf{0}|| = |0| \ ||\mathbf{0}|| = 0||\mathbf{0}|| = 0$. Since $0\mathbf{0} = \mathbf{0}$, we obtain $||\mathbf{0}|| = 0$, as desired. □

Example 4.12. Prove that if $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are vectors in $\mathbb{R}^n$ with a norm $||\cdot||$, then

$$||\mathbf{v}_1 + \cdots + \mathbf{v}_m|| \leq ||\mathbf{v}_1|| + \cdots + ||\mathbf{v}_m||.$$

Solution. We will prove this by induction on m.

Basis step: For $m = 1$ both sides of the inequality are $||\mathbf{v}_1||$. This proves the basis step.

Inductive Step: Let $\mathbf{v}_1, \ldots, \mathbf{v}_{m+1}$ be vectors in $\mathbb{R}^n$. Suppose

$$||\mathbf{v}_1 + \cdots + \mathbf{v}_m|| \leq ||\mathbf{v}_1|| + \cdots + ||\mathbf{v}_m|| \qquad (*)$$

By the Triangle Inequality we obtain:

$$||\mathbf{v}_1 + \cdots + \mathbf{v}_{m+1}|| \leq ||\mathbf{v}_1 + \cdots + \mathbf{v}_m|| + ||\mathbf{v}_{m+1}||.$$

Combining this with $(*)$ completes the inductive step. □

Example 4.13. Find Ker L, and Im L, where $L : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$L(\mathbf{v}) = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -1 & 0 \end{pmatrix} \mathbf{v}.$$

Solution. We row reduce the given matrix to obtain.

$$\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix}$$

This means the first two columns of the original matrix are linearly independent. Therefore, the image is 2-dimensional. Since the image is a subspace of $\mathbb{R}^2$, by Exercise 3.5 the image must be equal to $\mathbb{R}^2$.

For the kernel, we must solve the system

$$\begin{cases} x - 3z &= 0 \\ y + 3z &= 0 \end{cases}$$

This gives us $x = 3z$, and $y = -3z$. Therefore,

$$\text{Ker } L = \{(3z, -3z, z) \mid z \in \mathbb{R}\} = \text{span } \{(3, -3, 1)\}.$$

□

Example 4.14. Find the kernel and image of the rotation function $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$.

Solution. For kernel, suppose $R_\alpha(\mathbf{v}) = \mathbf{0}$ for some $\mathbf{v} \in \mathbb{R}^2$. But that means if we rotate $\mathbf{0}$ with angle $-\alpha$ we should obtain the vector $\mathbf{v}$, and thus $\mathbf{v} = \mathbf{0}$. The zero vector is in Ker R_α. Therefore, Ker $R_\alpha = \{\mathbf{0}\}$.

By the Rank-Nullity Theorem $\dim \text{Ker } R_\alpha + \dim \text{Im } R_\alpha = 2$. Thus $\dim \text{Im } R_\alpha = 2$, and since Im R_α is a subspace of $\mathbb{R}^2$ and $\dim \mathbb{R}^2 = 2$, we conclude that Im $R_\alpha = \mathbb{R}^2$, as desired. □

Example 4.15. Let $L : V \to W$ be a linear transformation. Prove that $L(\mathbf{0}) = \mathbf{0}$.

Solution. By linearity we have $L(\mathbf{0}) = L(0\mathbf{0}) = 0L(\mathbf{0}) = \mathbf{0}$, since the product of 0 and any vector is $\mathbf{0}$. $\square$

Example 4.16. Prove Theorem 4.2.

Solution. $(a) \Rightarrow (b)$: Assume L is linear. By additivity and homogeneity we have

$$\begin{aligned} L(\mathbf{u} + c\mathbf{v}) &= L(\mathbf{u}) + L(c\mathbf{v}) && \text{Additivity} \\ &= L(\mathbf{u}) + cL(\mathbf{v}) && \text{Homogeneity} \end{aligned}$$

$(b) \Rightarrow (c)$: Assume L satisfies (b), and let $\mathbf{u}, \mathbf{v} \in V, a, b \in \mathbb{R}$. Applying (b) to vectors $a\mathbf{u}, \mathbf{v}$ and the scalar b we obtain the following:

$$\begin{aligned} L(a\mathbf{u} + b\mathbf{v}) &= L(a\mathbf{u}) + bL(\mathbf{v}) \\ &= L(\mathbf{0} + a\mathbf{u}) + bL(\mathbf{v}) \\ &= L(\mathbf{0}) + aL(\mathbf{u}) + bL(\mathbf{v}) \end{aligned}$$

On the other hand if we set $a = b = 0$ in (b), we obtain $L(\mathbf{0}) = \mathbf{0}$. Thus, $L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v})$, which proves (c).

$(c) \Rightarrow (a)$: Letting $a = b = 1$ in (c) we obtain $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$, which is precisely the additivity. Letting $a = 0$ in (c), we obtain $L(0\mathbf{u} + b\mathbf{v}) = 0L(\mathbf{u}) + bL(\mathbf{v})$, which implies $L(b\mathbf{v}) = bL(\mathbf{v})$, which is the homogeneity. Therefore, L is linear. $\square$

Example 4.17. Let c be a scalar, A, B be two matrices with real entries and $\mathbf{v}$ be a column vector. Prove or disprove each of the following:

(a) If $c\mathbf{v} = \mathbf{0}$, then $c = 0$ or $\mathbf{v} = \mathbf{0}$.

(b) If $AB = 0$, then $A = 0$ or $B = 0$.

(c) If $A\mathbf{v} = \mathbf{0}$, then $A = 0$ or $\mathbf{v} = \mathbf{0}$.

Solution. (a) This is true. Suppose $c\mathbf{v} = 0$. If $c \neq 0$, then multiplying both sides by $1/c$ we obtain $1\mathbf{v} = \mathbf{0}$ and thus $\mathbf{v} = \mathbf{0}$.

(b) This is false. Consider the two matrices $A = (1\ 0)$ and $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Neither A nor B is zero, but $AB = 0$.

(c) This is false. The same example as the one in part (b) works. $\square$

Example 4.18. Prove that if a, b are real numbers with $b \neq 0$, then the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = ax + b$ is not linear.

Solution. There are multiple ways of doing this. One way would be to note $f(0) = b \neq 0$ and thus f cannot be linear by Example 4.15.

Another way would be to use Theorem 4.3: If f were linear, then there would be a 1×1 matrix A for which $f(x) = Ax$. Note that 1×1 matrices are just real numbers. Thus, we must have $ax + b = Ax$ for all $x \in R$ and thus $b = 0$, which is a contradiction.

We could also check that such f does not satisfy the homogeneity (or the additivity) condition. For example $f(2) = 2a + b \neq 2f(1) = 2a + 2b$, since $b \neq 0$. $\square$

Example 4.19. Determine if each of the following functions is linear:

(a) $L : \mathbb{R}^2 \to \mathbb{R}$, given by $L(x, y) = xy$.

(b) $L : \mathbb{R}^2 \to \mathbb{R}$, given by $L(x, y) = x + 3y$.

(c) $L : \mathbb{R}^n \to \mathbb{R}$, given by $L(x_1, \ldots, x_n) = x_1$.

Solution. (a) This is not linear. Note that $f(1, 0) = 0$, $f(0, 1) = 0$, and $f(1, 1) = 1$. This means $f(1, 1) \neq f(1, 0) + f(0, 1)$, which implies f is not additive and thus it is not linear.

(b) This is linear. We have

$$L(x, y) = (1\ 3)\begin{pmatrix} x \\ y \end{pmatrix}.$$

By Theorem 4.3 this function is linear.

(c) This is linear using Theorem 4.3 and the following:

$$L(x_1, \ldots, x_n) = (1\ 0\ \cdots\ 0)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

$\square$

Example 4.20. Let V, W be two vector spaces, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be a basis for V. Assume $S, T : V \to W$ are linear transformations. Prove that $S = T$ if and only if $S(\mathbf{v}_j) = T(\mathbf{v}_j)$ for $j = 1, \ldots, n$.

Solution. $\Rightarrow$: If $S = T$, then $S(\mathbf{v}_j) = T(\mathbf{v}_j)$, by definition of equality of two functions.

$\Leftarrow$: Suppose $S(\mathbf{v}_j) = T(\mathbf{v}_j)$ for $j = 1, \ldots, n$. Let $\mathbf{v} \in V$. Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis for V, there are scalars $c_1, c_2, \ldots, c_n$ for which $\mathbf{v} = \sum_{j=1}^{n} c_j \mathbf{v}_j$. By linearity of S and T, and the fact that $S(\mathbf{v}_j) = T(\mathbf{v}_j)$ we have

$$S(\mathbf{v}) = S(\sum_{j=1}^{n} c_j \mathbf{v}_j) = \sum_{j=1}^{n} c_j S(\mathbf{v}_j) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j) = T(\sum_{j=1}^{n} c_j \mathbf{v}_j) = T(\mathbf{v}).$$

Note also that S and T have the same domain and co-domain. Therefore $S = T$, as desired. $\square$

Example 4.21. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation for which $T(1,2) = (1,0,1)$ and $T(2,1) = (1,1,0)$. Find the matrix M_T.

Solution. We need to find $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$. We see

$$(1,0) = \frac{2}{3}(2,1) - \frac{1}{3}(1,2), \text{ and } (0,1) = \frac{2}{3}(1,2) - \frac{1}{3}(2,1).$$

By linearity of T we have

$$T(\mathbf{e}_1) = \frac{2}{3}T(2,1) - \frac{1}{3}T(1,2) = \frac{2}{3}(1,1,0) - \frac{1}{3}(1,0,1) = (1/3, 2/3, -1/3),$$

and

$$T(\mathbf{e}_2) = \frac{2}{3}T(1,2) - \frac{1}{3}T(2,1) = \frac{2}{3}(1,0,1) - \frac{1}{3}(1,1,0) = (1/3, -1/3, 2/3).$$

Therefore, by a theorem the matrix M_T is given by

$$M_T = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{pmatrix}$$

$\square$

4.5 Exercises

Exercise 4.1. *Determine if each of the following is a linear transformation. If it is linear, provide a proof. If it is not, by an example prove that it fails to satisfy one of the conditions of linear transformation.*

(a) $L : \mathbb{R}^3 \to \mathbb{R}^3$, $L(x,y,z) = (x+y, z, x^2)$.

(b) $L : \mathbb{R}^2 \to \mathbb{R}^3$, $L(x,y) = (x+2y, y, -x)$.

(c) $L : \mathbb{R}^3 \to \mathbb{R}^2$, $L(x,y,z) = (xy, xz)$.

(d) $L : \mathbb{R}^3 \to \mathbb{R}^2$, $L(x,y,z) = (x+y, z-1)$.

Exercise 4.2. *Find all linear transformations* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *satisfying all of the following:*

$$T(1,2,0) = (0,2), T(-1,1,1) = (-2,3), \text{ and } T(1,-2,-1) = (1,-3).$$

Exercise 4.3. *Let* $\alpha \in [0, 2\pi)$ *be an angle. Consider the transformation* $T_\alpha : \mathbb{R}^3 \to \mathbb{R}^3$ *which rotates every point around the* z*-axis with angle* α*. Assume we know* T_α *is linear. Find* M_{T_α}*.*

Exercise 4.4. *Find all* 2×2 *matrices* A *that commute with every other matrix. In other words, find all matrices* $A \in M_2(\mathbb{R})$*, for which* $AB = BA$*, for every* $B \in M_2(\mathbb{R})$*.*

Exercise 4.5. *True or false? If true provide a proof, and if false provide a counter-example.*

(a) If for a square matrix A we have $A^2 = 0$, then $A = 0$.

(b) If the two products AB and BA are defined, then A and B must be square matrices.

(c) $AB = BA$ for every two 2×2 matrices A and B

Exercise 4.6. *Find an example of three matrices A, B, C for which $AB = BA, AC = CA$, but $BC \neq CB$.*

Exercise 4.7. *Find an example of two matrices A, B for which A^2 and B commute but A and B do not commute.*

Exercise 4.8. *Prove that if two matrices A and B commute, then for every two positive integers m, n the two matrices A^n and B^m also commute.*

Exercise 4.9. *Using the definition of linearity, prove that if $S : V \to W$ and $T : W \to U$ are linear transformations of vector spaces, then $T \circ S : V \to U$ is also linear.*

Exercise 4.10. *Suppose $T : V \to W$ is a linear transformation between vector spaces. Using induction, prove that for every $c_1, \ldots, c_n \in \mathbb{R}$ and every $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, we have*

$$T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n).$$

Exercise 4.11. *Provide an example of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ for some n, m that is homogeneous, but it is not linear.*

Exercise 4.12. *Let V, W be two vector spaces. Suppose $L : V \to W$ is a function that is additive and satisfies $L(c\mathbf{v}) = cL(\mathbf{v})$, for all $\mathbf{v} \in V$ and all positive $c \in \mathbb{R}$. Does L have to be linear?*

Exercise 4.13. *Suppose $T : V \to W$ is a function between vector spaces that is additive. Prove that:*

(a) $T(-\mathbf{v}) = -T(\mathbf{v})$, for all $\mathbf{v} \in V$.

(b) For every positive integer n and every $\mathbf{v} \in V$, $T(n\mathbf{v}) = nT(\mathbf{v})$. (Hint: Use induction on n.)

(c) Combining parts (a) and (b), prove $T(n\mathbf{v}) = nT(\mathbf{v})$ for every $\mathbf{v} \in V$ and every $n \in \mathbb{Z}$.

(d) Prove that for every $r \in \mathbb{Q}$ and every $\mathbf{v} \in V$, we have $T(r\mathbf{v}) = rT(\mathbf{v})$.

Exercise 4.14. *Let V, W be vector spaces. Assume $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis for V, and let $\mathbf{w}_1, \ldots, \mathbf{w}_n \in W$. Prove that $T : V \to W$ defined by*

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n, \text{ for all } c_1, c_2, \ldots, c_n \in \mathbb{R}$$

is a linear transformation.

Exercise 4.15. *Suppose $f, g : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations. Prove that $f + g : \mathbb{R}^n \to \mathbb{R}^m$ is linear, and that $M_{f+g} = M_f + M_g$.*

Exercise 4.16. *Suppose $L : V \to W$ is a bijective linear transformation. Prove that $L^{-1} : W \to V$ is linear.*

Exercise 4.17. *Suppose A, B are matrices of size $m \times n$ and $n \times k$, respectively. Prove that $(AB)^T = B^T A^T$.*

Exercise 4.18. *Suppose $\mathbb{R}^n$ is equipped with a norm. Given every two distinct points $A, B \in \mathbb{R}^n$, prove there is a unique point M on segment AB that is equidistant from A and B. In other words, $||\overrightarrow{AM}|| = ||\overrightarrow{BM}||$.*

Definition 4.9. The point M in the above exercise is called the **midpoint** of segment AB.

Definition 4.10. Consider $\mathbb{R}^n$ equipped with a norm. Given a triangle ABC in $\mathbb{R}^n$, the segment connecting A and the midpoint of BC is called the **median** of ABC correspoding to vertex A.

Exercise 4.19. *Consider $\mathbb{R}^n$ equipped with a norm. Prove the medians of ABC are concurrent. In other words, prove there is a unique point $G \in \mathbb{R}^n$ for which G lies on all the three medians corresponding to vertices of triangle ABC.*

Exercise 4.20. *Suppose $\langle , \rangle$ is an inner product of $\mathbb{R}^n$. Prove that for every $\mathbf{w} \in \mathbb{R}^n$, there is a vector $\mathbf{x} \in \mathbb{R}^n$ for which $\langle \mathbf{w}, \mathbf{v} \rangle = \mathbf{x} \cdot \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n$.*

Exercise 4.21. *We know composition of every two linear transformations is linear. Is the converse true? In other words, suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^k$ are functions for which $g \circ f$ is linear. Can we conclude f and g are linear?*

Definition 4.11. A **distance** on $\mathbb{R}^n$ is a function d that assigns a nonnegative real number $d(\mathbf{v}, \mathbf{w})$ to every pair of points $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ that satisfies the following for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

(a) $d(\mathbf{v}, \mathbf{w}) > 0$ if $\mathbf{v} \neq \mathbf{w}$ (Positivity),

(b) $d(\mathbf{v}, \mathbf{w}) = d(\mathbf{w}, \mathbf{v})$ (Symmetry),

(c) $d(\mathbf{v}, \mathbf{w}) + d(\mathbf{w}, \mathbf{u}) \geq d(\mathbf{v}, \mathbf{u})$ (Triangle ineuqality).

Exercise 4.22. *Prove that any norm on $\mathbb{R}^n$ provides a distance d between every two $\mathbf{v}, \mathbf{w} \in V$ defined by:*

$$d(\mathbf{v}, \mathbf{w}) = |\mathbf{v} - \mathbf{w}|.$$

4.6 Challenge Problems

Exercise 4.23. *Prove that for every $n \geq 2$ there is no inner product on $\mathbb{R}^n$ that gives us the norm*

$$||(x_1, \ldots, x_n)|| = \max\{|x_1|, \ldots, |x_n|\}.$$

Exercise 4.24. *Prove that for every $n \geq 2$ there is no inner product on $\mathbb{R}^n$ that gives us the norm*

$$||(x_1, \ldots, x_n)|| = |x_1| + \cdots + |x_n|.$$

Definition 4.12. For two subspaces U and W of $\mathbb{R}^n$, define

$$U + W = \{\mathbf{x} \in V \mid \mathbf{x} = \mathbf{u} + \mathbf{w} \text{ for some } \mathbf{u} \in U, \text{ and } \mathbf{w} \in W\}.$$

Exercise 4.25. *Suppose U and W are subspaces of a vector space V. Prove that*

$$\dim(U+W)+\dim(U\cap W)=\dim U+\dim W.$$

Exercise 4.26. *Suppose $A \in M_n(\mathbb{R})$ is not invertible. Prove that there is a nonzero matrix $B \in M_n(\mathbb{R})$ for which $AB = BA = 0$.*

Exercise 4.27. *Suppose $L : \mathbb{R}^2 \to \mathbb{R}^2$ is a bijective linear transformation. Prove that the image of every ellipse is also an ellipse. What can we say about images of parabolas and hyperbolas?*

4.7 Summary

- To check if $||\mathbf{v}||$ is a norm, we need to check if it satisfies three properties: Positivity, Triangle Inequality, and Homogeneity.
- To prove $L : V \to W$ is linear we need to prove two properties for all $\mathbf{v}, \mathbf{w} \in V$ and $c \in \mathbb{R}$:
  - Additivity: $L(\mathbf{v}+\mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$, and
  - Homogeneity: $L(c\mathbf{v}) = cL(\mathbf{v})$
- To prove $L : V \to W$ is not linear, we need to show either the additivity or the homogeneity fails for at least *some* vectors and constants. We do not need to prove that both additivity and homogeneity fail.
- To find the product $A\mathbf{v}$, where A is an $m \times n$ matrix and $\mathbf{v}$ is an $n \times 1$ column we can use one of the following:
  - Using rows of A, write: $A = \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{pmatrix}$. Then we have $A\mathbf{v} = \begin{pmatrix} \mathbf{w}_1 \cdot \mathbf{v} \\ \vdots \\ \mathbf{w}_m \cdot \mathbf{v} \end{pmatrix}$.
  - Using columns of A, write: $A = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Then,

    $$A\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n.$$
- Every linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ is given by $L(\mathbf{v}) = A\mathbf{v}$, where A is an $m \times n$ matrix. The columns of the matrix A are $L(\mathbf{e}_1), \dots, L(\mathbf{e}_n)$. Every such function is linear.
- The (i, j) entry of AB is obtained by finding the dot product of the i-th row of A and the j-th column of B.
- For the matrix AB to be defined we need the number of columns of A and the number of rows of B to be the same.
- If A is a matrix of size $m \times n$ and B is a matrix of size $n \times k$, then the matrix AB is of size $m \times k$.

- Note that in general AB and BA are not the same matrices.

- Image and kernel of every linear transformation are subspaces.

- The Rank-Nullity Theorem states that for every linear transformation $L : V \to W$ we have

$$\dim \operatorname{Ker} L + \dim \operatorname{Im} L = \dim V.$$