MATH 340: Multivariable Calculus, Linear Algebra, and Differential Equations I

Omkar Pathak*

January 21, 2023

These are my notes for UMD's MATH 340, which covers proof-based multivariable calculus, linear algebra, and differential equations. These notes are taken live in class ("live-TEX"-ed). MATH 340 is the first course in a two-course honors sequence (340/341). As a reference, I may use Apostol's *Calculus, Volume II*, Axler's *Linear Algebra Done Right*, as well as last year's MATH 340 notes and Oliver Knill's Harvard MATH 22a notes. This course is taught by Prof. Roohollah Ebrahimian.

Contents

*Email: omkarp07@terpmail.umd.edu

§1 Monday, August 29, 2022

§1.1 Sets

Today, we discussed sets and covered two main definitions.

Definition 1.1. Let A and B be two sets for which the statement

$$\text{“If } x \in A \text{, then } x \in B\text{”}$$

is true. Then, we say that A is a **subset** of B, win 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 $A \subsetneqq Q$. The **union** of A and B, denoted by $A \cap B$ is the set consisting of all elements x that are in A or B (or both, because the word “or” is not exclusive). The **intersection** of A and B, denoted by $A \cup B$, and is the set consisting of all elements that are in both A and B. In other words,

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

The union and intersection of n sets is defined similarly:

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

Definition 1.2. We say that 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 1.3

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

Proof. By the definition of set equality, we must prove that $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$ and $(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$. We will start with the former.

By the definition of subset, we want to prove that if $x \in (A \cup B) \cap C)$, $x \in (A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$. Note that x is any general element of $(A \cup B) \cap C$. If $x \in (A \cup B) \cap C$, we have that $x \in (A \cup B)$ AND $x \in C$. Now, because $x \in (A \cup B)$, we have $x \in A$ OR $x \in B$. We can split this into two cases. In the first case, $x \in A$. This would mean that $x \in (A \cap C)$, because $x \in C$, as mentioned before. In the second case, $x \in B$, implying that $x \in B \cap C$. Therefore, because we have $x \in A$ OR $x \in B$, we have that $x \in (A \cap C) \cup (B \cap C)$.

Now, we'll prove that if $x \in (A \cap C) \cup (B \cap C)$, $x \in (A \cap B) \cup C$. If $x \in (A \cap C) \cup (B \cap C)$, we have $x \in (A \cap C)$ OR $x \in (B \cap C)$. We can also split this into two cases. If $x \in (A \cap C)$, $x \in A$ AND $x \in C$. If $x \in (B \cap C)$, $x \in B$ AND $x \in C$. Either way, $x \in C$. Therefore, x is either in A or B (or both) AND in C. As a result, we have that $x \in (A \cup B) \cap C$. $\square$

Definition 1.4. An **ordered pair** (a, b) of objects a and 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 any only if $a_i = b_i$ for all i.

Definition 1.5. The **Cartesian product** of n sets $A_1, A_2, \ldots, A_n$ is denoted by $A_1 \times A_2 \times \ldots \times A_n$ 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 set A is denoted by A^n.

Example 1.6

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

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}$$

Definition 1.7. 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 **pairwise disjoint** if for every $i \neq j$, A_i and A_j are disjoint.

Example 1.8

$\{1, 2\}, \{1, 3\}$, and $\{2, 3\}$ are **not** pairwise disjoint, because every set shares an element with two other sets. However, all three sets **are** disjoint because there is no common element between all sets. Note that if there sets are pairwise disjoint

Venn diagrams are also a great way to visualize and understand sets. However, since I (at the moment) do not know how to draw them in LaTeX, I'll hold off on drawing them for now.

§2 Wednesday, August 31, 2022

Definition 2.1. If A is a subset of the universal set U, the **complement** of A in U is the set consisting of all elements that are **not in** A, e.g. A^c.

Theorem 2.2 (DeMorgan's Laws)

DeMorgan's Laws state the following:

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

Let's now prove DeMorgan's Laws.

Proof. Recall that to prove two sets are equal, the sets must be subsets of each other, and that to prove that the sets are subsets of each other, we must prove that any arbitrary x being in one set must be in another. For the first part, let $x \in \left(\bigcap_{j=1}^{n} A_j\right)^c$. By the definition of a complement, $x \notin \bigcap_{j-1}^{n} A_j$. Because x is not in the intersection of the A_j's, we can say that there is some j $(1 \leq j \leq n)$ such that $\notin A_j$. Since $x \notin A_j$, $x \in A_j^c$. By the definition of the union, we have that $x \in \bigcup_{j=1}^{n} A_j^c$.

Now, we must prove that $x \in \bigcup_{i=1}^{n} A_j^c$ implies $x \in (\bigcap_{i=1}^{n} A_j)^c$. If $x \in \bigcup_{j=1}^{n} A_j^c$, we have that x is in one or multiple A_i^c's. This means that $x \notin$ one or multiple A_is, implying that $x \notin \bigcup_{i=1}^{n} A_i)$. Therefore, by the definition of the complement, we have that $x \in \left(\bigcap_{i=1}^{n} A_j^c\right)$. ☐

Proving part 2 is similar, and is left as an exercise to the reader.

Let's now talk about functions.

§2.1 Functions

Definition 2.3. Given two nonempty sets A and B, a **function** or **mapping** $f : A \to B$ is a rule that assigns every element $a \in A$ to 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, 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 must have $a_1 = a_2$. f is called **bijective** if it is surjective and injective.

Example 2.4

Show $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^3 + 2$ is bijective.

Let $b \in \mathbb{R}$. We want to show that $R_f = \mathbb{R}$. We also want to show that $b = f(a)$ for some $a \in \mathbb{R}$. We have that $b = a^3 + 2$. This means $a = \sqrt[3]{b-2}$. Therefore, we have $f = \sqrt[3]{b-2}) = \sqrt[3]{(b-2)^3} + 2 = b - 2 + 2 = b$. Thus, we have mapped an arbitrary real number to another real number, and have shown f is onto. To prove that f is injective, suppose $f(a_1) = f(a_2)$. Thus, we have $a_1^3+2 = a_2^3+2 \to a_1^3 = a_2^3 \to a_1 = a_2$. Thus, f is one-to-one/injective. As a result, we've proved that f is bijective.

Remark 2.5. We don't have to show that things like $\sqrt{}$, $\sqrt[3]{}$, etc. are real. We can assume certain properties of the real numbers.

The function $\text{id}_A : A \to A$ is 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 $f \circ g = \mathrm{id}_B$. The function g is called the **inverse** of f and is denoted by f^{-1}. Note: for a composition $f \circ g$ to be legitimate, we need $R_g \subseteq D_f$. Furthermore, for a function to have an inverse, we need the function to be surjective and injective, e.g. bijective.

Example 2.6

The function $f : A \times B \to A$, defined by $f(a, b) = a$ is called the **projection** onto the first component. Note that pi_1 is also equivalent to $A \times B \to A$. Similarly, the function $f : A_1 \times A_2 \times \cdots \times A_n \to A_i$, defined by $f(a_1, \ldots, a_n) = a_i$ is called the **projection** onto the i-th component. We can also use π to refer to a projection with multiple elements, with $pi_1 = A_1 \times A_2 \cdots \times A_n \to A_i$.

Definition 2.7. 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 both the pre-image and image are both sets.

Example 2.8

Say we have a domain $\{0, 1, 2, 3, 4\}$ and a co-domain $\{1, 2, 3, -1, 5, 0\}$, with a mapping $f(\{0, 1, 2\} = \{f(0), f(1), f(2)\} = \{1, 2\}$. Thus, the image of $\{0, 1, 2\}$ is $\{1, 2\}$. If we have $f\{0, 3, 4\} = \{1, 3, -1\}$, we say that the pre-image of $\{1, 3, -1\}$ under f is $\{0, 2, 4\}$. Note that f is **not** invertible; we use f^{-1} for the pre-image and invertible function, so you must state clearly what an inverse is referring to.

Example 2.9

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.

We have that $f(a, b) = b$. By the definition of the pre-image, we have that the pre-image of $\{b\}$ under f is $f^{-1}(B) = \{a \in A \mid f(a) \in B\}$. By the definition of f, we have that the pre-image is (a, b) such that $a \in A$ and $b \in B$. Formally, we write this as $f^{-1}(\{b\}) = \{(x, y) \in A \times B \mid f(x, y) \in B\} = \{(x, y) \in A \times B \mid y \in \{b\} = \{(x, y) \in A \times B \mid y = b\} = \boxed{A \times \{b\}}$.

Example 2.10

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

Note that $b = 2x + 3y$. Therefore, $f^{-1}(\{b\}) = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) \in B\} = \{(x, y) \in \mathbb{R}^2 \mid 2x + 3y \in B\}$. For any arbitrary b, if $2x + 3y = b$, we have $x = \frac{b-3y}{2}$ and

$y = \frac{b-2x}{3}$. This means that our pre-image is $\boxed{\left(\dfrac{b-3y}{2}, \dfrac{b-2x}{2}\right)}$. As b changes, our pre-image will change accordingly, based on the above coordinates.

Remark 2.11. We can map an ordered triple to an ordered pair, but it isn't very common. Furthermore, we can map an element onto itself, which is simply the identity function id.

Theorem 2.12

Suppose $f : A \to B$ is a function, $S \subseteq A$, and $T_i \subseteq B$ for $i = 1, \ldots, n$. Then,

1. $S \subseteq f^{-1}(f(S))$, and $f(f^{-1}(T_i)) \subseteq T_i$
2. $f^{-1}\left(\bigcup_{i=1}^n T_i\right) = \bigcup_{i=1}^n f^{-1}(T_i)$
3. $f^{-1}\left(\bigcup_{i=1}^n T_i\right) = \bigcup_{i=1}^n f^{-1}(T_i)$

Let's prove Part 1:

Proof. Let $x \in S$. If $x \in S$, we have $x \in A$, because $S \subseteq A$. Because, $f^{-1}(f(S))$ is the pre-image of S under f, e.g. $f^{-1}(f(S)) = \{a \in f \mid f(a) \in S\}$. The definition of the image of S under f is the set $\{f(s) \mid s \in S\}$. Notice that the image of S under f is equivalent to $f(S)$ (e.g. the set of all values of f with $s \in S$). Also notice that $f(s) \in f(S)$. Therefore, by the definition of the pre-image, $s \in f^{-1}(f(S))$ for all s (because $s \in S$). □

For the second part of Part 1:

Proof. Suppose $x \in f(f^{-1}(T_i))$. This means that $x = f(y)$ for some $y \in f^{-1}(T_i)$. As a result, $f(y) \in T_i$, which implies $x \in T_i$. □

Part 2 (Theorem 2.12):

Proof. $f^{-1}\left(\bigcup_{i=1}^n T_i\right) = \{x \in A \mid f(x) \in \bigcup_{i=1}^n T_i\}$, by the definition of the pre-image. Because $f(x) \in \bigcup_{i=1}^n T_i$, by the definition of the union, we have that $f^{-1}\left(\bigcup_{i=1}^n T_i\right)$ represents the set of all $x \in A$ such that $f(x) \in T_i$, for some T_i, by the definition of the union.

Furthermore, by the definition of the union, we have $\bigcup_{i=1}^n f^{-1}(T_i) = \{x \mid x \in f^{-1}(T_i) \text{ for some } i\}$. By the definition of the pre-image, we have that if $x \in f^{-1}(T_i)$ for some i, $x \in A$ and $f(x) \in T_i$ for some i.

Therefore, if $x \in f^{-1}\left(\bigcup_{i=1}^n T_i\right)$, we have $x \in A$ and $f(x) \in T_i$, for some i, which means that x is also in $\bigcup_{i=1}^i f^{-1}(T_i)$. Similarly, if $x \in \bigcup_{i=1}^n f^{-1}(T_i)$, we have $x \in A$ and $f(x) \in T_i$, for some i, which implies that x is also in $f^{-1}\left(\bigcup_{i=1}^n T_i\right)$. As a result, we have proved both sets are subsets of each other, and are done. □

Part 3 (Theorem 2.12):

Proof. This is done in a similar manner to Part 2. By the definition of the pre-image, we have $f^{-1}\left(\bigcap_{i=1}^{n} T_i\right) = \{x \in A \mid f(x) \in \bigcap_{i=1}^{n} T_i\}$. By the definition of the intersection, we have that $f^{-1}\left(\bigcap_{i=1}^{n} T_i\right)$ represents the set of all $x \in A$ such that $f(x) \in T_i$, for all i.

Furthermore, by the definition of the intersection, we have $\bigcap_{i=1}^{n} f^{-1}(T_i) = \{x \mid x \in \bigcap_{i=1}^{n} f^{-1}(T_i)$ for all $i\}$. By the definition of the pre-image, we have that if $x \in \bigcap_{i=1}^{n} f^{-1}(T_i)$ for all i, $x \in A$ and $f(x) \in T_i$ for all i.

Therefore, if $x \in f^{-1}\left(\bigcap_{i=1}^{n} T_i\right)$, $x \in A$ and $f(x) \in T_i$ for all i, which implies that x is also in $\bigcap_{i=1}^{n} f^{-1}(T_i)$. Similarly, if $x \in \bigcap_{i=1}^{n} f^{-1}(T_i)$, we have $x \in A$ and $f(x) \in T_i$, which implies that x is also in $f^{-1}(\bigcap_{i=1}^{n} T_i)$. As a result, we have proved both sets are subsets of each other, and are done.

□

The next section in the notes discuss proof-writing. I will not take notes on this, as I already know the content from previous math classes.

§3 Friday, September 2, 2022

§3.1 Vector Spaces

Definition 3.1. Let V be a nonempty set endowed by two operations, called **vector addition** (denoted with a $+$): $V \times V \to V$ (where $(\mathbf{x}, \mathbf{y})$ is mapped to $\mathbf{x} + \mathbf{y}$, and **scalar multiplication** (denoted with a $\cdot$): $\mathbb{R} \times V \to V$ (where (c, $\mathbf{x}$) is mapped to $c\mathbf{x}$). V is called a vector space if any of the following holds:

- For all $\mathbf{x}$, $\mathbf{y} \in V$, we have $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- For all $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in V$, $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$
- There is an element $\mathbf{e} \in V$ for which for al $\mathbf{x} \in V$, we have $\mathbf{x} + \mathbf{e} = \mathbf{x}$. This element $\mathbf{e}$ is called the **zero vector** and is denoted by $\mathbf{0}$.
- For every $\mathbf{x} \in V$, there is an element $\mathbf{y} \in V$ for which $\mathbf{x} + \mathbf{y} = 0$. This element $\mathbf{y}$ is called the **additive inverse** of $\mathbf{x}$ and is denoted by $-\mathbf{x}$
- For every $\mathbf{x} \in V$, we have $1\mathbf{x}$ $=\mathbf{x}$
- For every $a, b \in \mathbb{R}$ and $\mathbf{x}$, $\mathbf{y}$, $\in V$, we have $a(\mathrm{b}\mathbf{x}) = (ab)\mathbf{x}$ and $(a + b)\mathbf{x} = a\mathbf{x}$ $+b\mathbf{x}$

Example 3.2

Here are some examples of vector spaces:

- $\mathbb{R}^n$, along with the usual scalar multiplication and vector addition
- The set $\mathbb{P}_n$, consisting of all polynomials on one variable t with degree $\leq n$, along with polynomial addition and the usual multiplication of scalars and polynomials.
- The set $C[a, b]$ of all continuous functions $f : [a, b] \to \mathbb{R}$, along with the usual addition of functions and multiplication of scalars and functions. Note that $a < b$, and both a and b are two real numbers.

Theorem 3.3

If $\mathbf{x}$ is an element of a vector space V, then:

- $0\mathbf{x} = \mathbf{0}$. Note that the first zero is a scalar, e.g. an element of $\mathbb{R}$, whereas the second zero is a vector, e.g. an element of V
- $(-1)\mathbf{x} = -\mathbf{x}$

§3.2 $\mathbb{R}^n$ as a Vector Space

As we saw earlier, elements of $\mathbb{R}^n$ are n-tuples of the form $(x_2, x_2, \ldots, x_n)$, where each x_j is a real number. Each one of these elements is a vector; vectors in $\mathbb{R}^n$ can be added component-wise 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 multiplied by a scalar as well:

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

Like any other vector space, $\mathbb{R}^n$ satisfies all six properties of the vector space.

Note that sometimes we refer to elements of $\mathbb{R}^n$ as **points**. This is solely to conceptualize these objects; the math doesn't change. When elements of $\mathbb{R}^n$ are seen as points, the zero vector is sometimes referred to as the **origin**.

Generally, elements of $\mathbb{R}^2$ can be represented by points on a plane. Elements of $\mathbb{R}^3$ can be represented by points in space. To do this, we need three **axes**: the x, y, and z-axes. These three axes must satisfy the **right-hand-rule**. The coordinates of each point on these axes can be found by dropping perpendiculars to the axes. In the (x, y, z) plane, the set of all points with positive coordinates is called the **first octant**.

Remark 3.4. The three axes in the (x, y, z) plane must be **pairwise orthogonal**, or **pairwise perpendicular**.

§3.3 Subspaces

A subset W of a vector space V is called a **subspace** if W is also a vector space (e.g. it has the same operations as V).

Theorem 3.5 (Subspace Criterion)
Let V be a vector space. A subset W of V is a subspace if and only if it satisfies all of the following:

- W is nonempty
- W contains the zero vector of V
- 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 3.6
Considering all elements of $\mathbb{P}_n$ as functions from $\mathbb{R}$ to $\mathbb{R}$, we see that $\mathbb{P}_n$ is a subspace of $C[\mathbb{R}]$.

Theorem 3.7
If W and U are subspaces of V, then $W \cap U$ is also a subspace of V.

Proof. We will use the Subspace Criterion Theorem. First, note that $\mathbf{0}$ belongs to both U and W and thus it is in $W \cap U$. Therefore, $W \cap U$ is not empty. Next, suppose $\mathbf{x}$, $\mathbf{y} \in U \cap W$. By definition of intersection, $\mathbf{x}$ and $\mathbf{y}$ are both in W and U. Since W and U are both subspaces, by the Subspace Criterion Theorem, we have $\mathbf{x} + \mathbf{y} \in W$, $\mathbf{x} + \mathbf{y} \in V$, $c\mathbf{x} \in W$, and $c\mathbf{x} \in U$. Therefore, by the definition of intersection, $\mathbf{x} + \mathbf{y} \in W \cap U$ and $c\mathbf{x} \in W \cap U$, as desired. □

Theorem 3.8
The **length** of a vector $\mathbf{u} = (x, y, z) \in \mathbb{R}^3$ is given by $||\mathbf{u}|| = \sqrt{x^2 + y^2 + z^2}$. Length is also denoted by $||(x, y, z)||$, e.g. $(x, y, z) = \sqrt{x^2 + y^2 + z^2}$.

Proof. Say we want to find the length of a vector (a, b, c). We can drop a perpendicular to the xy plane and use the Pythagorean Theorem to get that the bottom length of the triangle is $\sqrt{a^2 + b^2}$ and that the height of the triangle is c. As a result, we have $||(a, b, c)|| = \sqrt{(\sqrt{a^2 + b^2})^2 + c^2} = \sqrt{a^2 + b^2 + c^2}$. □

The next section of the notes gives warm-up problems, along with more problems and examples. I will handwrite these to get practice with solving problems on paper.

§4 Wednesday, September 7, 2022

§4.1 Linear Dependence, Spanning, and Basis

Definition 4.1. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_n\}$ be a subset of the vector space $\mathbb{R}^n$. We say $\mathbf{w}$ is a **linear combination** of elements in S if $\mathbf{w}= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m$ for some $c_1, c_2, \ldots, c_m \in \mathbb{R}$. By definition, if S is the empty set, then the only linear combination of elements in S is $\mathbf{0}$, e.g. the zero vector.

We note that every vector $\mathbf{v} = (x, y, z)$ in $\mathbb{R}^3$ has essentially three components x, y, and z that are independent of each other. This can also 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" in more abstract vector spaces.

Definition 4.2. We say vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly dependent** if one of these vectors can be written as a linear combination of the others. Otherwise, we way $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly independent**.

Example 4.3

Prove that $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, and $\mathbf{e}_3 = (0, 0, 1)$ are linearly independent.

Proof. We will show that none of these vectors can be written as a linear combination of some of the others. We will proceed with a proof by contradiction. WLOG suppose that $(1, 0, 0)$ can be written as a linear combination of $(0, 1, 0)$ and $(0, 0, 1)$, e.g. $(1, 0, 0) = c_1(0, 1, 0) + c_2(0, 0, 1)$ for some $c_1, c_2 \in \mathbb{R}$. With this, we have that $(1, 0, 0) = (0, c_1, c_2)$. However, $1 \neq 0$, and we have a contradiction. The other two cases are similar. $\square$

Definition 4.4. We say a subset $\mathcal{S}$ of a vector space V is **spanning** (or **generating**) if every $\mathbf{v} \in V$ is a linear combination of some elements of $\mathcal{S}$.

Theorem 4.5

The vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_m \in \mathbb{R}^n$ are linearly dependent if and only if there are real numbers $c_1, c_2, \cdots, c_m$, 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 $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ are linearly independent if and only if the following statement is true:

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

Proof. We will start with the forward direction. Assume that they are linearly dependent; we will try to prove that $c_1, c_2, \cdots, c_m \in \mathbb{R}$ are all 0. By the definition of linear dependence, we have that one of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ is a linear combination of the others. WLOG assume that $\mathbf{v}_n = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{n-1}\mathbf{v}_{n-1}$. Now, we have that $c_1 + c_2\mathbf{v}_2 + \cdots + c_{n-1}\mathbf{v}_{n-1} - \mathbf{v}_n = 0$. Note that the coefficient of $\mathbf{v}_n$ is -1, completing the proof.

Now, we move onto the backwards direction. Assume that we have $c_1, c_2, \cdots, c_m \in \mathbb{R}$ such that at least one $c_i \neq 0$ and that the sum of all $c_i\mathbf{v}_i$ is $\mathbf{0}$. Now, we want to prove that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ are linearly dependent. WLOG assume that c_i is our nonzero coefficient. Then, we have that $-c_i\mathbf{v}_i = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{i-1}\mathbf{v}_{i-1}$. As $\mathbf{v}_i$ can be written as a linear combination of the other vectors (obtained when we divided both sides of the equation by $\frac{-1}{c_i}$, we are done. □

Definition 4.6. We say a subset $\mathcal{B}$ of a vector space V is a **basis** if $\mathcal{B}$ is both linearly independent and spanning.

Theorem 4.7

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

The proof of this theorem is left as an exercise.

Example 4.8

Prove $\{1, t, t^2, \cdots, t^n\}$ is a basis for $\mathbb{P}_n$.

Proof. This is left as an exercise for the reader. □

§4.2 Some Examples of Subspaces

Example 4.9 (Span of vectors)

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

The proof of this is left as an exercise. *Hint*: use the Subspace Criterion.

Definition 4.10. Let A be an $m \times n$ matrix. The **row space**, denoted by $\operatorname{Row}(A)$, is the subspace of $\mathbb{R}^n$ spanned by the rows of A. Analogously, the **column space**, denoted $\operatorname{Col}(A)$, is the subspace of of $\mathbb{R}^m$ spanned by the columns of A.

Example 4.11

Let $A = \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 2 & 5 & 3 \end{array}\right)$.
Then, $\operatorname{Col}(A) = \operatorname{span}(\{(1,2), (-1,5), (0,3)\})$ and
$\operatorname{Row}(A) = \operatorname{span}(\{(1,-1,0), (2,5,3)\})$.

> **Example 4.12** (Row space and column space)
> Prove that the row space and column space of every matrix is a vector space.

§5 Friday, September 9, 2022

§5.1 Solving Systems

Suppose we would like to solve the system of equations

$$T(n) = \{ \tag{1}$$

While we could use substitution and/or elimination to solve this system, these methods could get too computational, especially when the number of variables is too large. We would like to solve this system in a more organized fashion.

To do this, we can form a matrix with all of the coefficients, and have a separate part of the matrix with what the equations are equal to:

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

This matrix is called the **augmented matrix** of the system.

Using elimination, we add an appropriate multiple of one of the equations to another equation. This means that we can do the same thing to the rows of the augmented matrix. Note: each step is reversible, and we are not inserting or eliminating any solutions. In this process, the three operations that are used are listed below, and are called **elementary row operations**.

- **Row Addition**: adding a scalar multiple(s) of a row to another row
- **Row Interchange**: interchanging two rows
- **Row Rescaling**: Multiplying a row by a nonzero number

The objective is to obtain a matrix that satisfies the following:

- All zero rows are at the bottom
- The column of the first nonzero entry of each row has only one nonzero entry
- 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 **echelon form**.

If in addition to the above, the first nonzero entry of each row is 1, we say the matrix is in **reduced echelon form**.

There are numerous steps and techniques to apply this method. They are listed in Dr. Ebrahimian's notes.

Theorem 5.1

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 5.2. 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 5.3. A system of linear equations is called **homogeneous** if the right hand side (RHS) is all zeroes. In other words, any homogeneous system is of the following form:

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

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 solution, 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 5.4

Find all solutions of the system

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

We can solve this system using an augmented matrix; with this method, we can prove the following theorem:

Theorem 5.5

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

Corollary 5.6

Every $n + 1$ vectors in $\mathbb{R}^n$ are linearly dependent. In other words, if $m > n$, any m vectors in $\mathbb{R}^n$ are linearly dependent.

Proof. Let $(a_{11}, a_{12}, \cdots, a_{1n}), (a_{21}, a_{22}, \cdots, a_{2n}) \cdots, (a_{m1}, a_{m2}, \cdots, a_{mn})$ be m vectors in $\mathbb{R}^n$. We need to find $(c_1, c_2, \cdots, c_n)$ such that not all of the c_i's are 0 and such that

$$c_1(a_{11}, a_{12}, \cdots, a_{1n}) + c_2(a_{21}, a_{22}, \cdots, a_{2n}), (a_{m1}, a_{m2}, \cdots, a_{mn}) + \cdots + \\ c_m(a_{m1}, a_{m2}, \cdots, a_{mn}) = \vec{0}.$$

This is equivalent to

$$(c_1 a11 + c_2 a_{21} + \cdots + c_m a_{m1}, \ \cdots \ c_1 a_{m1} + c_2 a_{m2} + \cdots c_n a_{mn}) = 0.$$

This is, in turn, equivalent to

$$\begin{Bmatrix} c_1 a_{11} + c_2 a_{21} + \cdots c_m a_{m1} = 0 \\ \vdots \\ c_1 a_{1n} + c_2 a_{2n} + \cdots c_m a_{mn} = 0 \end{Bmatrix}$$

Note that we have n variables and m equations. By Theorem 5.5, this system has a nontrivial solution. Thus, the system of equations is true for some $c_1, c_2, \cdots, c_m$ that are not all 0. $\square$

§6 Monday, September 12, 2022

§6.1 Dimension of a Vector Space

Definition 6.1. 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 6.2

Find the dimension of each of the following vector spaces:

a) $\mathbb{R}^n$

b) $\{0\}$

c) The line $y = 3x$ in the xy-plane.

For a), our dimension is n, as one basis consists of $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$. For b), our dimension is 0, because $\vec{0}$ is not dependent on any other vectors. For c), we can create any point on the line $y = 3x$ with any other point, so our dimension is 1.

Theorem 6.3

Let V be a subspace of $\mathbb{R}^n$. Then,

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

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

The proof of this theorem is left as an exercise.

Example 6.4

If V is a subspace of $\mathbb{R}^n$, dim $V = 2$ or 1 or 0. dim $V = 2 \to \{\vec{v}, \vec{w}\}$ is a basis for V.

Theorem 6.5

Let A be a matrix. Then,

a) the dimension of $\text{Row}(a)$ 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 $\text{Row}(a)$.

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

Remark 6.6. To find a basis for $\text{Row}(A)$, we must use the nonzero rows of the **echelon form** of A. On the other hand, to find a basis for $\text{Col}(A)$, we must use the pivot columns of A (NOT the echelon form of A).

Example 6.7

Find a basis for $\text{Row}(a)$ and $\text{Col}(a)$, where

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

We can start to row-reduce this matrix.

Definition 6.8. The **rank** of a matrix A, denoted by rank A, is the dimension of $\text{Row}(A)$ (which is the same as the dimension of $\text{Col}(A)$) (why?).

Definition 6.9. 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.

Example 6.10

Find a basis for $\text{span}\{(1, 2, 0, 1), (-1, 1, 2, 1), (1, 5, 2, 3), (2, 1, 2, 0)\}$.

We can put these vectors in rows or columns. If you put them as rows, you can just look at the final result (echelon form) to find the basis. If you put the vectors as columns, you should follow the second part of the theorem, and look at the pivot columns of A (the original matrix A).

Example 6.11

Given an $m \times n$ matrix A, whose columns are $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n \in \mathbb{R}^n$, prove that the set of all vectors $\mathbf{u} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ for which

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

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

Proof. We already know that $(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, so this set of vectors is a subspace of $\mathbb{R}^n$. Proving closure under addition and scalar multiplication is trivial. $\square$

§7 Wednesday, September 14, 2022

So far, we've discussed linear independence and dependence, how to generate a basis, span, dimension, etc. However, there are more things that we can discover from vector spaces.

§7.1 Inner Products

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

Example 7.1

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 (recall what this states), prove that $x_1x_2 + y_1y_2 = \sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2}\cos(\theta)$.

Simply make a triangle, use the Law of Cosines, and do some algebraic manipulation to prove this.

Remark 7.2. $x_1x_2 + y_1y_2$ is known as the **dot product** of $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$.

Definition 7.3. 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 $\mathbf{x} \neq 0$ (Positivity)

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

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

Example 7.4

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

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

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

To prove that something is an inner product, we can take the inner product of a vector with itself.

Example 7.5

Prove $\langle(x_1, x_2, \cdots, x_n), (y_1, y_2, \cdots, y_n)\rangle = x_1y_1 + x_2y_2 + \cdots x_ny_n$ is an inner product in $\mathbb{R}^n$.

Proof. $\langle(x_1, x_2, \cdots, x_n), (x_1, x_2, \cdots, x_n)\rangle = \sum_{i=1}^{n} x_i^2$, which is greater than or equal to 0 if $x \neq 0$, proving positivity. To prove symmetry, we can prove that

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

Because multiplication is commutative, we are done. To prove linearity, we can take take the inner product

$$\langle(a(x_1, x_2, \cdots, x_n) + b(y_1, y_2, \cdots, y_n), (z_1, z_2, \cdots, z_n)\rangle$$

manipulate the inside of the product, and then see that the inner product satisfies linearity. □

Remark 7.6. The above inner product (the dot product) of $\mathbb{R}^n$ is called the **standard inner product** of $\mathbb{R}^n$.

Definition 7.7. The **length** of a vector $\mathbf{v} \in \mathbb{R}^n$ relative to an arbitrary inner product is given by $||v|| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$. Therefore, the length of a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ relative to the standard inner product is given by

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

Note that this definition matches the definition of Euclidean distance in $\mathbb{R}^n$.

Remark 7.8. In Example 7.1, note that when $\theta = \frac{\pi}{2}$, we have $\mathbf{v} \cdot \mathbf{w} = 0$. This suggests the following definition:

Definition 7.9. 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 $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are **orthogonal** if and only if $\mathbf{v}_i$ and $\mathbf{v}_j$ are orthogonal for every $i \neq j$.

Example 7.10

Show that $(1, 2, -1)$ and $(-1, 1, 1)$ are orthogonal vectors of $\mathbb{R}^n$ using the standard inner product of $\mathbb{R}^3$.

Solution. $1(-1) + 2(1) - 1(1) = 0$, so the vectors are orthogonal. □

Example 7.11

Let $e_i \in \mathbb{R}^n$ be the vector whose i-th component is 1 and whose all other components are zero. Then, $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ forms an orthogonal basis for $\mathbb{R}^n$.

Solution. We have previously shown that $\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \cdots, \vec{\mathbf{e}}_n$ form a basis for $\mathbb{R}^n$. Now, note that the dot product for every $\vec{\mathbf{e}}_i, \vec{\mathbf{e}}_j = 0$ for every $i, j \neq 0$. Thus, $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ forms an orthogonal basis for $\mathbb{R}^n$. $\square$

Theorem 7.12 (Pythagorean Theorem)

If vectors $\mathbf{v}$, $\mathbf{w}$, $\in \mathbb{R}^n$ are orthogonal relative to an inner product of $\mathbb{R}^n$, then

$$||v||^2 + ||w||^2 = ||v + w||^2$$

Proof. $||\vec{\mathbf{u}} + \vec{\mathbf{v}}||^2 = \langle \vec{\mathbf{u}} + \vec{\mathbf{v}}, \vec{\mathbf{u}} + \vec{\mathbf{v}} \rangle$. This, in turn, by linearity, is equal to $\langle \vec{\mathbf{u}}, \vec{\mathbf{u}} + \vec{\mathbf{v}} \rangle + \langle \vec{\mathbf{v}}, \vec{\mathbf{u}} + \vec{\mathbf{v}} \rangle$. Now, we can apply symmetry and linearity, and we are done. $\square$

Example 7.1 suggests that we should define the angle θ between two vectors $\mathbf{v}$, $\mathbf{w}$, by $\cos(\theta) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{||\mathbf{v}||||\mathbf{w}||}$. In order to define this, we need the following:

Theorem 7.13 (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 7.14. The angle between two 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 7.15

Find the angle between $(1, 2, -1)$ and $(1, 1, 3)$, once relative to the standard inner product and one relative to the inner product given by $\langle (x_1, y_1, z_1), (x_2, y_2, z_2) \rangle = x_1x_2 + 2y_1y_2 + 3z_13z_2$.

Solution. For the first part, we have $\theta = \cos^{-1}\left(\frac{1+2-3}{\sqrt{6}\cdot\sqrt{11}}\right) = \frac{\pi}{2}$. For the second part, we have that $\theta = \cos^{-1}\left(\frac{1+4-9}{\sqrt{1+8+3}\cdot\sqrt{30}}\right) = \cos^{-1}\left(\frac{-4}{\sqrt{360}}\right)$. Note that to find the denominators (e.g. the magnitudes of the vectors), we **must** use the **given** inner product. $\square$

§8 Friday, September 16, 2022

§8.1 Cauchy-Schwarz Inequality

Recall last class, where we introduced the **Cauchy-Schwarz Inequality** to allow us to create our formula for $\cos\theta$.

Let's now prove this inequality.

Proof. Suppose we have two vectors $\mathbf{v}$ and $\mathbf{w}$. We can **project** $\mathbf{v}$ onto $\mathbf{w}$ to get a third vector, known as $\text{proj}_{\mathbf{w}}\mathbf{v}$. The formula for a projection is $\text{proj}_{\mathbf{w}}\mathbf{v} = \left(\frac{\langle \mathbf{v},\mathbf{w}\rangle}{||v||^2}\mathbf{w}\right)$. This is because the projection must be a multiple of $\mathbf{w}$, as we are projecting a vector directly onto $\mathbf{w}$.

Let's now continue with the proof. Assume $\mathbf{w} \neq 0$. Then, we have that the components that make the triangle with the projection of $\mathbf{v}$ onto $\mathbf{w}$ and $\mathbf{v}$ are $\frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}$ and $\mathbf{v} - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}$. These are the "$x$" and "$y$" components of the vector $\mathbf{v}$. Taking the inner product $\left\langle \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}, \mathbf{v} - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}\right\rangle$. We can use linearity and symmetry as follows:

$$\begin{aligned}
\left\langle \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}, \mathbf{v} - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}\right\rangle &= \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\langle \mathbf{w},\mathbf{v}\rangle - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\langle \mathbf{w},\mathbf{w}\rangle && \text{Linearity} \\
&= \frac{\langle \mathbf{v},\mathbf{w}\rangle^2}{||\mathbf{w}||^2} - \frac{\langle \mathbf{v},\mathbf{w}\rangle^2}{||\mathbf{w}||^4}||\mathbf{w}||^2 && \text{Symmetry, Def. of Magnitude} \\
&= \frac{\langle \mathbf{v},\mathbf{w}\rangle^2}{||\mathbf{w}||^2} - \frac{\langle \mathbf{v},\mathbf{w}\rangle^2}{||\mathbf{w}||^2} \\
&= 0
\end{aligned}$$

Hence, the two vectors are orthogonal, allowing us to proceed with our proof.
By the Pythagorean Theorem, we have

$$||\mathbf{v}||^2 = \left|\left|\frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}||^2 + \right|\right|\mathbf{v} - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}\left|\right|^2 \geq \left|\left|\frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}\right|\right|^2$$

$$\begin{aligned}
\left|\left|\frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}\right|\right|^2 &= \left\langle \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}, \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\mathbf{w}\right\rangle \\
&= \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}\langle \mathbf{w},\mathbf{w}\rangle \\
&= \frac{\langle \mathbf{v},\mathbf{w}\rangle}{||\mathbf{w}||^2}||\mathbf{w}||^2 \\
&= \frac{\langle \mathbf{v},\mathbf{w}\rangle^2}{||\mathbf{w}||^2}
\end{aligned}$$

Thus, $||\mathbf{v}||^2 \geq \frac{\langle \mathbf{v},\mathbf{w}\rangle^2}{||\mathbf{w}||^2}$, implying $\langle \mathbf{v},\mathbf{w}\rangle \geq ||\mathbf{v}||^2\ ||\mathbf{w}||^2$, completing the proof. □

§8.2 Normed Spaces

To better understand the geometry of $\mathbb{R}^n$ better, we need to able to measure lengths. Given vectors (x_1, x_2) and (x_1, x_2, x_3) in $\mathbb{R}^n$, their lengths are given by $\sqrt{x_1^2 + x_2^2}$ and $\sqrt{x_1^2 + x_2^2 + x_3^2}$ using the distance formula. This suggests that we can define length (or magnitude or norm) in $\mathbb{R}^n$ in a similar fashion. But how about other vector spaces? For those, we will define a norm to be an assignment of nonnegative real numbers to vectors that satisfy the properties of such a geometric distance.

Definition 8.1. A **norm** on a vector space V is a function that assigns to any vector $\mathbf{v} \in V$ to a nonnegative real number $||v||$ that satisfies all of the following:

a) $||\mathbf{v}|| > 0$ for every nonzero $\mathbf{v} \in V$ (Positivity).

b) $||\mathbf{v} + \mathbf{w}|| \leq ||\mathbf{v}|| + ||\mathbf{w}||$ for every $\mathbf{v}, \mathbf{w} \in V$ (Triangle Inequality).

c) $||c\mathbf{v}|| = |c|||\mathbf{v}||$ for every $\mathbf{v} \in V$ and $c \in \mathbb{R}$ (Homogeneity)

Any vector space equipped with a norm is called a **normed space** (or a **normed vector space**).

The following theorem connects the notions of the inner product and the norm:

Theorem 8.2

If $\langle , \rangle$ is an inner product on a vector space V, then the function defined by $||v|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ is a norm.

Proof. This is left as an exercise. □

Example 8.3

The following are examples of a norm:

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

b) $||(x_1, x_2, \cdots, x_n)|| = \max\{|x_1|, |x_2|, \cdots |x_n|\}$ in $\mathbb{R}^n$.

Theorem 8.4

Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are non-zero, orthogonal vectors in an inner product space. Then, they are linearly independent.

Proof. Suppose $c_1, c_2, \ldots, c_n \in \mathbb{R}^n$ for which $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots c_n\mathbf{v}_n = 0$. Taking the inner product of both sides with $\mathbf{v}_1$, we obtain $\langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n, \mathbf{v}_1\rangle = c_1\langle \mathbf{v}_1, \mathbf{v}_1\rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_1\rangle + \cdots + c_n\langle \mathbf{v}_n, \mathbf{v}_1\rangle = c_1||v_1||^2$, since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are orthogonal.

By an exercise, $\langle 0, \mathbf{v}_1\rangle = 0$. Thus, $c_1||v_1||^2 = 0$. Since $\mathbf{v}_1 \neq \mathbf{0}$, $c_1 = 0$. Similarly, $c_2, c_3, \ldots, c_n = 0$, completing the proof. □

§9 Monday, September 19, 2022

§9.1 Gram-Schmidt Orthogonalization Process

Theorem 9.1 (Gram-Schmidt Orthogonalization Process)
Let $\langle , \rangle$ be an inner product on $\mathbb{R}^m$, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be linearly independent vectors in $\mathbb{R}^m$. Define vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ recursively as follows:

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

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

Corollary 9.2
Every finite-dimensional inner product space has an orthogonal basis.

Definition 9.3. 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 (e.g. all of them have norm 1).

§9.2 Linear Mappings

Remark 9.4. All vector spaces are subspaces of $\mathbb{R}^n$ for some positive integer n.

Definition 9.5. 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** if and only 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)
- $L(c\mathbf{v}) = cL(\mathbf{v})$ (Homogeneity)

Example 9.6

Determine which of the following functions are linear:

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

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

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

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

Solution. For part (a), we have $f(x + y) = c(x + y) = cx + cy = f(x) + f(y)$, so f is additive $f(ax) = cax = acx = f(x)$, so f satisfies homogeneity. For part (b), suppose (x, y) and $(z, t) \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then, we have $g((x, y) + (z, t)) = g(x + z, y + t) = 2(x + z) + 3(y + t) = (2x + 3y) + (2z + 3t) = g(x, y) + g(z, t)$, so g is additive. $g(c(x, y)) = g(cx, cy) = 2cx + 3cy = c(2x + 3y) = cg(x, y)$, so g satisfies homogeneity. In part (b), suppose we have x and y. Then, $f(x + y) = 2(x + y) + 3 = 2x + 2y + 3$, whereas $f(x) + f(y) = (2x + 3) + (2y + 3) = 2x + 2y + 6$. As a result, $f(x + y) \neq f(x) + f(y)$, meaning h is not additive, and is therefore not a linear transformation. For part (d), suppose we have $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$. Then, $k(\mathbf{v}_1 + \mathbf{v}_2) = \langle \mathbf{w}, \mathbf{v}_1 + \mathbf{v}_2 \rangle = \langle \mathbf{w}, \mathbf{v}_1 \rangle + \langle \mathbf{w}, \mathbf{v}_2 \rangle = k(\mathbf{v}_1) + k(v_2)$, meaning k is additive. Now, suppose we have $\mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then, $k(c\mathbf{v}) = \langle \mathbf{w}, c\mathbf{v} \rangle = c\langle \mathbf{w}, \mathbf{v} \rangle = ck(\mathbf{v})$, so k satisfies homogeneity. $\square$

Example 9.7

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

Solution. Suppose we have a vector $\mathbf{v} = (x, y, z)$ and f is linear. Then, we have

$$\begin{aligned} f(x, y, z) &= f(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) \\ &= xf(\mathbf{e}_1) + yf(\mathbf{e}_2) + zf(\mathbf{e}_3) \end{aligned}$$

Assume $f(\mathbf{e}_1) = (a_1, b_2)$, $f(\mathbf{e}_2) = (a_2, b_2)$, and $f(\mathbf{e}_3) = (a_3, b_3)$. Then, $f(x, y, z) = x(a_1, b_1) + y(a_2, b_2) + z(a_3, b_3) = (xa_1 + ya_2 + za_3, xb_1 + yb_2, zb_3)$. Note that $xa_1 + ya_2 + za_3$ is the standard inner product of (x, y, z) and (a_1, a_2, a_3) and $xb_1 + yb_2 + zb_3$ is the standard inner product of (x, y, z) and (b_1, b_2, b_3). Therefore, we have $f(\mathbf{v}) = (\mathbf{v} \cdot \mathbf{a}, \mathbf{v} \cdot \mathbf{b})$, where $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ are fixed vectors. Recall that $\mathbf{v} = (x, y, z)$. We've just showed that if f is linear, then there are vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ such that $f(\mathbf{v}) = (\mathbf{v} \cdot \mathbf{a}, \mathbf{v} \cdot \mathbf{b})$. We now have to show that this is linear, and if it is not linear, we have to show which vectors $\mathbf{a}$ and $\mathbf{b}$ work. However, because we are working with an inner product, by linearity of the inner product, we are done. $\square$

Example 9.8

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

Solution. Suppose we have $\mathbf{v} = (x_1, x_2, \dots, x_n)$. Then, we have $f(\mathbf{v}) = f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n$. Then, we have that

$$\begin{aligned} f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n &= x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \cdots + x_n f(\mathbf{e}_n) \\ &= (x_1, x_2, \cdots, x_n) \cdot (f(\mathbf{e}_1) f(\mathbf{e}_2), \dots, f(\mathbf{e}_n)) \end{aligned}$$

The above is a dot product of $\mathbf{v}$ and $(f(\mathbf{e}_1)f(\mathbf{e}_2), \dots, f(\mathbf{e}_n))$, where each $\mathbf{e}_i$ is a fixed vector. □

Example 9.9

Prove that if $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m \in \mathbb{R}^n$ are given vectors, then the mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$f(\mathbf{v}) = \begin{pmatrix} \mathbf{w}_1 \cdot \mathbf{v} \\ \mathbf{w}_2 \cdot \mathbf{w} \\ \vdots \\ \mathbf{w}_m \cdot \mathbf{v} \end{pmatrix}$$

is linear.

Proof. This is left as an exercise, as it is quite similar to other problems we've solved in this class. □

§9.3 Matrices

Definition 9.10. Given an $m \times n$ matrix

$$\begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_m \end{pmatrix}$$

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 given by

$$A\mathbf{v} = \begin{pmatrix} \mathbf{w}_1 \cdot \mathbf{v} \\ \mathbf{w}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{w}_m \cdot \mathbf{v} \end{pmatrix}$$

Theorem 9.11

A mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if there is an $m \times n$ matrix A for which $f(\mathbf{v}) = A\mathbf{v}$. Furthermore, for a given linear mapping f, the matrix A is unique. The columns of A are given by $f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n)$. In other words,

$$A = (f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n)).$$

Proof. We will only prove that A is unique (the rest of the proof was already done with Example 9.6), e.g. that the only A works is $A = (f(\mathbf{e}_1, f(\mathbf{e}_2, \cdots, f(\mathbf{e}_n))$. Let $f(\mathbf{v}) = A\mathbf{v}$. We have that $f(x_1, x_2, \cdots, x_n) = f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) = x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \cdots + x_n f(\mathbf{e}_n)$. Note that $f(\mathbf{e}_1) = A\mathbf{e}_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} \mathbf{e}_1$.

This, in turn, is equivalent to

$$\begin{pmatrix} (a_{11}, a_{12}, \cdots, a_{1n}) \cdots (1, 0, \cdots, 0) \\ (a_{21}, a_{22}, \cdots, a_{22}) \cdots (1, 0, \cdots, 0) \\ \vdots \\ (a_{m1}, a_{m2}, \cdots, a_{mn}) \cdots (1, 0, \cdots, 0) \end{pmatrix}$$

Note that the first column of A is $f(\mathbf{e}_1)$. Similarly, if we consider $f(\mathbf{e}_j) = A\mathbf{e}_j$, we conclude that the j-th column of A if $f(\mathbf{e}_j)$, proving that A is unique. Note that the set of vectors $\mathbf{u}$ is called the **null space** or **kernel** of A. □

§10 Wednesday, September 21, 2022

Recall last class, when we introduced the notions of matrix multiplication and linear mappings.

Example 10.1

Let $\alpha \in [0, 2\pi]$ be an angle. Suppose $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation of angle α about the origin. From geometry, we know R_α is linear. Find the matrix of R_α.

Solution. By Theorem 9.9, we have $M_{R_\alpha} = (R_\alpha(\mathbf{e}_1) R_\alpha(\mathbf{e}_2))$ (this isn't multiplication; it's simply saying that we're putting both matrices together). Now, suppose we want to rotate the point (x, y) (which is already at an arbitrary angle β) and we want to rotate this point to obtain (x', y') at angle α). Note that $x' = r\cos(\alpha + \beta)$ and $y' = r\sin(\alpha + \beta)$, where r is the length of the line from the origin to the point x', y'. Then, we have that $x' = r\cos(\alpha + \beta)$ and $y' = \sin(\alpha + \beta)$. We use the trigonometric identities for $\cos(\alpha + \beta)$ and for $\sin(\alpha + \beta)$, and substitute $x = r\cos(\beta)$ and $y = r\sin(\beta)$ to get that $x' = r\cos(\alpha) - r\sin(\alpha)$ and $y' = r\cos(\alpha) + r\sin(\alpha)$. As a result, we obtain $R_\alpha(\mathbf{e}_1) = \begin{pmatrix} \cos\alpha \\ -\sin\alpha \end{pmatrix}$ and $R_\alpha(\mathbf{e}_1) = \begin{pmatrix} \sin\alpha \\ \cos\alpha \end{pmatrix}$. Combining these, we have that $M_{R_\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$.

□

Remark 10.2. To rotate about an arbitrary point O', we can first translate O' to the origin along with all other points, perform the rotation, and inverse the translation to move O' and all other points to their initial positions.

Definition 10.3. The **matrix** A of the linear mapping f in Theorem 9.9 is called the matrix of f with respect to the standard basis denoted by M_f.

Definition 10.4. Let A be an $m \times n$ matrix and let B be an $n \times k$ matrix. Then, the matrix AB is an $m \times k$ matrix whose jth column is obtained by the j-th column of B. In other words, the (i, j)-th entry of AB is obtained by finding the standard inner product (dot product) of the i-th row of A with the j-th column of B.

Remark 10.5. 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 equal to the number of rows of B.

Example 10.6

Evaluate the matrices AB and BA, where

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

Solution. AB is not possible to compute, e.g. it is not defined. This is because A is a 2×2 matrix, and B is a 3×2 matrix. In other words, the number of columns of A (which is 2) is different from the number of rows of B, which is 3. We have $BA = \begin{pmatrix} -3 & -1 \\ 11 & 7 \\ 2 & 9 \end{pmatrix}$. □

Example 10.7

Consider a 2×3 matrix A and 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:

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

Solution. By Definition 9.8, we have

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

□

Remark 10.8. 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 10.9

If the mappings $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}_k$ are linear, then $g \circ f : \mathbb{R}^m \to \mathbb{R}^k$ 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 that the j-th column of the matrix $g \circ f$ is $g \circ f(\mathbf{e}_j)$. This equals $g(f(\mathbf{e}_j)) = M_g f(\mathbf{e}_j)$. Sine 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. □

Example 10.10

The matrix of the identity mapping $I : \mathbb{R}^n \to \mathbb{R}^n$, defined by $I(\mathbf{x}) = \mathbf{x}$ is the matrix

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

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

Theorem 10.11

For matrices A, B, C¡ and a real number r, we have the following rules:

- $A(BC) = (AB)C$ (associativity)
- $A(B + C) = AB + AC$ (distributivity)
- $r(AB) = (rA)B = A(rB)$
- $AI = IA = A$

All of these properties hold if and only if the appropriate addition and matrix multiplications are possible.

Proof. This is left as an exercise. □

§10.1 Kernel and Image

Definition 10.12. Given a linear mapping $L : V \to W$, the **kernel** of L is defined as $\text{Ker}L = L^{-1}(\{0\})$. In other words,

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

The **image** 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 10.13
Let $L : V \to W$ be a linear mapping of vector spaces. Then, $\text{Ker}L$ is a subspace of V and $\text{Im}L$ is a subspace of W.

§11 Friday, September 23, 2022

Let's prove Theorem 10.12.

Proof. We can apply the Subspace Criterion. To prove $\text{Ker}L$ is a subspace of V (e.g. the domain of L), first note that $\mathbf{0} \in \text{Ker}L$, by the definition of L. Now, suppose we have $\mathbf{x}, \mathbf{y} \in \text{Ker}L$. Then, we have that $L(\mathbf{x}+\mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}) = 0$, proving closure under addition. Similarly, we have $L(c\mathbf{x}) = cL(\mathbf{x}) = c \cdot 0 = 0$.

The proof of $\text{Im}L$ being a subspace of W is left as an exercise. □

Example 11.1
Find the kernel and image of the linear transformation given by $L(x, y, z) = (x + 2y + z, 2x - y - z)$.

Solution. To find $\text{Ker}L$, we can solve the system of equations

$$\begin{cases} x + 2y + z = 0 \\ 2x - y - z = 0 \end{cases}$$

Solving this system by putting the equations into a matrix and reducing it to echelon form, we get that $(x, y, z) = (\frac{1}{5}z, -\frac{3}{5}z, z) = z(\frac{1}{5}, -\frac{3}{5}, 1)$. Therefore, $\text{Ker}L = \text{span}\big\{\big(\frac{1}{5}, -\frac{3}{5}, 1\big)\big\}$.

To find the image, note that $(a, b) \in \text{Im}L$ if and only if $(a, b) = (x+2y+z, 2x-y-z) = x(1, 2) + y(2, -1) + z(1, -1)$. Thus, we have $\text{Im}L = \text{span}\{(1, 2), (2, -1), (1, -1)\}$.

We already say that these are two pivot entries for the echelon form of $\begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & -1 \end{pmatrix}$
Thus, by Theorem 6.5, $\text{Im}L$ is a two-dimensional. Since $\text{Im}L$ is a subspace of $\mathbb{R}^2$ and $\dim\mathbb{R}^2 = 2, \text{Im}L = \mathbb{R}^2$. □

Theorem 11.2
Suppose $L : V \to W$ is linear. $\text{Ker}L = \{\mathbf{0}\}$ implies L is one-to-one and $\dim \text{Im}L = \dim V$.

Proof. Recall how we prove that a function is one-to-one: if $f(a_1) = f(a_2)$, $a_1 = a_2$. Suppose $L(\mathbf{v}_1) = L(\mathbf{v}_2)$ for some $\mathbf{v}_1$ and $\mathbf{v}_2 \in V$. We would like to show that $\mathbf{v}_1 = \mathbf{v}_2$. If $L(\mathbf{v}_1) = L(\mathbf{v}_2)$, we have that $L(\mathbf{v}_1) = L(\mathbf{v}_2) = 0$. Applying additivity, we have that $L(\mathbf{v}_1 - \mathbf{v}_2) = 0$, implying that $\mathbf{v}_1 - \mathbf{v}_2 = 0$ (because $\text{Ker}L = \{\mathbf{0}\}$). Hence, $\mathbf{v}_1 = \mathbf{v}_2$.

Now, we will prove dim $\text{Im}L = \text{dim}V$. Here is a sketch: take a basis for V and map it to some vectors in W, and show that these vectors are a basis for $\text{Im}L$. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v})n$ form a basis for V. Then, $L(\mathbf{v}_1), L(\mathbf{v}_2, \ldots, L(\mathbf{v}_n)$ form a basis for $\text{Im}L$. If we prove this (e.g. that these vectors span $\text{Im}L$ and are linearly independent, we will have proved that they are a basis for $\text{Im}L$, proving dim $\text{Im}L = \text{dim}V$. ☐

Theorem 11.3

Suppose we have $L : \mathbb{R}^n \to \mathbb{R}^n$ and $L(\mathbf{v}) = A\mathbf{v}$, where A is $m \times n$. Then, we have:

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

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

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

Proof. Recall the proof of part b) from Homework 3. Here is a sketch of the proof of part a): note that $L(\mathbf{v}) = A\mathbf{v}$, e.g. vectors of the image are of the form $A\mathbf{v}$. We would like to show that these are exactly the vectors as in the column space of A. This is because

$$(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

proving that elements in $A\mathbf{v}$ are also in $\text{Col}(A)$. We now have to prove that $\text{Col}(A) \subseteq \text{Im}L$. This is left as an exercise.

To prove part c), recall that $\dim \text{Col}(A) = \dim(\text{Row}(A))^\perp + \dim(Row(A) = n$, by an earlier theorem. This is because $\text{Row}(A)$ is a subspace of $\mathbb{R}^n$. ☐

Theorem 11.4 (Rank-Nullity Theorem)

Suppose we have $L : V \to W$. Then, we have $\dim \text{Ker}L + \dim \text{Im}L = \text{dim}V$.

Proof. The idea behind this proof is similar to that Theorem 11.4; try to prove this as an exercise. ☐

§11.1 Determinants

In this section, we would like to define the determinant of a square matrix. One interpretation of determinants is volume. Given n vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, we want the $n \times n$ determinant $\det(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ to determine the volume of the parallelpiped determined by these n vectors. We expect an reasonable value to follow the properties discussed below.

Define a function $D : M_2(\mathbb{R}^2) \to \mathbb{R}$. Our objective is to have

$$D\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$$

represent the area of the parallelogram based on side $\mathbf{u}$ and base $\mathbf{v}$. Suppose we scale our base by c; we now want to prove

$$D\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = cD\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$$

which is essentially homogeneity.

§12 Monday, September 26, 2022

Definition 12.1. Let $D : M_n(\mathbb{R}) \to \mathbb{R})$ be a function.

(a) We say D is **multilinear** if D is linear with respect to each row. In other words, we have

$$D\begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ a\mathbf{v}_i + b\mathbf{w} \\ \vdots \\ \mathbf{v}_n \end{pmatrix} = aD\begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_i \\ \vdots \\ \mathbf{v}_n \end{pmatrix} + bD\begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{w} \\ \vdots \\ \mathbf{v}_n \end{pmatrix}$$

(b) We say D is **alternating** if $D\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} = 0$ when $\mathbf{v}_i = \mathbf{v}_j$ for some $i \neq j$.

To keep notations more compact, instead of writing $D\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix}$, we write $D(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, inserting commas to indicate $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are rows and not columns.

Example 12.2

Let $D : M_2(\mathbb{R}) \to \mathbb{R}$ be an alternating, multi-linear function. Prove that

$$D(\mathbf{u}, \mathbf{v}) = -D(\mathbf{v}, \mathbf{u}).$$

Proof. First, note that we can't always define D to be volume. This is because $-D(\mathbf{v}, \mathbf{u})$ cannot be defined. Now, let's continue to the proof. Note that $D(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}) = 0$, because D is alternating. Also note that

$$\begin{aligned} D(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}) &= D(\mathbf{u}, \mathbf{u}+\mathbf{v}) + D(\mathbf{v}, \mathbf{u}+\mathbf{v}) && \text{Linearity with } \mathbf{u} \\ &= D(\mathbf{u}, \mathbf{u}) + D(\mathbf{u}, \mathbf{v}).D(\mathbf{v}, \mathbf{u}) + D(\mathbf{v}, \mathbf{v}) && \text{Multilinearity} \\ &= 0 + D(\mathbf{u}, \mathbf{v}) + D(\mathbf{v}, \mathbf{u}) + 0 \to D(\mathbf{u}, \mathbf{v}) = -D(\mathbf{v}, \mathbf{u}) \end{aligned}$$

□

Example 12.3

Find all alternating, multilinear functions $D : M_2(\mathbb{R}) \to \mathbb{R}$ satisfying $D(I) = 1$, where I is the identity matrix.

Solution. Assume that we have some D that satisfies these conditions. Then, we have $D\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}(1,0) = a_{12}(0,1) = (a_{11}, a_{12})$ and $a_{22}(1,0)+a_{22}(0,1) = a_{12}+a_{22}$. Then, we have that $D\begin{pmatrix} a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 \\ a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 \end{pmatrix} = a_{11}D\begin{pmatrix} \mathbf{e}_1 \\ a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 \end{pmatrix} + a_{12}D\begin{pmatrix} \mathbf{e}_2 \\ a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 \end{pmatrix}$ $= a_{11}a_{21}D\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_1 \end{pmatrix} + a_{11}a_{22}D\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} + a_{12}a_{21}D\begin{pmatrix} \mathbf{e}_2 \\ \mathbf{e}_1 \end{pmatrix} + a_{12}a_{22}D\begin{pmatrix} \mathbf{e}_2 \\ \mathbf{e}_2 \end{pmatrix} = 0 + a_{11}a_{22}D\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} + a_{12}a_{12}D\begin{pmatrix} \mathbf{e}_2 \\ \mathbf{e}_1 \end{pmatrix} = a_{11}a_{22}D(I) - a_{12}a_{21}D(I) = a_{12}a_{22} - a_{12}a_{21}$.

Note that we were able to obtain the term $a_{11}a_{22}D(I) - a_{12}a_{21}D(I)$ from Example 12.2. Before finishing the solution, we must check that this in fact does indeed satisfy the conditions of D. This is purely computational, and is left as an exercise. □

Theorem 12.4

Let $D : M_n(\mathbb{R}) \to \mathbb{R})$ be alternating and multilinear. Then, D satisfies the following properties:

(a) Swapping two rows negates D. In other words,
$$D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = -D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n).$$

(b) Rescaling a row by c rescales D by c. In other words,
$$D(\mathbf{v}_1, \mathbf{v}_2, \dots, c\mathbf{v}_i, \dots, \mathbf{v}_n) = cD(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$$

(c) Adding a multiple of one row to another does not change D. In other words,
$$D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i + c\mathbf{v}_j, \dots, \mathbf{v}_n) = D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n).$$

(d) $D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = 0$ if $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are linearly dependent.

Proof. The proof of (a) is nearly identical to that of Example 12.2. We can prove pars (b) by using linearity. Let's prove parts (c) and (d). To prove part (c), note that

$$\begin{aligned} D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i + c\mathbf{v}_j, \dots, \mathbf{v}_n) &= D(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) + D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) \\ &= D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) + 0 \\ &= D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \end{aligned}$$

To prove (d), WLOG assume $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_n\mathbf{v}_n$ for some $\mathbf{c}_2, \dots, \mathbf{c}_n \in \mathbb{R}$. Then, we have that

$$\begin{aligned} D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) &= D(c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_n\mathbf{v}_n, \mathbf{v}_2, \dots, \mathbf{v}_n) \\ &= c_2 D(\mathbf{v}_2, \mathbf{v}_2, \dots, \mathbf{v}_n) + c_3 D(\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n) + \cdots + c_n D(\mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n) \\ &= 0 + 0 + \cdots + 0 = 0 \end{aligned}$$

The last step follows because D is not alternating.

□

Theorem 12.5

For every positive integer n, there is a unique, multilinear, alternating function $D : M_n(\mathbb{R}) \to \mathbb{R}$ satisfying $D(I) = 1$.

Proof. The proof of this theorem relies on combinatorics, and is left as a (challenging) exercise. □

Definition 12.6. Let D be the function in the above theorem. Then, the **determinant** of a $n \times n$ matrix A whose rows are $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is defined as $D(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$.

Example 12.7

Evaluate

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

§12.1 Row Operations and Matrix Multiplication

The outcome of each row operation to matrix A is a matrix E as follows:

- If the operation is interchanging rows i with row j with $i < j$, then

$$E = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_j \\ \vdots \\ \mathbf{e}_i \\ \vdots \\ \mathbf{e}_n \end{pmatrix}$$

- If the operation is rescaling the i-th row by a factor of c, then

$$E = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ c\mathbf{e}_i \\ \vdots \\ \mathbf{e}_n \end{pmatrix}$$

- If the operation is adding a multiple of the j-th row to the i-th row, then

$$E = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_j + c\mathbf{e}_j \\ \vdots \\ \mathbf{e}_n \end{pmatrix}$$

Definition 12.8. Any matrix E of one of the forms above is called an **elementary matrix**.

Combining the above and Theorem 12.5, we conclude that $\det(EA) = \det(E)\det(A)$, for every $n \times n$ matrix A and $n \times n$ matrix E as above.

§13 Wednesday, September 28, 2022

We can generalize the above statement to get Theorem 13.1:

Theorem 13.1 (Multiplicativity)
Let A and B be two $n \times n$ matrices. Then, $\det(AB) = \det(A)\det(B)$.

Proof. The idea is to row-reduce the matrix A using elementary row operations (e.g. row interchange, rescaling, and addition) to achieve an elementary matrix. Suppose the rows of A are linearly independent. Because A is a square matrix and its rows are linearly independent, its reduced echelon form is the identity matrix.

Therefore, there are elementary matrices E_1, E_2, $\ldots$, E_m such that when you apply $E_1, E_2, \ldots, E_m A$, you get the identity matrix. Note that $\det(E_1 E_2 \cdots E_m A) = \det(E_1)\det(E_2 E_3 \cdots E_m A) = \det(E_1)\det(E_2)\cdots\det(E_3 \cdots E_n)$
$= \det(E_1)\det(E_2)\cdots\det(E_m)\det(A) = 1$. If we apply a similar idea to $E_1 E_2 \cdots E_m B$, we get $\det(E_1)\det(E_2)\cdots\det(E_m)\det(AB) = \det(B)$. Hence, $\det(A)\det(B) = [\det(E_1)\det(E_2)\cdots\det(E_m)]^{-1}\det(B)$, implying $\det(AB) = \det(A)\det(B)$. Note that the superscript -1 of the product of the determinants of $E_1, E_2, \ldots, E_n$ is taking the reciprocal of this number.

If the rows of A are linearly dependent, $\det(A) = 0$. Then, $\det(A)\det(B) = 0$. Assume $A = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$. Since the rows of A are linearly dependent, there are scalars $c_1, c_2, \ldots, c_n$, not all zero, such that $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots c_n\mathbf{a}_n = \begin{pmatrix} c_1 & c_2 & \ldots & c_n \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \ldots \\ \mathbf{a}_n \end{pmatrix} = 0$. Thus, we also have $\begin{pmatrix} c_1 & c_2 & \cdots & c_n \end{pmatrix} AB = 0$. This is because if the rows of A are linearly independent, the rows of AB are also linearly independent. Thus $\det(AB) = \det(A)\det(B) = 0 = \det(A)\det(B)$. $\square$

Determinants can be evaluated using co-factor expansions. Here is an example:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$
$$a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{31} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

In other words, we can write the determinant of a 3×3 matrix as follows:

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$$

§13.1 Co-Factor Expansion

Theorem 13.2 (Co-Factor Expansion Along a Row or a Column)
Let $A = (a_{ij})_{n\times n}$ be an $n \times n$ matrix with a_{ij} as its (i, j) entry. Then, for every $1 \le i \le n$, we have

$$\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{ij} \det A_{ij}$$

where A_{ij} is obtained by removing the i-th row and j-th row of A. Similarly, for every j with $1 \le j \le n$, we have

$$\det A = \sum_{i=1}^{n} (-1)^{j+1} a_{ij} \det A_{ij}$$

Proof. This is left as an exercise. □

§13.2 Inverse Matrices

Definition 13.3. A square matrix A is called **invertible** or **nonsingular** if there is a square matrix B such that $AB = BA = 1$. When A has an inverse, we denote it as A^{-1}.

Example 13.4 (Finding an inverse matrix)
Suppose A is a matrix. Find A^{-1}.

Assume that the rows of A are linearly independent. Then, apply row operations $E_n, E_{n-1}, \cdots, E_2E_1$ to A to obtain I, e.g. $E_n, E_{n-1}, \cdots, E_2E_1A = I$. Note that each E_j is invertible (why?). We can iteratively multiply both sides of this equation by I

If the rows of A are linearly dependent, then

Example 13.5
Find the inverse of $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 2 & 1 & 0 \end{pmatrix}$.

Theorem 13.6
For a square matrix A the following are equivalent:

(a) A is invertible

(b) $\det A \neq 0$

(c) Columns of A are linearly independent

(d) Rows of A are linearly independent

§13.3 Cramer's Rule

Theorem 13.7 (Cramer's Rule)
Let $A = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$ be an invertible matrix. Then, for every column $\mathbf{b}$, the only solution to $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

where $x_j = \frac{\det(\mathbf{a}_1 \mathbf{a}_2, \cdots, \mathbf{a}_{j-1}, \mathbf{b} \mathbf{a}_{j+1}, \cdots, \mathbf{a}_n)}{\det(A)}$

Proof. This is left as an exercise. □

Example 13.8
Solve the following system of equations using Cramer's rule:

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

§14 Friday, September 30, 2022

Today is a review day, with questions from the study guide being gone over.

Some errata about the study guide:

- #6 was copied incorrectly, and should be Exercise 1.16 from the notes.
- #7: $\text{span}(S) = \text{span}(T) \iff S \subseteq \text{span}(T)$ and $T \subseteq \text{span} S$
- #15: refers to $\text{Ker} L$ and $\text{Im} L$

§15 Monday, October 3, 2022

Exam #1 is today.

§16 Wednesday, October 5, 2022

§16.1 Limits and Continuity

In order to be able to define a limit of a function at a given point, we need to be able to *approach* that point. For example, consider the following function:

$$f(x) = \begin{cases} x+1 & \text{if } x < 0 \\ 2 & \text{if } x = 1 \\ 2x-1 & \text{if } 2 \leq x \end{cases}$$

Note that $\lim_{x\to 1^+}$ and $\lim_{x\to 1^-}$ cannot be defined, because 1 is an isolated point. Also note that $\lim_{x\to 2^+}$ and $\lim_{x\to 0^-}$ can be defined.

Definition 16.1. Let a be a point in $\mathbb{R}^n$. Then, the **open ball** of radius r centered at a is defined by

$$B_r(a) = \{x \in \mathbb{R}^n \mid ||x-a|| < r\}$$

Analogously, the **closed ball** of radius r centered at a is defined by

$$\overline{B_r(a)} = \{x \in \mathbb{R}^n \mid ||x-a|| \leq r\}$$

Definition 16.2. Let D be a subset of $\mathbb{R}^n$. A point a in $\mathbb{R}^n$ is called a **limit point** of D if and only if every open ball centered at a contains at least one point of D other than a.

Example 16.3

Find all limit points of $(0,1)$ in $\mathbb{R}$.

Solution. We claim a is a limit point of $(0,1)$ if and only if $a \in [0,1]$. If $a > 1$, then $B_{a-1}(a) \cap (0,1) = \emptyset$. Suppose on the contrary we have $x \in B_{a-1}(a) \cap (0,1)$. Then, we have that $|x-a| < a-1$, implying $1-a < x < a-1$, meaning we have $1 < x < 2a-1$, which is a contradiction. Similarly, if $a < 0$, then a is not a limit point. Now, we must show that everything inside $[0,1]$ is a limit point. Suppose we have $a = 1$ and $r > 0$. Since $0 < 1$ and $1-r < 1$, there is some real number x that $0 < x < 1$ and $1-r < x < 1$. For example, x could be $\max\left(\frac{1}{2}, \frac{2-r}{2}\right)$, where $\frac{1}{2}$ and $\frac{2-r}{2}$ represent the average of the intervals $(0,1)$ and $(1-r,1)$. Thus, $B_r(1)$ contains a point in $(0,1)$, which is clearly other than one, by the definiion $\square$

Definition 16.4. Let D be a subset of $\mathbb{R}^n$, let $f : D \to \mathbb{R}^n$ be a function, and let a be a limit point of D, and let $b \in \mathbb{R}^m$. We say that b is the limit of f at a, written $\lim_{x\to a} f(x) = b$ if and only if for every $\epsilon > 0$, there is a $\delta > 0$ such that if for some $x \in D$, we have $0 < ||x-a|| < \delta$ implies $||f(x) - b|| < \epsilon$.

Remark 16.5. Note that since the limit of a function at a only depends on the functional values near a; if two functions f and g are "the same, except possibly at a, then their limits at a are the same.

Definition 16.6. Let $x_k \in \mathbb{R}^n$ with $k = 1, 2, \ldots$ be a sequence and $a \in \mathbb{R}^n$. We say x_k converges to a, written as $x_k \to a$ or $\lim_{n\to\infty} = a$ iff the following holds:

$$\forall \epsilon > 0 \ \exists N \in \mathbb{N} \text{ such that } \forall k \in \mathbb{N},\ k \geq N \to ||x_k - a|| < \epsilon$$

Example 16.7

Prove $\lim_{x\to 1} 3x + 2 = 5$

Proof. We want to prove that $\forall \epsilon > 0 \ \exists \delta > 0$ such that $0 < |x - 1| < \delta \to |(3x + 2) - 5| < \epsilon$. Note that $|(3x + 2) - 5| = 3|x - 1|$. Note that $|x - 1| < \delta$, from our assumption. Thus, $3|x - 1| < 3\delta$. Thus, we can let $\delta = \frac{\epsilon}{3}$, proving our statement for $|(3x + 2) - 5|$. Note that to actually prove this statement, we must start from the proposition $\delta = \frac{\epsilon}{3}$ and prove that $||x - 1|| < \delta \to ||(3x + 2) - 5|| < \epsilon$. □

Example 16.8

Prove that $\lim_{x\to 1} \frac{1+x}{1+2x} = \frac{2}{3}$.

This was proven in class; try to prove it as an exercise.

This provides more examples of epsilon-delta proofs; read it over to get a better idea of how to do them.

§17 Friday, October 7, 2022

Let's continue with more epsilon-delta proof examples.

Example 17.1

Prove $\lim_{(x,y)\to(1,-1)} x^2 + y^2 = 2$.

Proof. We want to prove that

$$\forall \epsilon > 0 \exists \delta > 0 \text{ such that } 0 < ||(x, y) - (1, -1)|| < \delta \to |x^2 + y^2 - 2| < \epsilon$$

Note that we can can rewrite our assumption as $||(x - 1, y + 1)|| < \delta$ and treat this value as a norm. This means that we have $\sqrt{(x - 1)^2 + (y + 1)^2} < \delta$, implying $|x - 1| < \delta$ and $|y + 1| < \delta$. If $\delta \leq 1$, we get $|x - 1| < 1$, implying $0 < x < 2$. If $|y + 1| < 1$, we get $-2 < y < 0$, implying $y - 1 < -1$. Therefore, $|x^2 - 1 + y^2 - 1| = |(x + 1)(x - 1) + (y + 1)(y - 1)| \leq |(x - 1)(x + 1)| + |(y + 1)(y - 2)| \leq \delta|x + 1| + \delta|y - 1| \leq 3\delta = 3\delta = 6\delta$. Thus, we can let $\delta = \min(1, \frac{\epsilon}{6})$. Once again, to complete this proof, we must start with the proposition $\delta = \min\left(1, \frac{\epsilon}{6}\right)$ and prove $||(x, y), -(1, -1)|| < \delta \to ||(x^2 + y^2) - 2|| < \epsilon$. □

Example 17.2

Show $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ does not exist.

Solution. We will proceed with a contradiction. Suppose that on the contrary, $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = L$. Then, we want to prove that

$$\forall \epsilon > 0 \exists \delta > 0 \text{ such that } 0 < ||(x,y) - (0,0)|| < \delta \to \left|\frac{xy}{x^2+y^2} - L\right| < \epsilon.$$

Note that the $0 <$ in the first inequality is quite important, as if $(x,y) = (0,0)$, then $\frac{xy}{x^2+y^2}$ isn't defined. Set $\epsilon = \frac{1}{4}$. Then, $\exists \delta > 0$ such that whenever $\sqrt{x^2+y^2} < \delta$, $\left|\frac{xy}{x^2+y^2} - L\right| < \frac{1}{4}$. Let $x = y$. We now want to find (x,x) such that $0 < \sqrt{x^2+x^2} < \delta \to \sqrt{2}|x| < \delta$. Thus, we have $\left|\frac{\frac{\delta}{2}\cdot\frac{\delta}{2}}{\frac{\delta}{2}^2+\frac{\delta}{2}^2} - L\right| < \frac{1}{4} \to \left|\frac{1}{2} - L\right| < \frac{1}{4} \to \frac{1}{4} < L < \frac{3}{4}$. Now, let $x = \frac{\delta}{2}$ and $y = 0$. Also note that $0 < \sqrt{\frac{\delta}{2}^2 + 0^2} = \frac{\delta}{2} < \delta$. This gives us $\left|\frac{\frac{\delta}{2}\cdot 0}{\frac{\delta}{2}^2+0^2} - L\right| < \frac{1}{4}$, which gives us $-\frac{1}{4} < L < \frac{1}{4}$. This contradicts the fact that $\frac{1}{4} < L < \frac{3}{4}$, meaning that the limit does not exist. □

§18 Monday, October 10, 2022

Theorem 18.1

Suppose D is a subset of $\mathbb{R}^n$, a is a limit point of D, and $f : D \to \mathbb{R}^m$ is a function. If there are two sequences $x_k, y_k \in D - \{a\}$ for which $x_k \to a$ and $y_k \to a$, but the limits $\lim_{k\to\infty} f(x_k)$ and $\lim_{k\to\infty} f(y_k)$ are not the same, then $\lim_{x\to a} f(x)$ does not exist.

Definition 18.2. Given a function $f : D \to \mathbb{R}^m$, where D is a subset of $\mathbb{R}^n$, we write $f = (f_1, f_2, \ldots, f_m)$ if $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$ for all $x \in D$. We say functions $f_1, f_2, \ldots, f_m$ are **coordinate functions** of f.

Theorem 18.3

Let D be a subset of $\mathbb{R}^n$, a be a limit point of D, and $b = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$. Assume $f = (f_1, f_2, \ldots, f_n) : D \to \mathbb{R}^n$ is a function. Then, $\lim_{x\to a} f(x) = b$ if and only if $\lim_{x\to a} f_i(x) = b$.

Theorem 18.4 (Squeeze Theorem)
Suppose D is a subset of $\mathbb{R}^n$ and a is a limit point of D. Let $f, g, h\colon D \to R$ be functions for which

$$f(x) \leq g(x) \leq h(x)$$

for all $x \in D - \{a\}$. If $\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = L$ for some real number L, then $\lim_{x \to a} g(x) = L$.

Definition 18.5. Let a be a limit point of a subset D of $\mathbb{R}^n$. We say a function $f : D \to \mathbb{R}^n$ is **continuous** at a if $\lim_{x \to a} f(x) = f(a)$ for every point a inside its domain.

Example 18.6
Prove that the following functions are continuous:

(a) $\pi_i : \mathbb{R}^n \to \mathbb{R}$ defined by $\pi_i(x_1, x_2, \ldots, x_n) = x_i$ where $1 \leq i \leq n$ is fixed

(b) $p : \mathbb{R}^2 \to \mathbb{R}$ defined by $p(x, y) = xy$

(c) $s : \mathbb{R}^2 \to \mathbb{R}$ defined by $s(x, y) = x + y$.

Theorem 18.7
Let D be a subset of $\mathbb{R}^n$ and let a be a limit point of D. The mapping $f : D \to \mathbb{R}^n$ is continuous at a if and only if each coordinate function of f is continuous at a.

Proof. Suppose we have $f : \mathbb{R}^n \to \mathbb{R}^m$, where $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$. Note that f is continuous if and only if $f_1, f_2, \ldots, f_n$ are continuous. By the $\epsilon - \delta$ definition of the limit, we have $||f(x) - f(a)|| = \sqrt{\sum_{j=1}^{n} (f_j(x) - f_j(a))^2} \geq |f_j(x) - f_j(a)|$. Note that we have that $f_j(x) - f_j(a) < \epsilon$, by the definition of the limit. This implies that $\sqrt{(f_1(x) - f_1(a))^2 + (f_1(x) - f_1(x))^2} < \sqrt{\epsilon^2 + \epsilon^2} = \epsilon\sqrt{2}$. FINISH THIS ASAP. $\square$

Theorem 18.8
Suppose D_1 and D_2 are subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $f : D_1 \to \mathbb{R}^m$ and $g : D_2 \to \mathbb{R}^k$. Let a be a limit point of D_1, as well as of the domain of $g \circ f$. Suppose $\lim_{x \to a} f(x) = b$, $b \in D_2$, and g is continuous at b. Then, $\lim_{x \to a} g \circ f = g(b)$. Note that we need g to be continuous for the limit to exist.

Theorem 18.9

Let D be a subset of $\mathbb{R}^n$ and a be a limit point of D. Suppose f and g are real-valued functions on D. Then,

$$\lim_{x\to a} f(x) + g(x) + \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$$

and

$$\lim_{x\to a} f(x)g(x) = (\lim_{x\to a} f(x))(\lim_{x\to a} g(x))$$

assuming both $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist.

Theorem 18.10

All of the following single-variable real-valued functions are continuous over their domains: polynomials and root functions, trigonometric functions and their inverses, and exponential functions and their inverses.

Example 18.11

Prove $f : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $f(x, y, z) = (x + y, \sin(xy) + \cos(z))$ is continuous, as its coordinate functions are all continuous.

Proof. Using Theorem 17.12, we have that $x + y$ is a polynomial that is continuous. $\sin(xy)$ is continuous because it is a composition of $f(x, y) = xy$ and $f(k) = \sin(k)$. $\cos z$ is continuous directly by Theorem 17.12. Hence, we have $(x + y, \sin(xy) + \cos z)$ is continuous, as. $\square$

Example 18.12

Prove that $f : \mathbb{R}^3 \to \mathbb{R}^2$ are continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$.

Example 18.13

Every polynomial $p(x_1, x_2, \ldots, x_n)$ is a continuous function from $\mathbb{R}^n$ to $\mathbb{R}$.

§18.1 Topology of $\mathbb{R}^n$

Definition 18.14. A subset A of $\mathbb{R}^n$ is called **open** if given any point $a \in A$, there exists an open ball $B_r(a)$ (with $r > 0$) that is completely contained in A.

Example 18.15

For any positive real number r and any $a \in \mathbb{R}^n$, the ball $B_r(a)$ is open.

Proof. Suppose we have $b \in B_r(a)$. Let $s = r - ||b-a||$. Let $x \in B_s(b)$. By the definition of B_s, we have that $||x-b|| < s$. Now, we want to prove that $B_s(B) \subseteq B_r(a)$. By the Triangle Inequality, note that $||x-a|| = ||(x-b)+(b-a)|| \leq ||x-b||+||b-a||$. Also note that $||x-a|| + ||b-a|| < s + ||b-a|| = r$, implying $||x-a|| < r$, which then implies $B_r(b) \subseteq B_r(a)$ and that $B_r(a)$ is open. □

Theorem 18.16

Open sets in $\mathbb{R}^n$ satisfy the following properties:

(a) $\emptyset$ and $\mathbb{R}^n$ are open

(b) The union of any collection of open sets is open

(c) The intersection of any finite number of open sets is open

Proof. To prove part (a), we want to prove that $x \in \emptyset \to \exists r > 0$ such that $B_r(x) \subseteq \emptyset$. Note that in this case, the assumption cannot be satisfied, because we cannot have $x \in \emptyset$, by the definition of the empty set. Thus, because the assumption is false, this statement is **vacuously true**. Note that $\mathbb{R}^n$ is open because every open ball of radius r is contained in $\mathbb{R}^n$. To prove part c), let $x \in \bigcap_{i=1}^n A_i$. By the definition of the intersection, we have $x \in A_i$ for some $i = 1, 2, \ldots n$. Since A_i is open, $\exists r_i$ such that $B_{r_i}(x) \subseteq A$. If we let $r = \min(r_1, \ldots, r_n)$, we get that $B_r(x) \subseteq B_{r_i}(x) \subseteq A$, meaning $B_r(x) \subseteq \bigcap_{i=1}^n A_i$. To prove Part (b), suppose we have $U = \bigcup_{i=1}^n A_i$, If $x \in U$, we have that x is in at least one of the open sets A_i, meaning that there is a ball $B_r(a)$ centered at x that is fully contained within A_i. As $B_r(a)$ is fully contained in A_i, it must also be fully contained in U, meaning U is open. □

Example 18.17

By an example show that the intersection of a collection of open sets may not be open.

Definition 18.18. A subset A of $\mathbb{R}^n$ is said to be **closed** if $\mathbb{R}^n - A$ is open.

Example 18.19

Prove that $[a, b]$ is closed in $\mathbb{R}$.

Proof. We will first prove (a, b) is open. Let $x \in (a, b)$. Then, we have $a < x < b$, Then, we have $a < x < b$. Let $r_1 = x - a > 0$ with $a \leq x - r_1 < x$ and $r_2 = b - x > 0$ with $x < x + r_2 \leq b$. Let $r = \min(r_1, r_2) > 0$. Then, we have $a \leq x - r_1 \leq x - r < x < x + r \leq x + r_2 \leq b$. Now, we will show $B_r(x) \subseteq (a, b)$. Let $0 < \epsilon < r$. Then, we have $a \leq x - r < x - \epsilon < x < b$ and $a < x < x + \epsilon < x + r \leq b$. Thus, we have $\forall 0 < \epsilon < r$, $x \pm \epsilon \in (a, b)$, meaning $B_r(x) \subseteq (a, b)$, meaning (a, b) is open. Note that $(-\infty, a) \cup (b, \infty)$ is a union of open intervals, meaning it is an open set, by Part (b) of Theorem 18.16. This is because we can prove $\bigcup_{i=1}^{\infty}(a-i,a)$ is an open set, as each individual set is open. Similarly, $\forall_{i=1}^{\infty}(a,a+i)$ is open, as it is the union of many individual sets. Thus, as we have $\mathbb{R}-[a,b]$ is open, $[a,b]$ is closed. $\square$

§19 Wednesday, October 12, 2022

Today, we'll continue discussing the topology of $\mathbb{R}^n$.

Theorem 19.1

A subset A of $\mathbb{R}^n$ is **closed** if and only if it contains all of its limit points.

Theorem 19.2

Closed subsets in $\mathbb{R}^n$ satisfy the following properties:

(a) $\emptyset$ and $\mathbb{R}^n$ are closed

(b) The union of any finite number of closed sets is closed

(c) The intersection of any collection of closed sets is closed

Theorem 19.3

Let $f:\mathbb{R}^n\to\mathbb{R}^m$ be a function.

1. f is continuous if and only if given any open subset U of $\mathbb{R}^m$, the inverse image $f^{-1}(U)$ is an open subset of $\mathbb{R}^n$.
2. f is continuous if and only if given any closed subset C of $\mathbb{R}^n$, the inverse image $f^{-1}(C)$ is a closed subset of $\mathbb{R}^n$.

Example 19.4

Prove that the circle x^2+y^2 is a closed subset of $\mathbb{R}^2$.

Example 19.5

Prove that every closed ball in $\mathbb{R}^n$ is a closed subset of $\mathbb{R}^n$.

Definition 19.6. A subset A of $\mathbb{R}^n$ is called **compact** if every infinite subset of A has a limit point which lies in A.

Example 19.7

Prove that $\mathbb{R}$ and $(0,1)$ are not compact.

Definition 19.8. A subset A of $\mathbb{R}^n$ is said to be **bounded** if and only if it is inside an open ball centered at the origin.

Theorem 19.9

A subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed.

Theorem 19.10 (Extreme Value Theorem)

Suppose A is a compact subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ is continuous. Then, f attains its maximum and minimum values. In other words, there exist $x_0, y_0 \in A$ such that $f(x_0) \le f(x) \le f(y_0)$ for all $x \in \mathbb{A}$.

§20 Friday, October 14, 2022

Now, we'll start discussing curves in $\mathbb{R}^n$.

§20.1 Curves in $\mathbb{R}^n$

Recall that for a function $f : \mathbb{R} \to \mathbb{R}$, we define its derivative at a by

$$f'(a) = \lim_{h\to 0} \tfrac{f(a+h)-f(h)}{h}$$

This can also be written as

$$\lim_{h\to 0} \tfrac{f(a+h)-f(a)-f'(a)h}{h} = 0$$

In other words, the value of $f(a+h) - f(a)$ is very close to $f'(a)h$, where h is small. Note that $f'(a)h$ is a linear function in terms of r.

Definition 20.1. Given a function $f : I \to \mathbb{R}^n$, where $I \subseteq \mathbb{R}$ is an open interval, the **derivative** of f at point $a \in I$ is given by

$$f'(a) = \lim_{h\to 0} \tfrac{f(a+h)-f(a)}{h}$$

If this limit does not exist, we say f is not **differentiable** at a. The n-th derivative of f at a, denoted by $f^n(a)$ is recursively defined as the derivative of f^{n-1} at a. Note that for the n-th derivative of f to exist at a, the $n-1$st derivative of f must exist on an open interval centered at a.

Theorem 20.2

Suppose $f = (f_1, \ldots, f_n) : I \to \mathbb{R}^n$ is a function, where $I \subseteq \mathbb{R}^n$ is an open interval. Then, f is differentiable at a point $a \in I$ if and only if f_j is differentiable at a for all j, $j = 1, \ldots, n$. Furthermore, if f is differentiable at a, then $f'(a) = (f_1'(a), \ldots, f_n'(a))$.

Theorem 20.3 (Properties of derivatives)

Let a be a number in an open interval I. Suppose $f, g : I \to \mathbb{R}^n$, and $\varphi : I \to \mathbb{R}$ are differentiable at a. Then,

a) $(f + g)'(a) = f'(a) + g'(a)$

b) $(f \cdot g)'(a) = f'(a) \cdot g(a) = f(a) \cdot g'(a)$ (recall that $\cdot$ denotes the standard inner product of $\mathbb{R}^n$)

c) $(\varphi f)'(a) = \varphi'(a) f(a) = \varphi(a)' f(a)$ (note that (φf) refers to the function f being multiplied by φ, NOT the standard inner product of $\mathbb{R}^n$)

Proof. Algebraic manipulation + definition of derivative. □

Theorem 20.4 (Chain Rule)

Suppose I and J are open intervals, $\varphi : I \to J$ is differentiable at $a \in I$, and $f : J \to \mathbb{R}^n$ is differentiable at $\varphi(a)$. Then, $(f \circ \varphi)'(a) = \varphi'(a) f'(\varphi(a))$.

Proof. Algebraic manipulation + definition of derivative. □

Definition 20.5. Let I be an open interval, and $f : I \to \mathbb{R}^n$ a function differentiable at a point $a \in I$. The linear function $L : \mathbb{R} \to \mathbb{R}^n$, defined by $L(h) = f'(a)h$ is denoted by df_a, and is called the **differential** of f at a.

Theorem 20.6

The mapping $f : I \to \mathbb{R}^n$ is differentiable at some $a \in I$, where I is an open interval, if and only if there exists a linear mapping $L : \mathbb{R} \to \mathbb{R}^n$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - L(h)}{h} = 0$$

Furthermore, when such a linear mapping exists, it is unique and $L(h) = f'(a)h$.

Example 20.7

Evaluate the derivative and differential of $f(x) = (\sin x, x^2, x + \cos x)$.

Solution. The derivative of f is the derivative of all of its coordinate functions, which is $(\cos x, 2x, 1 - \sin x)$. The differential of f is the differential of all of its coordinate functions, which is $(\cos x\ h, 2x\ h, 1 - \sin x\ h)$. □

Remark 20.8. Consider the identity function $x : \mathbb{R} \to \mathbb{R}$. We have $dx_a(h) = 1h = h$. If $\varphi : I \to \mathbb{R}$ is differentiable at a point $a \in I$, then $d\varphi_a = \varphi'(a)h$, which means $d\varphi_a(h) = \varphi'(a)dx_a(h)$, or $d\varphi_a = \varphi'(a)dx_a$. This is quite similar to the notion $\varphi'(x) = \frac{d\varphi}{dx}$.

If $f : \mathbb{R} \to \mathbb{R}^n$ is differentiable at $\varphi(a)$, then

$$d(f \circ \varphi)_a(h) = (f \circ \varphi)'(a)h = f'(\varphi(a))\varphi'(a)h = df_{\varphi(a)}(d\varphi_a(h)) = df_{\varphi(a)} \circ d\varphi_a(h)$$

Therefore, we have $d(f \circ \varphi)_a = df_{\varphi(a)} \circ d\varphi_a$.

§21 Monday, October 17, 2022

§21.1 Directional Derivatives

Definition 21.1. Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a function, with $a \in \mathbb{R}^n$ and $\mathbf{0} \leq \mathbf{v} \in \mathbb{R}^n$, where $\mathbf{v}$ is a vector. The **directional derivative** of F with respect to $\mathbf{v}$ at a is

$$D_{\mathbf{v}}F(a) = \lim_{h\to 0} \frac{F(a+h\mathbf{v})-F(a)}{h}$$

When $\mathbf{v} = \mathbf{e}_i$. this directional derivative is denoted by

$$D_{\mathbf{e}_i}F(a) = D_iF(a) = \frac{\partial F}{\partial x_i}(a) = F'_{x_i}(a)$$

Example 21.2

Evaluate the partial derivatives (with respect to both x and y) of $F(x, y) = x^2 - xy - y^3$.

Solution. $\frac{\partial F}{\partial x} = 2x + 1$ and $\frac{\partial F}{\partial y} = -1 - 3y^2$. □

Example 21.3

Evaluate the directional derivative of the following function with respect to the vector $(1, 2)$ at the origin:

$$F(x, y) = \begin{cases} \frac{x^2}{x^2+y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{otherwise} \end{cases}$$

Solution. Fill this in ASAP. □

Theorem 21.4
Let U be an open subset of $\mathbb{R}^n$. Given a function $F : U \to \mathbb{R}^m$, a vector $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$, a point $a \in U$, and $0 \neq c \in \mathbb{R}$, we have $D_{c\mathbf{v}}F(a) = cD_{\mathbf{v}}F(a)$.

Proof. Fill this in ASAP. Follows from the definition of the directional derivative. ☐

We know from the definition of the directional derivative that

$$\frac{F(a+h\mathbf{v})-F(a)-hD_{\mathbf{v}}F(a)}{h} = 0$$

This brings us to the following definition:

Definition 21.5. Let a be a point in an open subset U of $\mathbb{R}^n$. We say $F : U \to \mathbb{R}^n$ is **differentiable** at a if there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{F(a+h)-F(a)-L(h)}{||h||} = \mathbf{0}$$

Theorem 21.6
The linear transformation L in Definition 21.5 is unique.

Proof. Suppose not, and that we have L_1 and L_2 both satisfying the conditions of the previous definition. By Definition 21.5, we have $\lim_{h\to 0} \frac{F(a+h)-F(a)-L_1(h)}{||h||} = \lim_{h\to 0} \frac{F(a+h)-F(a)-L_2(h)}{||h||} = \mathbf{0}$. If $L_1(h) = L_2(h)$, we have

$$\lim_{h\to 0} \frac{L_2(h)-L_1(h)}{||h||} = \mathbf{0}$$

Let $\mathbf{v} \in \mathbb{R}^n$, and $h = t\mathbf{v}$ and let $t \to 0$. This means that h will also go to 0. This gives us

$$\begin{aligned}
\lim_{t\to 0} \frac{L_2(t\mathbf{v}) - L_1(t\mathbf{v})}{||t\mathbf{v}||} &= \mathbf{0} \\
&= \lim_{t\to 0^+} \frac{t(L_2(\mathbf{v}) - L_1(\mathbf{v}))}{t||\mathbf{v}||} \\
&= \frac{L_2(\mathbf{v}) - L_1(\mathbf{v})}{||\mathbf{v}||}
\end{aligned}$$

This means that $L_2(\mathbf{v}) = L_1(\mathbf{v})$, which is a contradiction. Note that we don't need to show that this also holds for $t \to 0^-$, as we only used the fact $t \to 0^+$ in the beginning of the proof.

☐

Definition 21.7. The linear transformation in Definition 21.5 and Theorem 21.6 is called the **differential** of F at a, and is denoted by dF_a. Its matrix is called the **derivative** of F at a, and is denoted by $F'(a)$.

Remark 21.8. Note that if $F : U \to \mathbb{R}^m$, where U is an open subset of $\mathbb{R}^n$ and $a \in U$, then $dF_n : \mathbb{R}^n \to \mathbb{R}^m$ is linear, and its matrix $F'(a)$ is an $m \times n$ matrix. We also have $dF_a(\mathbf{h}) = F'(a)h$, where $\mathbf{h}$ is a column vector in $\mathbb{R}^n$.

Theorem 21.9

Let a be a point in an open subset U of $\mathbb{R}^n$. If $F = (F_1, \dots, F_m) : U \to \mathbb{R}^m$ is differentiable at a, then

$$D_{\mathbf{v}}F(a) = dF_a(\mathbf{v}) = F'(a)\mathbf{v}$$

Furthermore, the (i, j) entry of $F'(a)$ is $\frac{\partial F_i}{\partial x_j}(a)$. A matrix representing $F'(a)$ is in Dr. Ebrahimian's notes.

Proof. We will prove the first part of the theorem, e.g. $D_{\mathbf{v}}F(a) = dF_a(\mathbf{v}) = F'(a)\mathbf{v}$. By the definition of the directional derivative, we have $D_{\mathbf{v}}F(a) = \frac{F(a+h)-F(a)-F(a)h}{||h||} = 0$. Substitute $t\mathbf{v}$ with h and let $t \to 0$. Then, we have

$$\lim_{t\to 0} \frac{F(a+tv)-F(a)-F'(a)t\mathbf{v}}{||tv||} = \mathbf{0}.$$

Note that because $||v||$ is a constant, we can multiply both sides by it and get rid of it. If $t \to 0^+$, $\lim_{t\to 0^+} \frac{F(a+t\mathbf{v})-F(a)-tF'(a)\mathbf{v}}{t} = \mathbf{0}$. We can split this equation to see that

$$\lim_{t\to 0^+} \frac{F(a+t\mathbf{v})}{t} - \frac{tF'(a)\mathbf{v}}{t} = 0 \to D_{\mathbf{v}}F(a) = F'(a).$$

which completes the proof for $t \to 0^+$. The proof for $t \to 0^-$ is similar, and we also get $F'(a)\mathbf{v}$, meaning that $D_{\mathbf{v}}(a) = F'(a)\mathbf{v}$. $\square$

Definition 21.10. The matrix in Theorem 21.9 is called the **Jacobian** matrix of F and a.

§22 Wednesday, October 19, 2022

Let's start with an example involving a Jacobian matrix.

Example 22.1

Assume we know $f : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $F(x, y) = (x^2 + y, x - 1, y^2)$ is differentiable everywhere. Find its derivative $F'(1, 2)$ and differential $dF_{(1,2)}$, as well as $D_{(2,3)}F(1, 2)$.

Solution. We will first find the Jacobian of f. Let $F_1 = x^2 + y$, $F_2 = x - 1$, and $F_3 = y^2$. The matrix is $\begin{pmatrix} \frac{\partial F_1}{x_1} & \frac{\partial F_1}{y_1} \\ \frac{\partial F_2}{x_2} & \frac{\partial F_2}{y_2} \\ \frac{\partial F_3}{x_3} & \frac{\partial F_3}{y_3} \end{pmatrix}$, which, after computing each partial derivative, is $\begin{pmatrix} 2x & 1 \\ 1 & 0 \\ 0 & 2y \end{pmatrix}$. Now, we can evaluate this at the point $(1, 2)$. This is equivalent to $\begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 4 \end{pmatrix}$, which is the derivative of F at $(1, 2)$. To compute the differential $dF_{(1,2)}$, note that the differential is intrinsically a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$. Then, we have $dF_{(1,2)}(x, y) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x \\ 4y \end{pmatrix}$. To find $D_{(2,3)}F(1, 2) = F'(1, 2) \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \\ 12 \end{pmatrix}$. Alternatively, one could note $D_{(2,3)}F(1, 2) dF_{(1,2)} = \begin{pmatrix} 7 \\ 2 \\ 12 \end{pmatrix}$. □

Corollary 22.2

Suppose U is an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is a function that is differentiable at some $a \in U$. Then, for every $0 \neq \mathbf{v} \in \mathbb{R}$, we have

$$D_{\mathbf{v}}(a) = \mathbf{v} \cdot (D_1 f(a), \ldots, D_n f(a)).$$

By definition of the directional derivative, we have $D_{\mathbf{v}} f(a) = f'(a)\mathbf{v}$. Assume $v = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix}$. Now, note that the derivative of f is a $1 \times n$ matrix; it is equivalent to

$$\left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right) = f'(a) \cdot \mathbf{v}$$

Definition 22.3. In the case above, we call $f'(a)$ the **gradient** of f at a, and is denoted by $\nabla f(a)$.

Definition 22.4. A **direction** is a unit vector $\mathbf{u}$. The directional derivative of a function F in the direction of a nonzero vector $\mathbf{v}$ at point a is $D_{\mathbf{u}}F(a)$, where $\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||}$.

Theorem 22.5

Let $a \in U$, where U is an open subset of $\mathbb{R}^n$. Suppose $f : U \to \mathbb{R}$ is differentiable at a and that $\nabla f(a) \neq 0$. Then, the maximum directional derivative of f at a is in the direction of $\nabla f(a)$, and this maximum directional derivative is equal to $||\nabla f(a)||$. Similarly, the minimum directional derivative of f at a is in the direction $-\nabla f$, and this minimum directional derivative is equal to $-||\nabla f(a)||$.

Example 22.6

Find the maximum and minimum directional derivative of the function $f : \mathbb{R}^3 \to \mathbb{R}$, given by $f(x, y, z) = \sin(xyz) + x^2 + yz$ at $(2\pi, 1, 3)$.

Solution. $\nabla f = (yz\cos(xyz) + 2x, xz\cos(xyz) + z, xy\cos(xyz) + y)$. At $(2\pi, 1, 3)$, $\nabla f = (3 + 4\pi, 6\pi + 3, 2\pi + 1)$. By Theorem 22.7, maximum directional derivative of f at a is equal to $||\nabla f||$ and the minimum directional derivative is at $-||\nabla f||$. We can compute both of these values. □

Definition 22.7. Suppose U is an open subset U of $\mathbb{R}^n$. A function $F : U \to \mathbb{R}^m$ is said to be **continuously differentiable** at a if all partial derivatives $D_1F, D_2F, \ldots, D_nF$ exist on U and they are all continuous at a.

Theorem 22.8

If F is continuously differentiable at a point $a \in U$, then F is differentiable at a.

Example 22.9

Prove $F(x, y) = (x^2 + y, 2xy, y^2 - x)$ is differentiable on $\mathbb{R}^2$.

Proof. We can check all of the partial derivatives of F to see if F is continuously differentiable. As each of these partial derivatives exist (the work is trivial, and omitted here), and are continuous everywhere, by Theorem 22.8, F is differentiable. □

Definition 22.10. Suppose U is an open subset of $\mathbb{R}^n$ and $F : U \to \mathbb{R}^n$ is differentiable. A point $a \in U$ is called a **critical point** of f if $\nabla f(a) = 0$.

Definition 22.11. Let $f : U \to \mathbb{R}$ be a function, where U is an open subset of $\mathbb{R}^n$. We say f attains a **local minimum** (with respect to a **local maximum**) at a if there is an open subset V of U for which $f(a) \leq f(x)$ for all $x \in V$. If f has a local maximum or local minimum at a, we say f has a **local extremum** at a.

Theorem 22.12

If f attains a local minimum or maximum at a point $a \in U$, a is a critical point of f.

Definition 22.13. Let U be an open subset of $\mathbb{R}^n$ and let $F : U \to \mathbb{R}^m$ be differentiable. Suppose $a \in U$. Then, the approximation

$$F(x) \approx F(a) + dF_a(x - a)$$

is called the **linear plane approximation** of F near a.

Example 22.14

Approximation $\sqrt{1.95 \times 2.01 \times 4.01}$ using the tangent plane approximation.

Solution. We can let $f(x,y,z) = \sqrt{xyz}$, where $f : (0,\infty)^3 \to \mathbb{R}$. Now we must prove that f is differentiable. We can do this by taking each of the partial derivatives of f and showing that these partial derivatives are continuous to see f is continuously differentiable, and therefore differentiable. Now, note that $f'(2,2,4) = (1,1,\frac{1}{2})$. Thus, $f(1.95, 2.01, 4.01) \approx f(2,2,4) + \left(1, 1, \frac{1}{2}\right)\begin{pmatrix} -0.05 \\ 0.01 \\ 0.01 \end{pmatrix} = 4 + (-0.05 \cdot 1, 0.01 \cdot 1, 0.01 \cdot \frac{1}{2}). = 3.965.$ □

§23 Friday, October 21, 2022

Given a function $f : U \to \mathbb{R}$, where U is an open subset of $\mathbb{R}^n$ and $a \in U$, we have the following:

$$df(\mathbf{h}) = D_{\mathbf{h}}f = \nabla f \cdot \mathbf{h} = \sum D_i f h_i = D_i f dx_i(\mathbf{h})$$

Therefore, we can write $df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dx_i$.

Definition 23.1. Let $f : U \to \mathbb{R}$ be a differentiable function, where U is an open subset of $\mathbb{R}^n$. The mapping L given by $L(a) = df_a$, which assigns any point a to the linear mapping $df_a : \mathbb{R}^n \to \mathbb{R}^m$, is called a **differential form**.

§23.0.1 The Chain Rule

Theorem 23.2 (Chain Rule)

Suppose U and V are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Suppose $F : U \to \mathbb{R}^m$ and $G : V \to \mathbb{R}^k$ are differentiable at points $a \in U$ and $F(a) \in V$, respectively. Then, their composition $H = G \circ F$ is differentiable at a and $dH_a = dG_{F(a)} \circ dF_a$. Furthermore, $H'(a) = G'(F(a))F'(a)$.

Example 23.3

Write the chain rule for functions $f = (f_1, \ldots, f_m) : \mathbb{R} \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}$.

Solution. Note that both $f \circ g$ and $g \circ f$ exist. $(g \circ f)' = \frac{d}{dx}(g \circ f)(a) = g'(f(a))f'(a)$. Note that the size of $g'(f(a))$ is $1 \times m$ and the size of $f'(a)$ is $m \times 1$. Finish this ASAP. □

Example 23.4

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the usual polar coordinate mapping defined by $T(r,\theta)$ defined by $(r\cos\theta, r\sin\theta)$. For a function $f(x,y)$ from the Cartesian plane $\mathbb{R}^2 \to \mathbb{R}$, find the partial derivatives of $f(r\cos\theta, r\sin\theta)$ with respect to r and θ.

Definition 23.5. Given two points $a, b \in \mathbb{R}^n$, the **segment** L from a to b is the set given by

$$L = \{c \in \mathbb{R}^n | c = tb + (1-t)a \text{ where } 0 \le t \le 1\}$$

Definition 23.6. An open subset U of $\mathbb{R}^n$ is called **connected** if for every $a, b \in U$ there is a differentiable function $\varphi : \mathbb{R} \to U$ such that $\varphi(0) = a$ and $\varphi(1) = b$.

Definition 23.7. A function $F : U \to \mathbb{R}^m$ is called **constant** if there is a $c \in \mathbb{R}^m$ for which $F(x) = c$ for all $x \in U$.

§24 Monday, October 24, 2022

Theorem 24.1
Let U be an open and connected subset of $\mathbb{R}^n$. A differentiable function $F : U \to \mathbb{R}^m$ is constant if and only if $F'(x) = 0$ for all $x \in U$.

Proof. The forward direction is relatively straightforward. If $F(x) = c$, where $c \in \mathbb{R}^m$, we want to show

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{||h||} = 0$$

As $F(x+h) = (c_1, c_2, \ldots, c_m)$, we have $F(x+h) - F(x) = c - c = 0$.

To prove the backwards direction, suppose we have $F'(x) = 0$ for all $x \in U$. Consider the points $a < x_1 < x_2 < b$. Since F' exists on (a, b), we have x is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2). Therefore, by the Mean Value Theorem (which we will introduce next), we have that there exists a c such that

$$F'(c) = \tfrac{F(x_2) - F(x_1)}{x_2 - x_1}$$

By assumption, we have $F'(c) = 0$, which implies

$$\tfrac{F(x_2) - F(x_1)}{x_2 - x_1} = 0 \to F(x_1) = F(x_2)$$

Since x_1 and x_2 were arbitrary, we can conclude that $F(x_1) = F(x_2)$ for all $x_1, x_2 \in U$, meaning $F(x)$ is constant over the interval U. □

Theorem 24.2 (Mean Value Theorem)
Suppose U is an open subset of $\mathbb{R}^n$, and a, b are two points in U such that U contains the line segment L from a to b. If $f : U \to \mathbb{R}$ is differentiable, then there is a point $c \in L$ for which

$$f(b) - f(a) = f'(c)(b - a) = \nabla f(c) \cdot (b - a)$$

Example 24.3

Find all second partial derivatives of $f(x, y) = x^3 - 5y\ln(xy)$.

Solution. $f_x = 3x^2 - 5y\frac{1}{xy} \cdot y = 3x^2 - \frac{5y}{x}$. $f_{xx} = 6x + \frac{5y}{x^2}$. $f_{xy} = (f_x)_y = -\frac{5}{x}$. $f_y = -5y\ln(xy) - 5y\frac{1}{xy} = -5\ln(xy)$. □

Theorem 24.4 (Clairaut's Theorem)

Suppose U is an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ has continuous first and second partial derivatives. Then, for every i, j, we have $D_jD_if(a) = D_iD_jf(a)$ for all $a \in U$.

Remark 24.5. Note that $f_{xy} = (f_x)_y$, and $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$.

Example 24.6

Let $f(x, y)$ with $f : \mathbb{R}^2 \to \mathbb{R}$ be a function with continuous first an second partial derivatives and let $g(u, v) = f(Au + Bv, Cu + Dv)$, where A, B, C, D are constants. Prove that

$$\frac{\partial^2 g}{\partial u \partial v} = AB\frac{\partial^2 f}{\partial x^2} + CD\frac{\partial^2 f}{\partial y^2} + (AD + BC)\frac{\partial^2 f}{\partial x \partial y}$$

Solution. Let $x = Au+Bv$ and $y = Cu+Dv$. Then, note that f is a function of x and y, and x and y are each functions of u and v. Note that $\frac{\partial g}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} = A\frac{\partial f}{\partial x} + C\frac{\partial f}{\partial y}$. □

§24.1 Critical Points in Two Dimensions

We would like to classify critical points of a function $f : \mathbb{R}^2 \to \mathbb{R}$. Recall that if a point is a local extremum, then it must be a critical point. Let's first look at a simple case when

$$f(x, y) = ax^2 + 2bxy + cy^2, \text{ where } a, b, c \text{ are constants.}$$

Such a function is called a **quadratic form**. We note that $(0, 0)$ is a critical point of this function, and $f(0, 0) = 0$. Now, we ask: under what conditions on a, b, c can we guarantee that $f(x, y) \geq 0$ for points (x, y) near the origin? Completing the square, we obtain the following:

$$f(x, y) = \frac{(ax+by)^2 + (ac-b^2)y^2}{a}$$

This gives the following:

- If $a > 0$ and $ac - b^2 \geq 0$, then $f(x, y)$ has a local (and absolute) minimum at $(0, 0)$.

- If $a < 0$ and $ac - b^2 \geq 00$, $f(x, y)$ has a local (and absolute) maximum at $(0, 0)$¿ If $ac - b^2 < 0$, $f(x, y)$ has neither a local minimum nor a local maximum at $(0, 0)$.

Definition 24.7. A **quadratic form** is a function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \text{ where } a_{ij} \in \mathbb{R} \text{ is a constant}$$

Definition 24.8. A quadratic form $f(x)$ is called **positive-definite** or **negative-definite** if $f(x) > 0$ or $f(x) < 0$ (respectively) for all $0 \neq x \in \mathbb{R}^n$. It is called **nondefinite** if it has both positive and negative values. The above discussion gives us the following theorem:

Theorem 24.9

The quadratic form $f(x, y) = ax^2 + 2bxy + cy^2$ is

- positive-definite if $a > 0$ and $ac - b^2 > 0$
- negative-definite if $a < 0$ and $ac - b^2 > 0$
- nondefinite if $ac - b^2 < 0$

Example 24.10

Determine and classify all critical points of $f(x, y) = x^2 - y^2$.

Solution. $f_x = 2x$ and $f_y = -2y$. Hence, $(0, 0)$ is the only critical point. $f(0, 0)$ is both positive and negative near 0. For every ball centered at the origin, you can find something in the ball that is positive, and something else that is negative. For example, $f\left(\frac{r}{2}, 0\right) = \frac{r^2}{4} > 0 = f(0, 0)$, but $f\left(0, \frac{r}{2}\right) = -\frac{r^2}{4} < f(0, 0)$, meaning $f(0, 0)$ is neither a local maximum nor minimum. It is a saddle point, which we will now define. □

Definition 24.11. A critical point a of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is called a **saddle point** if every open ball containing a contains points x and y for which $f(x) < f(a) < f(y)$.

§25 Wednesday, October 26, 2022

Theorem 25.1 (Second Partials Test)
Let U be an open subset of $\mathbb{R}^2$ and $f : U \to \mathbb{R}$, where f is twice continuously differentiable. Suppose $a \in U$ is a critical point of f. Let

$$\nabla = \frac{\partial^2 f}{\partial x^2}(a) \cdot \frac{\partial^2 f}{\partial y^2}(a) - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 (a)$$

Then, f has

- a local minimum at a if $\nabla > 0$ and $\frac{\partial^2 f}{\partial x^2}(a) > 0$
- a local maximum at a if $\nabla > 0$ and $\frac{\partial^2 f}{\partial x^2}(a) < 0$
- a saddle point at a if $\nabla < 0$

Note that if $\nabla = 0$, the test is inconclusive.

Remark 25.2. ∇ is equal to the determinant of the matrix

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}$$

This matrix is called the **Hessian matrix**, and its determinant is called the **Hessian determinant**. Note that $\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$

Example 25.3
Classify all critical points of $f(x, y) = xy + 2x - y$.

Solution. We have $f_x = y + 2$ and $f_y = x - 1$. Setting both of these equal to 0 and solving, we have $(1, -2)$ is the only critical point. Now, we can apply the second partials test. $f_{xx} = 0$ and $f_{xy} = 1$, and $f_{yy} = 0$. We have the that $\nabla = \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1 < 0 \to (1, -2)$ is a saddle point. $\square$

Remark 25.4. To understand quadratic forms, note that for a quadratic form $f(x_1, \ldots, x_n)$, we have

$$f(cx_1, \ldots cx_n) = c^2 f(x_1, \ldots, x_n)$$

Thus, in order to understand if the origin is a local minimum or local maximum, we need to understand f over the unit sphere $x_1^2 + \ldots + x_n^2$.

§25.1 Lagrange Multipliers

Theorem 25.5

Let S be a subset of $\mathbb{R}^n$. Assume f is a differentiable real-valued function defined on some open set containing S, and f has a local maximum (or a local minimum) on S at a (i.e. $f(x) \geq f(a)$ for all $x \in S \cap B_r(a)$ if a is a local minimum, and $f(x) \leq f(a)$ if a is local maximum). Then, the gradient vector $\nabla f(a)$ is orthogonal to all tangent lines to all curves on S that pass through a. In other words, if $\varphi : \mathbb{R} \to S$ is a differentiable curve with $\varphi(0) = a$, then $\nabla f(a)$ is orthogonal to $\varphi'(0)$.

Proof. Suppose $f \circ \varphi(t)$ (from $\mathbb{R} \to \mathbb{R}$) has a local minimum a $t = 0$. Then, $(f \circ \varphi)'(0) = 0$. Now, we have $(f \circ \varphi)'(0) = f'(\varphi(0))\varphi'(0) = 0 \to \nabla f(a) \cdot \varphi(a) = 0 \to \nabla f(a) \perp \varphi'(0)$, completing the proof (the case for the local maximum is identical). $\square$

Example 25.6

Find the maximum and minimum values of $f(x, y) = xy$ subject to the constraint $x^2 + y^2 = 1$.

Solution. Before we find the maximum and minimum values, we need to ensure that maximum and minimum values actually exist. We can try to use the Extreme Value Theorem. Note that f is continuous (as it is a polynomial). Now, we need to prove f is compact. Note that $f = \{(x, y) \mid x^2 + y^2 = 1\} = g'(\{1\})$, where $g(x, y) = x^2 + y^2$. Therefore, f is closed, since $f^{-1}(\{1\})$ (the pre-image of f under $\{1\}$ is closed. Since we also have $\{(x, y) \mid x^2 + y^2 = 1\} \subseteq B_2(0, 0)$, we have that f is bounded. Therefore, f is compact. By the Extreme Value Theorem, the maximum and minimum values of f exist. Now, note that $\nabla f = (y, x)$. If (a, b) is an extremum for f on the unit circle, by Theorem 25.5, (b, a) is perpendicular to the tangent line. As the radius of $x^2 + y^2 = 1$ is perpendicular to the tangent line of the circle (by geometric properties of the radius), we have $(b, a) = c(a, b)$ (e.g. the gradient is a multiple of (a, b). This is because the gradient is perpendicular to the tangent line, and the radius is also perpendicular to the tangent line; as there is only one line perpendicular to the tangent line, we have that the gradient is a scalar multiple of the tangent line.

To find a and b, we can use the following system of equations:

$$\begin{cases} b = ca \\ a = cb \\ a^2 + b^2 = 1 \end{cases}$$

Note that we obtain $a^2 + b^2 = 1$ because (a, b) is on the unit circle. Substituting in $b = c^2 b \to b(1 - c^2)$, we get $b = 0$ or $c = \pm 1$.

We now have the following cases:

- $b = 0 \to a^2 = 1 \to a = \pm 1 \to f(\pm 1, 0) = 0$

- $c = 1 \to a = b \to a = b = \pm\frac{1}{\sqrt{2}} \to f\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right) = \frac{1}{2}$
- $c = -1 \to a = -b \to f\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right) = -\frac{1}{2}$

Thus, our minimum is $-\frac{1}{2}$ and maximum is $\frac{1}{2}$. □

Example 25.7

Find the equation of the plane tangent to the surface $x^2 + 2y^2 + 3z^2 = 6$ at $(1, -1, 1)$.

Solution. First, we must prove $f(x, y, z) = x^2+2y^2+3z^2$ actually attains its minimum and maximum values at $(1, -1, 1)$. After doing this, we note that by Theorem 25.5, all tangent lines to S are orthogonal to $\nabla f(1, -1, 1)$. Let (x, y, z) be on the tangent plane. We have $\big((x, y, z) - (1, -1, 1)\big) \perp \nabla f(1, -1, 1)$. Now, note that $\nabla f = (2x, 4y, 6z)$. Thus, $\nabla f(1, -1, 1) = (2, -4, 6)$. Therefore, $(x - 1, y + 1, z - 1) \cdot (2, -4, 6) = 0 \to 2(x - 1) - 4(y + 1) + 6(z - 1) = 0$. □

Definition 25.8. A **k-dimensional manifold** (or a k-**manifold**) M is a subset of $\mathbb{R}^n$ for which for every point $a \in M$, there is an open subset U of $\mathbb{R}^n$ containing a for which $U \cap M$ "looks like" the k-dimensional space $\mathbb{R}^k$.

Example 25.9

A sphere in $\mathbb{R}^3$ is a 2-dimensional manifold.

§26 Friday, October 28, 2022

Theorem 26.1

If M is a k-dimensional manifold in $\mathbb{R}^n$ and $a \in M$, then M has a k-dimensional tangent plane at a. In other words, all lines tangent to curves on M at a that pass through a form the translation of a k-dimensional subspace of $\mathbb{R}^n$.

Theorem 26.2

Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. If M is the set of all points x with both $g(x) = 0$ and $\nabla g(x) \neq 0$, then M is an $(n - 1)$-manifold. Given $a \in M$, the gradient vector $\nabla g(a)$ is orthogonal to the tangent plane to M at a.

Theorem 26.3 (Lagrange Multipliers Theorem, simplified)
Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, and let M be the set of all points $x \in \mathbb{R}^n$ such that both $g(x) = 0$ and $\nabla g(x) \neq 0$. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Assume f attains a local maximum or minimum on M at a point $a \in M$. Then, $\nabla f(a) = \lambda \nabla g(a)$ for some scalar λ.

Example 26.4
Find the maximum and minimum values of $f(x, y, z) = x + 3y + z$ under the constraint $x^2 + y^2 + z^2 = 1$.

Solution. $f(x, y, z) = x + 3y + z$ is continuously differentiable, and therefore differentiable since it is a polynomial. The surface $S' = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ is closed, since $S' = g^{-1}(\{0\})$, and $\{0\}$ is a finite set. Also, S' is bounded, since $S' \subseteq B_2(0)$. Thus, as S is compact and g is continuous, g will attain its maximum and minimum values, by the Extreme Value Theorem. Now, we have $\nabla g = (2x, 2y, 2z) = 0 \to x = y = z = 0$. However, as $(0, 0, 0) \neq S'$, by the Lagrange Multipliers Theorem, we have $\nabla f = \lambda \nabla g$. We now have the following system:

$$\begin{cases} 1 = 2\lambda x \\ 3 = 2\lambda y \\ 1 = 2\lambda z \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

Solving this system, we obtain $(x, y,) = \pm \left(\frac{\sqrt{11}}{11}, \frac{3\sqrt{11}}{11}, \frac{\sqrt{11}}{11} \right) \to$ maximum value of f is $\sqrt{11}$ and minimum value is $-\sqrt{11}$. $\square$

Theorem 26.5 (Lagrange Multipliers Theorem, rigorous)
Suppose $G = (G_1, \ldots, G_m) : \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable, and let M be the set of all points $x \in \mathbb{R}^n$ such that $G(x) = 0$ and the gradient vectors $\nabla G_1(a), \ldots, \nabla G_m(a)$ are linearly independent. Assume the differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ attains a local minimum or maximum on M at $a \in M$, then $\nabla f(a)$ is a linear combination of $\nabla G_1(a), \ldots, \nabla G_m(a)$.

Example 26.6
Find the highest and lowest points of the ellipse of intersection of the cylinder $x^2 + y^2 = 1$ and the plane $x + y + z = 1$.

Solution. Drawing a diagram of this intersection helps immensely with this solution. We have $G(x, y, z) = (x^2 + y^2 - 1, x + y + z - 1)$ and $f(x, y, z) = z$. Thus, $M = \{(x, y, z) \in \mathbb{R}^3 \mid G(x, y, z) = (0, 0)\}$. We know that G attains its maximum and minimum values because $M = G^{-1}(\{0,0\})$ and $\{0,0\}$ is a closed set, meaning M is closed. As $M \subseteq B_2(0,0,0)$ and G is a continuous function, we have G attains its maximum and minimum values, by the Extreme Value Theorem. Now, we have $\nabla(x^2+y^2-1) = (2x,2y,0)$ and $\nabla(x+y+z-1) = (1,1,1)$. Now, let λ_1 and λ_2 be constants. We have have $\lambda_1(2x,2y,0)+\lambda_2(1,1,1) = (0,0,0) \to (2\lambda_1 x+\lambda_2, 2\lambda_1 y+\lambda_2, \lambda_2) = (0,0,0) \to \lambda_2 = 0$ and $2\lambda_1 x = 2\lambda_1 y = 0$. If $\lambda_1 = 0$,then $x = y = 0$, which does not satisfy $x^2+y^2 = 1$. Thus, we can solve the system

$$\begin{cases} (0,0,1) = \lambda_1(2x,2y,0)+\lambda_2(1,1,1) \\ x^2+y^2 = 1 \\ x+y+z = 1 \end{cases}$$

and plug the solutions of the system back into f to find its maximum and minimum points. □

§27 Monday, October 31, 2022

The Second Partials test for a function $f : \mathbb{R}^2 \to \mathbb{R}$ allows us to determine if a critical point is a local minimum, a local maximum, or a saddle point. Now, we will turn our focus to functions $f : \mathbb{R}^n \to \mathbb{R}$.

Definition 27.1. A matrix A is called symmetric if $A^T = A$. In other words, if the (i,j) entry of A is the same as its (j,i) entry for all i,j.

Remark 27.2. Any quadratic form $\sum_{1\le i\le j\le n} a_{ij}x_ix_j$ can be written as $q(x) = x^TAx$ for a symmetric matrix A, where x is a column vector, and the (i,j) entry of A is $\frac{a_{ij}}{2}$ or $\frac{a_{ji}}{2}$, depending on whether $i<j$ or $j<i$, and the (i,i) entry of A is a_{ii}.

Example 27.3

Write the quadratic form of the below in the form $q(x) = x^TAx$:

$$q(x,y,z) = x^2+2y^2-z^2+3xy+xz-yz$$

Solution. We will start by constructing a matrix A. This matrix is

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

Now, we have $x^TA = \begin{pmatrix} x+\frac{3}{2}y+\frac{1}{2}z & \frac{3}{2}x+2y-\frac{1}{2}z & \frac{1}{2}x-\frac{1}{2}y-z \end{pmatrix}$.

As our "x" in the formula $q(x) = x^TAx$, we have

$$q(x) = \begin{pmatrix} x+\frac{3}{2}y+\frac{1}{2}z & \frac{3}{2}x+2y-\frac{1}{2}z & \frac{1}{2}x-\frac{1}{2}y-z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$ □

Definition 27.4. Given a symmetric $n \times n$ matrix A, the quadratic form $q(x) = x^T Ax$ is called the quadratic form **associated** with A. We also say A is the matrix associated with q. The linear transformation given by $L(x) = Ax$ is called the linear transformation associated with q.

Note that 0 is a critical point of q. Also, for a quadratic form q, scalar c, and a vector $\mathbf{x}$, we have $q(c\mathbf{x}) = c^2 q(\mathbf{x})$. Therefore, to determine if 0 is a local minimum or maximum, we need to determine the maximum and minimum of q over the unit sphere given by $||\mathbf{x}|| = 1$. This can be done using Lagrange Multipliers.

Theorem 27.5

Let q be a quadratic form associated with the $n \times n$ symmetric matrix. If q attains its maximum of minimum values on the unit sphere in $\mathbb{R}^n$ at a point v (with $||v|| = 1$), $Av = \lambda v$ for some $\lambda \in \mathbb{R}$.

Definition 27.6. Given a square matrix A, we say a non-zero vector $\mathbf{v}$ is an **eigenvector** of A if there is a $\lambda \in \mathbb{R}$ for which $A\mathbf{v} = \lambda\mathbf{v}$. The number lambda is called an eigenvalue of A and the pair $(\mathbf{v}, \lambda)$ is called an **eigenpair** of A.

Note that if $(\mathbf{v}, \lambda)$ is an eigenpair of matrix A associated to a quadratic form, $q(\mathbf{v}) = \lambda||\mathbf{v}||^2$.

Theorem 27.7

A real number λ is an eigenvalue of a square matrix A if and only if $\det(A - \lambda I) = 0$, where I is the identity matrix.

Example 27.8

Find the maximum and minimum values of $q(x, y) = 3x^2 + 2y^2 - 2xy$ subject to the condition $x^2 + y^2 = 1$.

Solution. Fill this in ASAP. □

Corollary 27.9

Let A be the matrix associated with a quadratic form q. Then, the maximum and minimum values of $q(x)$, where x is on the unit sphere is the largest and smallest real root of the equation $\det(A - \lambda I) = 0$. If all eigenvalues are positive, $q(x)$ is positive-definite; if all eigenvalues are negative, $q(x)$ is negative definite. If there are both positive and negative eigenvalues, $q(x)$ is non-definite.

Example 27.10

Consider the quadratic form $q(x, y, z) = 2x^2 + 4xy - y^2 + z^2$. Find the maximum and minimum value of this quadratic form over the unit sphere. Determine whether 0 is a local maximum, local minimum, or saddle point.

Solution. $\square$

Definition 27.11. Let A be an $n \times n$ matrix. For every $k \leq n$, we denote the determinant of the upper left-hand $k \times k$ submatrix of A by Δ_k.

Definition 27.12. We say a quadratic form is **positive definite** is $q(x) > 0$ for all non-zero $x \in \mathbb{R}^n$, **negative definite** if $q(x) < 0$ for all non-zero $x \in \mathbb{R}^n$. If q is neither positive nor negative definite, we say q is **non-definite**.

Theorem 27.13

Let $q(x) = x^T Ax$ be a quadratic form whose matrix A is invertible (i.e. $\det A \neq 0$). Then, q is

- positive-definite iff $\Delta_k > 0$ for all k
- negative-definite iff $(-1)^k\Delta_k > 0$ for all k.
- nondefinite iff neither of the previous two conditions is satisfied

Example 27.14

Determine if 0 is local minimum, maximum or saddle point for $q(x, y, z) = x^2 + 5y^2 + z^2 + 4xy + 3xz$.

Solution. First, we must set up the matrix A associated with this quadratic form. This matrix is $\begin{pmatrix} 1 & 2 & \frac{3}{2} \\ 2 & 5 & 0 \\ \frac{3}{2} & 0 & 1 \end{pmatrix}$. To find the eigenvalues, we can try to solve $\det(A - \lambda I) = 0$, but this would result in an ugly cubic that would be rather difficult to solve. However, we don't need to find the eigenvalues of A, and can apply Theorem 27.13. Note: we must make sure the determinant of this matrix is nonzero. We will find each Δ_k for $k \leq n$:

$$\Delta_1 = 1 > 0$$
$$\Delta_2 = 5 - 4 = 1 > 0$$
$$\Delta_3 = (-1)^{3+3} \cdot \det \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} + (-1)^{3+1}\frac{3}{2} \det \begin{pmatrix} 2 & \frac{3}{2} \\ 5 & 0 \end{pmatrix} = 1 - \frac{45}{4} < 0$$

Since $\Delta_3 \neq 0$, we have the $(0,0)$ is a saddle point (e.g. neither a local maximum or minimum). $\square$

To classify a critical point a of a function f, we approximate the function f with a quadratic form and determine if the quadratic form is positive-definite, negative-definite, or nondefinite.

Definition 27.15 (Hessian Matrix)**.** Let U be an open subset of $\mathbb{R}^n$. Suppose $f : U \to \mathbb{R}$ is a function with continuous first, second, and third partial derivatives. The **Hessian matrix** of f at a point $a \in U$ is the $n \times n$ matrix whose (i, j) entry is $D_i D_j f(a)$. The determinant of this matrix is called the **Hessian determinant** of f at a.

Theorem 27.16

Let U be an open subset of $\mathbb{R}^n$. Suppose $f : U \to \mathbb{R}$ is a function with continuous first, second, and third partial derivatives, and let $a \in U$ be a critical point of f. Suppose the Hessian determinant of f at n is nonzero. Then,

- If the Hessian matrix of f at a is positive-definite, then f has a local minimum at a
- If the Hessian matrix of f at a is negative-definite, then f has a local maximum at a
- If the Hessian matrix of f at a is nondefinite, then f has a saddle point at a

Example 27.17

Consider the function $f(x, y, z) = 2x^2 + 5y^2 + 2z^2 + 2xz + x^4 + \sin(y^4)$. Prove $(0, 0, 0)$ is a critical point of f and classify this critical point.

Solution. We have $f_x = 4x + 2z + 4x^2$, $f_y = 10y + 4y^3 \cos(y^3)$, and $f_z = 4z + 2x$. As we have $\nabla f(0, 0, 0) = 0$, $(0, 0, 0)$ is a critical point. To classify this point, we can analyze the Hessian matrix at $(0, 0, 0)$. Our Hessian matrix is $\begin{pmatrix} 4 & 0 & 2 \\ 0 & 10 & 0 \\ 2 & 0 & 4 \end{pmatrix}$. We have

$$\begin{aligned} \Delta_1 &= 4 > 0 \\ \Delta_2 &= 40 > 0 \\ \Delta_3 &= 120 > 0 \end{aligned}$$

Thus, 0 is a local minimum. However, we *cannot* say 0 is a global minimum, because... $\square$

§28 Wednesday, November 2, 2022

§28.1 Area and Volume

Consider a solid E in $\mathbb{R}^3$ that lies between the planes $x = a$ and $x = b$. Suppose the cross-sectional area of the solid at x is given by $A(x)$. Then, the volume of this solid is $\int_b^a A(x)\, dx$.

Now, assume E lies above a rectangle $R = [a, b] \times [c, d]$, and below the graph $z = f(x, y)$. We see that $A(x) = \int_c^d f(x, y)\, dy$. This means that the volume of E is equal to

$$\int_a^b \left(\int_c^d f(x,y)\, dy \right)\, dx$$

Example 28.1

Find the volume of the solid bounded above by the surface $z = xy$ that lies above the rectangle in the xy-plane given by $0 \leq x \leq 1$ and $1 \leq y \leq 2$.

Solution. By what we discussed above, our volume is

$$\begin{aligned}
\int_0^1 \int_1^2 xy\, dydx &= \int_0^1 \frac{xy^2}{2}\Big|_{y=1}^{y=2}\, dx \\
&= \int_0^1 \frac{3}{2}x\, dx \\
&= \frac{3}{4}x\Big|_{x=0}^{x=1} \\
&= \boxed{\frac{3}{4}}
\end{aligned}$$

□

If the region R is bounded, but not a rectangle, we place R inside a rectangle S and defined $f(x, y) = 0$ for every (x, y) that lies in S, but does not lie in R.

Example 28.2

Let R be the triangle in the xy-plane whose vertices are $(0, 0)$, $(1, 0)$ and $(1, 1)$. Evaluate the volume of the solid bounded above by the pane $z = x + y$ that lies above the region R.

§29 Friday, November 4, 2022

Today is a review day for the second midterm exam on Monday, November 7.

§30 Wednesday, November 9, 2022

§30.1 Double Integrals

§30.1.1 Motivation

Definition 30.1. Let $f(x, y)$ be a function over the rectangle $[a, b] \times [c, d]$. A **partition** of R is a collection of rectangles $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ for which $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ is a partition of $[a, b]$ and $c = y_0 < y_1 < \cdots < y_{n-1} < y_n = d$ is a partition of $[c, d]$. Let c_{ij} be a point in the rectangle R_{ij}. Then, the quantity

$$\sum_{i,j=1}^{n} f(c_{ij}) = \Delta A_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} f(x_i, y_i)\delta$$

where $\Delta A_{ij} = \Delta x_i \Delta y_j$ is the area of the rectangle R_{ij} is called a **Riemann sum** of f on R corresponding to this partition of R. f is called **integrable** on R provided the limit of the Riemann sums as $(\Delta x_i, \Delta y_j) \to (0, 0)$ exists and is a real number. The limit of these Riemann sums is denoted by $\iint\limits_R f(x, y)$.

Remark 30.2. The above definition can be written more mathematically using the ϵ-δ definition of the limit.

Theorem 30.3

Let f be a continuous function on a closed rectangle R. Then, f, is integrable.

Example 30.4

Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x, y) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if} x \notin \mathbb{Q} \end{cases}$$

Show that $\iint\limits_{[0,1]^2} dA$ does not exist.

Solution. Consider $0 < x_0 < x_1 < \cdots < x_n = 1$ and $0 = x_0 < x_1 < \cdots < x_n = 1$. In a Week 7 homework problem, we proved that between every two rational numbers is an irrational number. Suppose we have a and b rational, and an irrational number $\frac{r}{\sqrt{2}}$, where r is rational. Then, we have $a\sqrt{2} < r < b\sqrt{2}$, meaning that in between any two irrational numbers, there is a rational number. Now, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} f(r_i, y_i)\Delta A_{ij} = \sum_{i=1}\sum_{j=1} \Delta A_{ij}$$

The second summation above is equivalent to the area of $[0,1]^2$, which is 1. Now, suppose we have an $x \leq s_i \leq$, where s_i is irrational. Then, $\sum_{i=1}\sum_{j=1} f(s_i, y_i)\Delta A_{ij} = 0$, as s_i is irrational. Thus, for every choice of x_i, x_j, there is a Riemann sum that is 1 and one that is 0, meaning f is not integrable. □

Definition 30.5. Let X be a subset of $\mathbb{R}^2$. We say X has **zero area** if for every $\epsilon > 0$ there is a sequence of closed rectangles $R_1, R_2, \ldots$ for which $X \subseteq \bigcup_{n=1}^{\infty} R_n$ and the sums of areas of R_n is less than ϵ.

Theorem 30.6

Let f be a function that is bounded over a rectangle R for which the points of discontinuity of f in R has zero area. Then, f is integrable over R. Note: the value of the area at the points of discontinuity does not matter.

Theorem 30.7 (Fubini's Theorem)

Let $f(x,y)$ be a bounded function on $R = [a,b] \times [c,d]$ and let S be the set of all points of discontinuity of f on R. Assume S has zero area, and suppose every line parallel to the z and y-axes intersects S in finitely many points. Then,

$$\iint\limits_R f(x,y) = \int_a^b \int_c^d f(x,y)\ dydx = \int_c^d \int_a^b f(x,y)\ dxdy$$

Theorem 30.8 (Properties of Double Integrals)

Suppose f and g are integrable functions over a rectangle R, and c is a constant. Then,

- $f+g$ is integrable, and $\iint\limits_R (f+g)\ dA = \iint\limits_R f\ dA + \iint\limits_R g\ dA$
- c is integrable, and $\iint\limits_R cf\ dA = c\iint\limits_R f\ dA$
- If $f \leq g$, $\iint\limits_R f\ dA \leq \iint\limits_R g\ dA$
- $|f|$ is integrable over R, and $\left|\iint\limits_R f\ dA\right| \leq \iint\limits_R |f|\ dA$

Definition 30.9. A region D in $\mathbb{R}^2$ is called **elementary** if it can be described in one of the following ways:

Type I (**Horizontally Simple**):

$$D = \{(x, y) \mid a \le x \le b, \delta_1(x) \le y \le \delta_2(x)\}$$

where δ_1 and δ_2 are continuous over $[a, b]$.

Type II (**Vertically Simple**):

$$D = \{(x, y) \mid \gamma_1(y) \le x \le \gamma_2(y), c \le y \le d\}$$

where δ_1, δ_2 are continuous over $[c, d]$.

Theorem 30.10 (Fubini's Theorem)
Let D be an elementary region in $\mathbb{R}^2$ an f a continuous function on D. Then,

- If D is a Type I region, we have

$$\iint\limits_D f \, dA = \int_a^b \int_{\delta_1(x)}^{\delta_2(x)} f(x, y) \, dydx$$

- If D is a Type II region, we have

$$\iint\limits_D f \, dA = \int_c^d \int_{\gamma_1(y)}^{\gamma_2(y)} f(x, y) \, dxdy$$

§31 Friday, November 11, 2022

Let's start with an example of computing a double integral.

Example 31.1
Suppose D is the region enclosed by $y = x^2$ and $y = \sqrt{x}$. Evaluate $\iint\limits_D (xy + y^2) \, dA$.

Solution. Fill this in ASAP. □

Remark 31.2. Let R be a region in $\mathbb{R}^2$. If $\delta(x, y)$ is the density of a thin metal surface placed at R, then the total mass of this surface is $\iint\limits_R \delta(x, y) \, dA$. When $\delta(x, y) = 1$, we get the area of R. The reason for this is because $\delta(x, y)$ is equivalent to $\frac{\text{mass}}{\text{area}}$. As dA is area, we have $\iint\limits_R \delta(x, y) \, dA$ is equal to the mass density multiplied by the area, which is equal to the mass of the metal plate.

§31.1 Changing the Order of Integration

Sometimes, we can us double integrals to evaluate iterated double integrals, i.e. integrals of form $\iint f(x,y)\ dxdy$ or $\iint f(x,y)\ dydx$.

Example 31.3
Evaluate $\int_0^1 \int_x^1 e^{y^x}\ dydx$.

Solution. Note that we cannot evaluate $\int_0^1 e^{y^2}\ dy$, as there does not exist a function whose derivative is e^{y^2}. However, we can relate this iterated integral with some double integral, analyze the region created by the functions in the double integral, and compute this double integral. Let's first identify an integrable region. From the bounds of integration, we have that $x \leq y \leq 1$ and that $0 \leq x \leq 1$, meaning $0 \leq y \leq 1$. As the region is horizontally simple, we have

$$\begin{aligned}\int_0^1 \int_x^1 e^{y^2}\ dydx &= \iint_R e^{y^2}\ dA \\ &= \int_0^1 \int_0^y e^{y^2}\ dydx \\ &= \int_0^1 xe^{y^2}|_{x=0}^{x=y}\ dy \\ &= \int_0^1 ye^{y^2}\ dy \\ &= \frac{1}{2}e^{y^2}|_0^1 \\ &= \frac{e-1}{2}\end{aligned}$$

□

Remark 31.4. The integrals of e^{y^2}, $\sin(y^2)$, and $\cos(y^2)$ cannot be evaluated. These will be important in 341.

§31.2 Triple Integrals

§31.2.1 Motivation

Similar to double integrals, we can define triple integrals over boxes. Suppose we have a solid cube with its faces parallel to each of the coordinate planes, called E. Now suppose E is the rectangular prism given by $a \leq x \leq b$, $c \leq y \leq d$, and $\alpha \leq z \leq \beta$. This is the simplest 3-dimensional region that we can integrate over using $\iiint_E f(x,y,z)dV$.

Remark 31.5. If $\delta(x,y,z)$ represents the density of a thin metal plate E, $\iiint f(x,y,z)dV$ represents the total mass of the region E, as $\delta(x,y,z) = \frac{\text{mass}}{\text{volume}}$ at (x,y,z).

Now, let's formally define the triple integral.

Definition 31.6. Let $f(x, y, z)$ be a function over the closed box $B = [a, b] \times [c, d] \times [p, q]$. A **partition** of B is a collection of boxes $B_{ijk} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$ for which

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b \text{ is a partition of } [a, b]$$
$$c = y_0 < y_1 < \cdots < y_{n-1} < y_n = b \text{ is a partition of } [c, d]$$
$$p = z_0 < z_1 < \cdots < z_{n-1} < z_n = b \text{ is a partition of } [p, q]$$

Let c_{ijk} be a point in the box B_{ijk}. Then, the quantity

$$\sum_{i,j,k=1} f(c_{ijk})\Delta V_{ijk} = \sum_{i=1} \sum_{j=1} \sum_{k=1} f(c_{ijk})\Delta V_{ijk}$$

where $\Delta V_{ijk} = \Delta x_i \Delta y_j \Delta z_k$ is the volume of the box B_{ijk}, is called a **Riemann sum** of f on B corresponding to this partition of B.

Example 31.7

Evaluate the volume of the solid that lies in first octant (i.e. $x, y, z > 0$) and inside the cylinders $x^2 + y^2 = 1$ and $y^2 + z^2 = 1$.

Solution. Fill this in ASAP. □

Example 31.8

Find the volume of the solid that lies above the surface $z = x^2 + y^2$ and above the plane $z = 1$.

Solution. Fill this in ASAP. □

§32 Monday, November 14, 2022

§32.0.1 Change of Variables

Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a continuously differentiable function and D is a region in $\mathbb{R}^2$. We would like to find a relation between $\iint_D f \, dA$ and $\iint_{T(D)} f \, dA$. Let $T(x, y) = (u, v)$. Essentially, we wnat to find a relation between $dxdy$ and $dudv$.

Example 32.1

Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, and D is a parallelogram formed by vectors u and v from the origin in $\mathbb{R}^2$. Prove that the area of $T(D)$ is equal to $|\det T| \cdot$ area of D.

Solution. Note that $T(D) = \begin{pmatrix} T(u) \\ T(v) \end{pmatrix}$ is a parallelogram formed by $T(u)$ and $T(v)$, e.g. $T(u) = Au$ and $T(v) = Av$. This means that the area of $T(D)$ is equal to $\Big|\det(Au\ Av)\Big| = \Big|\det(A(u\ v))\Big|$. By properties of the determinant, we have that the last expression is equal to $\det(A)\det(u\ v)\big| = \big|\det A\big| \cdot$ Area of D. □

Example 32.2

Consider the transformation $T(r, \theta) = (r\cos\theta, r\sin\theta)$. Find a relation between the area of R and $T(R)$ if R is a rectangle given by $0 \leq r \leq a$ and $\leq \theta \leq \beta$, where a is a positive constant, and $0 \leq \alpha < \beta \leq 2\pi$ are constants.

Solution. First, we must determine the graph of R in the xy-plane, and compare the area of this graph with the area of the rectangle in polar coordinate. The area of R is $a \cdot (\beta - \alpha)$. The area of the sector is $\pi a^2 \cdot \frac{(\beta-\alpha}{2\pi} = \frac{a^2(\beta-\alpha)}{2}$. Thus, we have that the area of $T(R)$ is equal to $\frac{a}{2} \cdot$ Area of R. □

Definition 32.3. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a continuously differentiable function given by $T(u, v) = (x(u, v), y(u, v))$. The **Jacobian** determinant of T is given by $\frac{\partial(x,y)}{\partial(u,v)} = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial u} \end{pmatrix}$

Theorem 32.4

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a continuously differentiable function given by $T(u, v) = (x(u, v), y(u, v))$.Suppose D and $T(D)$ are elementary regions in the uv and xy-planes, respectively. Then,

$$\iint\limits_{T(D)} f(x, y)\ dA = \iint\limits_{D} f(u, v)\left|\frac{\partial(x, y)}{\partial(u, v)}\right|$$

This is often summarized as

$$dxdy = \left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv$$

Remark 32.5. $dxdy = rdrd\theta$. This occurs when we do double integrals in polar coordinates.

Example 32.6

Evaluate $\iint\limits_R xy\, dA$, where R is

- the parallelogram whose vertices are $(0,0)$, $(1,1)$, $(1,2)$, and $(2,3)$
- the region in the first quadrant bounded by the lines $y = x$ $y = 2x$, and the hyperbolas $xy = 1$ and $xy = 2$.

Solution. For Part (a), it helps to draw the region of the parallelogram. Now, we can bound the rectangle between the lines $y = x + 1$ and $y = x$ (rather than trying to make the region horizontally or vertically similar). Take $u = x - y$ and $v = 2x - y$. The Jacobian $\frac{\partial(x,y)}{\partial(u,v)}$ is $\begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}$. Thus, the Jacobian determinant is $-1 + 2 = 1$. Now, note that $y = x$ becomes $u = 0$, $y = x + 1$ becomes $u = -1$, $y = 2x$ becomes $v = 0$, and $y = 2x - 1$ becomes $v = 1$. Now, our integral becomes

$$\iint\limits_R xy\, dA = \iint\limits_{T(R)} (v = u)(v = 2u)|1|dA = \int_0^1 \int_{-1}^0 (v-u)(v-2u)dudv$$

For Part (b), we can also draw the region bounded by the lines and hyperbolas. While we could attempt to break up the region into two vertically-simple regions, we could also attempt to bound the region like we did with the parallelogram. Consider $u = xy$ and $v = y$. From these bounds, we have $x = \frac{u}{v}$ and $y = v$ (unless $v = 0$). The Jacobian determinant of $\frac{\partial(x,y)}{\partial(u,v)}$ is equal to $\frac{1}{v}$. Now, we must find the limits of integration by looking at the boundaries of intersection; $xy = 1$ becomes $u = 1$; $xy = 2$ becomes $u = 2$; $y = x$ becomes $u = v^2$ (which gives us a parabola); $y = 2x$ becomes $v = \frac{2u}{v} \to u = \frac{1}{2}v^2$. Then, we have

$$\iint\limits_R xy\, dA = \iint\limits_S u\frac{1}{v}\, dA = \int \int_1^2 \int_{\sqrt{u}}^{\sqrt{2u}} \frac{u}{v} dvdu \tag{2}$$
$$= \int_1^2 u \ln v|_{v=\sqrt{u}}^{v=\sqrt{2u}} du \tag{3}$$
$$= \int_1^2 u\ln(\sqrt{2u}) - u\ln(\sqrt{u})du \tag{4}$$
$$= \int_1^2 u \ln\left(\frac{\sqrt{2u}}{\sqrt{u}}\right) du \tag{5}$$
$$= \ln(\sqrt{2}) \int_1^2 u\, du \tag{6}$$
$$= \frac{3}{2} \ln \sqrt{2} \tag{7}$$

Note that we use $dvdu$ because the region created by the equations with u and v is vertically simple. $\square$

For every point P in $\mathbb{R}^3$, we assign a triple (r, θ, z) called the **cylindrical coordinates** of P, where (r, θ) are the polar coordinates of the point (x, y). Similarly, we assign a triple (ρ, φ, θ), called the **spherical coordinates** of P, where ρ is the distance to the origin, φ is the angle that the vector OP makes with the positive direction of the z-axis, and θ is the same angle as in the polar coordinates of (x, y). Note that we have $\rho \geq 0$, $0 \leq \varphi \leq \pi$, and $0 \leq \theta < 2\pi$. There are several useful formulas that we obtain from cylindrical and spherical coordinates that are in Dr. Ebrahimian's notes.

§33 Wednesday, November 16, 2022

Let's start with an example:

Example 33.1

Evaluate the volume of a sphere (or a ball) of radius a.

Solution. Let E be the ball $B_a(0)$ in $\mathbb{R}^3$. We want to find the volume of this ball. We have $\iiint\limits_E 1\ dV$ represents the volume of the ball. It's easiest to use spherical coordinates here. When we do this, we have a factor of $\rho^2 \sin\varphi$ (this is the Jacobian determinant of the spherical coordinates). Our integral now becomes

$$\iiint\limits_E \rho^2 \sin\phi d\varphi d\theta$$

We will fix our theta for some arbitrary θ and arbitrary φ, we obtain a ray from the origin. Thus, $0 \leq \varphi$. Fixing θ, we obtain a half-plane. The limits of φ are $0 \leq \varphi \leq \pi$. Thus, our integral becomes

$$\int_0^{2\pi} \int_0^{\pi} \int_0^{a} e^2 \sin\varphi\ d\rho d\varphi d\theta$$

$\square$

Example 33.2

Find the volume of the solid that lies inside both the cylinder $r = 1$ and the sphere $\rho = 2$.

Solution. We will proceed with cylindrical coordinates. Fill this in ASAP. $\square$

§33.1 Applications of Integration

Definition 33.3. The **average value** of a function f over a region $D \subseteq \mathbb{R}^2$ is given by

$$f_{\text{avg}} = \frac{\iint Df(x,y)\ dA}{\text{Area of D}}$$

The average value a function f over a solid $E \subseteq \mathbb{R}^3$ is given by

$$f_{\text{avg}} = \frac{\iiint\limits_E f(x,y,z)\ dV}{\text{Volume of E}}$$

Example 33.4

Find the average value of the function $f(x,y,z) = z$ over the solid E, where E is the solid that lies inside the surface given by $x^2 + y^2 + z^2 = 2z$ and is above the surface given by $z = \sqrt{x^2 + y^2}$.

Solution. We must first find the volume of E. We can first sketch the surfaces. To do so, we have to identify the region $x^2 + y^2 + z^2 = 2z$. Completing the square, we obtain $x^2 + y^2 + (z-1)^2 = 1$. This surface is thus the sphere centered at $(0,0,1)$ with radius 1. The surface $z = \sqrt{x^2. + y^2}$ can be rewritten as $x^2 + y^2 - z^2 = 0$, which is a cone with vertex at $(0,0,0)$. We can now set up the triple integral. We should proceed with spherical coordinates. Fixing θ, φ, we get a ray from the origin itself. Farthest from the origin is is on the sphere. We must determine ρ on the sphere; we can do this using $x^2 + y^2 + z^2 = 2z$. We have $\rho^2 = 2\rho\cos\varphi$. This gives us $\rho = 2\cos\varphi$. To find the smallest and largest φ, we can □

Definition 33.5. A **point mass** is a mass concentrated at a single point. The **moment** of a point mass m located at point x on the number line with respect to the origin is mx.

Definition 33.6. The **center of mass** of a finite number of point masses on a number line is a point such that if the total masses were concentrated there, the total moment of point masses would be the same as the moment of the mass located at the center of mass.

§34 Friday, November 18, 2022

Let's start with an example of center of mass.

Example 34.1

Find the center of mass of a hemisphere with radius r.

Solution. Assume that the density of the hemisphere is 1, e.g. $\delta(x,y,z) = 1$. We have $x = y = 0$; if you look at the numerator, we have $\iiint\limits_E x\ dV$ is 0, by symmetry (as the function of the hemisphere is odd). To find $\overline{z}$, we can calculate

$$\begin{aligned}\overline{z} &= \frac{\iint Ez\ dV}{\iint 1\ dV} \\ &= \frac{\iint\limits_E z\ dV}{\iint\limits_E 1\ dV} \\ &= \frac{\iiint \rho\cos\varphi\rho^2\sin\varphi d\rho d\varphi d\theta}{\frac{1}{2}\cdot\frac{4}{3}\pi a^3} \\ &= \frac{\int_0^{2\pi}\int_)^{\frac{\pi}{2}}\int_0^a \rho cos\varphi\rho^2\sin\varphi d\rho d\varphi d\theta}{\frac{1}{2}\cdot\frac{4}{3}\pi a^3}\end{aligned}$$

Note that we used spherical coordinates in the integrals in the numerator, as well as the fact that the derivative of the volume is area in the integrals in the denominator. The calculation of this integral is left as an exercise for the reader. ☐

Example 34.2

An object is located in the first octant and below the plane $x + y + z = 3$. Suppose the mass density of this object is given by $\delta(x, y, z) = \sqrt{x^2 + y^2 + z^2}$. Find the center of mass of this object.

Solution. By symmetry, we have that the coordinate (x, y, z) of the center of mass will satisfy $x = y = z$. Thus, we only need to find one coordinate. We have

$$\overline{x} = \frac{\iiint x\sqrt{x^2+y^2+z^2}\ dV}{\iiint\limits_E \sqrt{x^2+y^2+z^2}\ dV} \tag{8}$$

$$\tag{9}$$

Using Cartesian coordinates is acceptable here. This is because the top part of our region is $z = 3 - x - y$ and the bottom part of this region is $z = 0$. The projection of this region on the xy-plans is a triangle made with $y = 3 - x$, $y = 0$, and $x = 0$. Our integral expression becomes

$$\frac{\int_0^3\int_0^{3-x}\int_0^{3-x}\int x\sqrt{x^2+y^2+z^2}\ dzdydx}{\int_0^3\int_0^{3-x}\int_0^{3-x-y}\sqrt{x^2+y^2+z^2}\ dzdydx}$$

Evaluating this integral is left as an exercise. ☐

Now, we will discuss **line integrals**. There are two types of line integrals: scalar line integrals and vector line integrals. We will start by discussing scalar line integrals.

§34.1 Scalar Line Integrals

Scalar line integrals evaluate total mass of a wire.

Suppose we have a wire c given by the parametrization $r(t)$ such that $a \le t \le b$. To find the mass of the wire, we need the density of the wire. We will proceed in 2D, although all of the following can be extended to 3D. Suppose $\delta(x, y)$ is the mass density. Let the wire have negligible thickness; then we have $\delta(x, y) = \frac{\text{mass}}{\text{length}}$. To find the total mass of the wire, we can divide the wire into several region, choose a point in each place, and evaluate the density at each point. Our **total mass** is approximated by $\sum_{i=1}^{n} \delta(r(t_i)) = \Delta S_i$, where ΔS_i is the length of the region. Recall the formula for arc length: $\int_a^b \sqrt{1 + f'(x)^2}\, dx$, where $y = f(x)$, $a \le x \le b$. The formula for arc length for parametric curves is $\int_a^b ||r'(t)||dt$ (the justification for this can be done using Riemann sums). If we take the derivative of arc length, e.g. $\frac{ds}{dt}$, we obtain $||r'(t)||$, by the Fundamental Theorem of Calculus. This tells us that the total mass of the wire, if we let $n \to \infty$, is $\int\limits_C \delta(x, y)\, ds = \int_a^b \delta(x(t), y(t))||r'(t)||\, dt$, where we replaced ds with $||r'(t)||\, dt$ in the second integral.

Remark 34.3. When we want to find the total mass of the wire, we must have a parametrization of the wire.

Definition 34.4. Consider a curve parametrized by $x : [a, b] \to \mathbb{R}^n$. The **scalar line integral** of a real-valued function f over this curve is given by

$$\int_a^b f(x(t))||x'(t)||\, dt$$

We denote this integral $\int\limits_x f\, ds$

Example 34.5

Find the total mass of a wire located at the unit circle $x^2 + y^2 = 1$ whose density is given by $\delta(x, y) = x^2$.

Solution. The total mass is given by $\int\limits_C x^2\, ds$. To parametrize the circle, we must convert to polar coordinates: let $r(t) = (cos(t), \sin(t))$. We have $0 \le t \le 2\pi$. We can now plug this into the formula for mass:

$$\begin{aligned}\int\limits_C x^2\, ds &= \int_0^{2\pi} \cos^2(t)||(-\sin t, \cos t)||\ s \\ &= \int_0^{2\pi} \cos^2 t\ dt = \int_0^{2\pi} \frac{1 + \cos(2t)}{2}\end{aligned}$$

Note that we determined the bounds of integration by evaluating density on the points of our curve (e.g. the circle should only be defined when $0 \le \theta \le 2\pi$). We also used the $\square$

§34.2 Vector Line Integrals

Recall that the work done by a constant force $\mathbf{F}$ with displacement vector $\mathbf{D}$ is given by $\mathbf{F} \cdot \mathbf{D} = ||\mathbf{F}|| \cos\theta ||\mathbf{D}||$. However, this only works when the force is constant.

Definition 34.6. A **vector field** is a function $\mathbf{F} : D \to \mathbb{R}^n$, where D is a subset of $\mathbb{R}^n$. Essentially, vector fields assign vectors in $\mathbb{R}^n$ to vectors in a subset of $\mathbb{R}^n$

An intuitive example of this is explained with fluid flow in a river. Every given point in the river has a velocity. A similar example is gravitational force, where every force has a different gravitational force. Suppose we have a path where the force at each point is given by $\mathbf{F}(r(t))$, where $r(t)$ is the displacement at that point. Note that we define displacement $r'(t)$ as displacement/sec.

Definition 34.7. The **vector line integral** of a vector field $\mathbf{F} : D \to \mathbb{R}^n$ over r to be

$$\int_C \mathbf{F} \cdot ds = \int_a^b \mathbf{F}(r(t)) \cdot r'(t)\ dt$$

If we change the orientation of c, the vector line integral will be negated. The process of computing the integral will be the same, but the orientation of the line must be taken into account. In other words, if we reverse the points that we want to find the work done, we can simply negate our integral.

Example 34.8

Find the work done by the force $\mathbf{F}(x, y, z) = (x, y, 2z^3)$ along the curve given by $x(t) = (t, t^2, t)$ from $(0, 0, 0)$ to $(1, 1, 1)$.

Solution. Note that work done is $\int_{\mathbf{x}} \mathbf{F} \cdot d\vec{s}$. First, we need to parametrize $\mathbf{F}(x, y, z)$. Our limits of integration become 0 to 1, as $x(0) = (0, 0, 0)$ and $x(1) = (1, 1, 1)$. Our integral becomes

$$\begin{aligned} \int_0^1 \mathbf{F}(t, t^2, t) \cdot (1, 2t, 1)\ dt \\ &= \int_0^1 (t, t^2, 2t^3) \cdot (1, 2t, 1)\ dt \\ &= \int_0^1 t + 2t^7 + 2t^3\ dt \end{aligned}$$

□

Suppose our vector field consists of 3 components, s.g. $\vec{F}(x, y, z) = M(x, y, z)\vec{i} + N(x, y, z)\vec{j} + P(x, y, z)\vec{k}$. Let $x = (x', y', z')$. Then, $F \cdot (x', y', z'$ is equal to

$$Mx' + Ny' + Pz'$$

This implies

$$\begin{aligned}\int \vec{F} \cdot \vec{x}dt &= \int (Mx' + Ny' + Pz')\ dt \\ &= \int Mdx + Ndy + Pdz\end{aligned}$$

Note that we removed the dt term above, as we can now express everything with dx, dy, and dz. Essentially, instead of $\int\limits_C \vec{F} \cdot d\vec{s}$, we write $\int_C Mdx + Ndy + Pdz$. The process of computing the integral is still rather simple: parametrize x, y, z, and evaluate the integral.

Remark 34.9. Up until now, we've assumed all of these functions are differentiable; this may not always be the case, and you may have to work around this (e.g. split the region into integrable regions and sum the integrals of these regions).

§35 Monday, November 21, 2022

Today, we'll begin with certain definitions regarding differentiability and parametrization.

Definition 35.1. A function $x : [a, b] \to \mathbb{R}^n$ is said to be **piecewise continuously differentiable** if x is continuous, and the interval $[a, b]$ can be partitioned into finitely many intervals $a = t_0 < t_1 < \cdots < t_n = b$ for which x is continuously differentiable on each interval (t_i, t_{i+1}).

Definition 35.2. Let $x : [a, b]$ be a piecewise continuously differentiable path. We say another piecewise continuously differentiable path $y : [c, d] \to \mathbb{R}^n$ is a **reparametrization** of x if there is a bijective continuously differentiable function $u : [c, d] \to [a, b]$ whose inverse is also continuously differentiable such that $y = x \circ u$. If $y(c) = x(a)$ and $y(d) = x(b)$, then we say y is **orientation-preserving**. If $y(c) = x(b)$ and $y(d) = x(a)$, we say y is **orientation-reversing**.

Example 35.3

$y : [0, 1] \to \mathbb{R}^3$ given by $y(t) = (t, 2t, 3t)$ is a reparametrization of $x : [2, 4] \to \mathbb{R}^3$ given by $x(t) = \left(0.5t - 1, t - 2, \frac{3t}{2} - 3\right)$.

Theorem 35.4

Suppose $f : U \to \mathbb{R}$ is a continuous scalar-valued function over an open subset U of $\mathbb{R}^n$. Suppose x is a path whose image is inside U. If y is a reparametrization of x, then

$$\int_y f\ ds = \int_x f\ ds$$

Theorem 35.5

Suppose $\mathbf{F} : U \to \mathbb{R}$ is a continuous vector field over an open subset U of $\mathbb{R}^n$. Suppose x is a piecewise continuously differentiable path whose image is inside U. If y is an orientation-preserving reparametrization of x, $\int_y \mathbf{F} \cdot ds = \int_x \mathbf{F} \cdot ds$.

If y is orientation-reversing,

$$\int_y \mathbf{F} \cdot ds = -\int_x \mathbf{F} \cdot ds$$

.

Theorem 35.6

Let C be a curve given by a parametrization $x : [a, b] \to \mathbb{R}^n$. Suppose $c \in (a, b)$ and the restriction of x to intervals $[a, c]$ and $[c, b]$ divides C into two curves C_1 and C_2. Then, for every vector field $\mathbf{F}$ and every scalar function f, we have

$$\int_C \mathbf{F} \cdot ds = \int_{C_1} \mathbf{F} \cdot ds + \int_{C_2} \mathbf{F} \cdot ds$$

and

$$\int_C f\ ds = \int_{C_1} f\ ds + \int_{C_2} f\ ds \tag{10}$$

§35.1 Green's Theorem

Definition 35.7. A piecewise continuously differentiable path $x : [a, b] \to \mathbb{R}^n$ is said to be **closed** if $x(a) = x(b)$. It is said to be **simple** if x is one-to-one except possible $x(a)$ may be equal to $x(b)$. If x is one-to-one except possible at finitely many points of $[a, b]$, we say its image is a **curve** C, in which case x is said to be a parametrization of C. We say C is closed or simple if it has a parametrization that has the corresponding property.

Theorem 35.8 (Green's Theorem)
Let D be a closed and bounded region in $\mathbb{R}^2$, whose boundary ∂D consists of finitely many simple, closed, piecewise continuously differentiable cuves. Suppose ∂D is oriented in such a way that D lies on the left as one traverses ∂D. Let $F : M\vec{i} + N\vec{j}$ be a continuously differentiable vector field on D. Then,

$$\int_{\partial D} M\ dx + N\ dy = \iint_{D} \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\ dA$$

Remark 35.9. We often write $\oint_C$ instead of $\int_C$ to indicate C is a union of finitely many closed curves.

Example 35.10
Evaluate $\oint_C (x^2 - y^2)\ dx + \int_C (x^2 + y^2)\ dy$, where C is the boundary of the square whose vertices are $(0, 0)$, $(1, 0)$, $(1, 1)$ and $(0, 1)$ oriented clockwise.

§36 Monday, November 28, 2022

Recall how we defined and computed scalar and vector line integrals. We will use these techniques when examining **conservative vector fields** to evaluate line integrals.

§36.1 Conservative Vector Fields

Definition 36.1. A vector field $\mathbf{F}$ is said to have **path-independent** line integrals if

$$\int_{C_1} \mathbf{F} \cdot ds = \int_{C_2} \mathbf{F} \cdot ds$$

for any two simple piecewise continuously differentiable curves C_1, C_2, lying in the domain of $\mathbf{F}$ that have the same initial and terminal point.s

Essentially, this states that the path of the curve does not matter when computing the line integral of the path.

Example 36.2

Check if each of the following vector fields has path-independent line integrals:

(a) $xy\mathbf{i} + y\mathbf{j}$

(b) $x\mathbf{i} + y\mathbf{j}$

Solution. To check if a vector field has path-independent line integrals, we can first take two points that the path is dependent on, say a and b, and suppose the vector field has parametrization $\vec{x}(t) = (x(t), y(t))$ on each path. For part (a), we have that by the definition of the line integral,

$$\begin{aligned}\int_a^b \mathbf{F} \cdot ds &= \int (x(t)y(t)\mathbf{i} + y(t)\mathbf{j}) \cdot (x'(t), y'(t))\ dt \\ &= \int_a^b \left[x(t)y(t)x'(t) + y(t)y'(t)\right]\ dt\end{aligned}$$

Now, we can examine if the integrand is path-independent. $y(t)y'(t)$ is the derivative of $\frac{y(t)^2}{2}$. Our integral now becomes

$$\int_a^b x(t)y(t)y'(t) + \left[\left(\frac{y(t)^2}{2}\right)'\right] dt$$

As $\int_a^b \frac{y(t)^2}{2}' dt = \frac{y(b)^2}{2} - \frac{y(a)^2}{2}$, we have that it does not depend on the path. We suspect $x(t)y(t)y'(t)$ is not independent of path. We can analyze two different paths from two different points, say $(0,0)$ and $(1,1)$. Let $x = y = t$, where $0 \leq t \leq 1$. We have $\int\limits_{C_1} (xy\mathbf{i} + y\mathbf{j}) \cdot ds$ is equal to

$$\begin{aligned}\int_0^1 (t^2\mathbf{i} + t\mathbf{j}) \cdot (1,1)\ ds &= \int_0^1 (t^2 + t)\ dt \\ &= \frac{1}{3} + \frac{1}{2} = \frac{5}{6}\end{aligned}$$

Now, let $C_2 = (t, t^2)$, where $0 \leq t \leq 1$. Computing $\int\limits_{C_1} (xy\mathbf{i} + y\mathbf{j} \cdot ds$, we obtain $\frac{3}{4}$. As the integral over C_1 and the integral over C_2 are different, we have that these vector fields are **not** path independent.

We can do part (b) similarly, starting from a point A and going to a point B. Let C be given by $(x(t), y(t))$, where $a \leq t \leq b$. We can now create our line integrals:

$$\begin{aligned}\int\limits_C \mathbf{F} \cdot ds &= \int_a^b (x(t)\mathbf{i} + y(t)\mathbf{j}) \cdot (x'(t), y'(t))\ dt \\ &= \int_a^b \left[x(t)x'(t) + y(t)y'(t)\right]\ dt\end{aligned}$$

This is similar to the second portion of part (a); our integral becomes $\int_a^b \frac{1}{2}\left[x(t)^2 + y(t)^2\right] dt = \frac{1}{2}\left[x(b)^2 + y(b)^2 - x(a)^2 + y(b)^2\right]$. Thus, the integral is path-independent. □

Theorem 36.3
A vector field $\mathbf{F}$ has path-independent line integrals if and only if $\oint\limits_C \mathbf{F} \cdot ds = 0$ for every simple, piecewise continuously differentiable closed curve C in the domain of F.

Proof. Let's start with the forward direction: assume the line integral is independent of path. We can look at two regions: one from A to a point C_1, where C_1 is counterclockwise to A, and one from C_2, which is clockwise to A. We have $\int\limits_C \mathbf{F} \cdot ds = \int\limits_{C_1} \mathbf{F} \cdot ds + \int\limits_{C_2} \mathbf{F} \cdot ds$. This is equal to

$$\int\limits_{C_1} \mathbf{F} \cdot ds - \int\limits_{-C_2} \mathbf{F} \cdot ds$$

The above integral is 0, because the line integrals are independent of path. Now, we can proceed with the backwards direction. This is quite similar. □

Definition 36.4. A continuous vector field $\mathbf{F}$ is said to be **conservative** if $\mathbf{F} = \nabla f$ for some continuously differentiable real-valued function f. We call f a **potential function** of $\mathbf{F}$.

Theorem 36.5 (Fundamental Theorem of Line Integrals)
Suppose $\mathbf{F}$ is a continuous vector field over an open and connected subset u of $\mathbb{R}^n$. Then, $\mathbf{F}$ is conservative on U, if and only if $\mathbf{F}$ has path-independent line integrals over curves in u. Furthermore, if C is a piecewise continuously differentiable curve in U from point A to point B and $\mathbf{F} = \nabla f$, we have

$$\int\limits_C \mathbf{F} \cdot ds = f(B) - f(A)$$

Proof. Fill this in ASAP. □

Example 36.6
Determine whether $(x^2 + 2, y - 1)$ is conservative or not.

Solution. We will proceed with Definition 36.4 and attempt to find a potential function $f(x, y)$ for $(x^2+2, y-1)$. We have $\frac{\partial f}{\partial x} = \frac{x^3}{3} + 2x$. Let $f(x, y) = \frac{x^3}{3} + 2x + g(y)$

and $\frac{\partial f}{\partial y} = g'(y)$. Now, we must find an appropriate $g(y)$. We have $g(y) = \frac{y^2}{2} - y$ works. Thus, $f(x,y) = \frac{x^3}{3} + 2x + \frac{y^2}{2} - y$ is a potential function. The line integral is then, by the Fundamental Theorem of Line Integrals, $f(2,1) - f(1,0) = \frac{8}{3} + 4 - \frac{1}{2} - 1 - (\frac{1}{3} + 2) = \frac{4}{3}$. $\square$

Definition 36.7. A region U in $\mathbb{R}^2$ or $\mathbb{R}^3$ is called **simply connected** if it is connected and every simply closed curve in U can be continuously shrunk to a point while remaining in U. In other words, if $x : [a,b] \to U$ is a parametrization of a simple closed curve, then there is a continuous function $\varphi : [a,b] \times [0,1] \to U$ for which $\varphi(0,1) = x(t)$ for all t, and $\varphi(t,1)$ is a constant.

Definition 36.8. Let $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ be two vectors in $\mathbb{R}^3$. The **cross product** $\mathbf{u} \times \mathbf{v}$ is defined as

$$\mathbf{u} \times \mathbf{v} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_2 \\ v_1 & v_2 & v_3 \end{pmatrix}$$

Theorem 36.9 (Properties of Cross Products)

For every three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ and every scalar $c \in \mathbb{R}$, we have the following:

(a) $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$

(b) $\mathbf{v} \times \mathbf{u} = -(\mathbf{u} \times \mathbf{v})$

(c) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$

(d) $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$

Proofs of these properties are relatively simple using the definitions of the cross product and properties of the determinant.

Definition 36.10. Let $\mathbf{F}(x,y,z) = M(x,y,z)\mathbf{i} + N(x,y,z)\mathbf{j} + P(x,y,z)\mathbf{k}$ be a vector field in $\mathbb{R}^3$. The **curl** of $\mathbf{F}$, denoted by curl $\mathbf{F}$, is defined as

$$\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ M & N & P \end{pmatrix}$$

In the two-dimensional case, the curl is similarly defined with $P = 0$.

The concept of the curl will become more relevant when we get to Stokes' Theorem. However, the curl does give us the criterion we were looking for to determine whether a vector field is conservative or not.

Theorem 36.11

Suppose U is a simply connected region in $\mathbb{R}^2$ or $\mathbb{R}^3$. Let $\mathbf{F}$ be a continuously differentiable vector field on U. Then, $\mathbf{F} = \nabla f$ for some real-valued function f if and only if curl $\mathbf{F} = 0$ on U.

Example 36.12

Without evaluating a potential function, show that the vector field $\mathbf{F}(x, y, z) = (3x^2 + y\sin(xy))\mathbf{i} + (2y + x\sin(xy))\mathbf{j} + (2z+1)\mathbf{k}$ is conservative. Evaluate $\int_C \mathbf{F} \cdot ds$, where C is the curve given by $\mathbf{x}(t) = (t^2, e^t, 2t)$ from $(0, 1, 0)$ to $(1, e, 2)$.

§37 Wednesday, November 30, 2022

§37.1 Parametrized Surfaces

Definition 37.1. Let D be a subset of $\mathbb{R}^2$ that consists of an open connected set along with some or all of its boundary. A **parametrized surface** in $\mathbb{R}^3$ is a continuous function $X : D \to \mathbb{R}^3$ that is one-to-one on D except possible along ∂D. We say $X(D)$ is a surface **parametrized** by X.

Example 37.2

Find a parametrization for the unit sphere $x^2 + y^2 + z^2 = 1$ and the cylinder $x^2 + y^2 = 1$.

Solution. Fill this in ASAP. □

Definition 37.3. Let $X : D \to \mathbb{R}^3$ given by $X(s, t)$ be a parametrization of the surface $S = X(D)$. An ***s*-coordinate curve** at $t = t_0$ is the curve given by $s \mapsto X(s, t_0)$. t-coordinate curves are defined similarly.

Example 37.4

Find a parametrization of a torus. Use that to find its coordinate curves.

We know that partial derivatives X_t and X_s give us vectors that are tangent to the coordinate curve. Therefore, to find the vector normal to both coordinate vectors we need to evaluate $X_s \times X_t$.

§38 Friday, December 1, 2022

Surface Integrals

§39 Monday, December 5, 2022

Recall that to evaluate a surface integral of a scalar function, e.g. $\iint\limits_{\Sigma} f(x,y,z)\ ds$, we mut find a parametrization $X(s,t)$ for the surface, and compute $\iint\limits_{D} f(X(s,t))||X_s \times X_t||$ using an iterated integral.

Example 39.1

Suppose S is part of the paraboloid $z = x^2 + y^2 - 4$ that lies below the xy-plane (e.g. $z \le 0$)). Evaluate the surface integral of $z + 4$ over S, e.g. $\iint\limits_{S} (z+4)\ dS$.

Solution. After drawing a diagram, we must first parametrize this surface. There are two ways in which we can do this. First, let us proceed with cylindrical coordinates. Let $X(r,\theta) = (r\cos\theta, r\sin\theta, r^2 - 4)$. We must also find the limits of the parameters. We have $0 \le r \le 2$ and $0 \le \theta \le 2\pi$. Now, we have to find the partials of X with respect to r and θ. We have $X_r = (\cos\theta, \sin\theta, 2r)$ and $X_\theta = (-r\sin\theta, r\cos\theta, 0)$. Taking the cross product of these two, we obtain $X_r \times X_\theta = (-2r^2\cos\theta, -2r\sin\theta, r)$. Now, we have to take the magnitude $||X_r \times X_\theta||$, which is $\sqrt{4r^4 + r^2}$. Finally, we can set up our integral. We have

$$\begin{aligned}
\iint\limits_{S} r^2\ dS &= \int_0^{2\pi}\int_0^2 \sqrt{4r^4+r^2}\ drd\theta \\
&= \int_0^{2\pi}\int_1^1 7\frac{u-1}{4}\sqrt{u}\ \frac{du}{8}d\theta \\
&= \frac{1}{32}\int_0^{2\pi}\int_1^{17} u^{\frac{3}{2}} - u^{\frac{1}{2}}\ dud\theta
\end{aligned}$$

Note that we made the substitution $u = 4r^2 + 1$, $du = 8rdr$, and $r^2 = \frac{u-1}{4}$.

Another parametrization for this curve is $X(x,y) = (x, y, x^2 + y^2 - 4)$. The limits for our parameter become $-2 \le x \le 2$ and $-\sqrt{4-x^2} \le y \le \sqrt{4-x^2}$, or vice versa (with x replacing y). Now, we have $X_r \times X_y = (-2x, -2y, 1)$. We can now set up our integral:

$$\iint\limits_{S} (z+4)\ dA = \iint\limits_{x^2+y^2\le 4} (x^2+y^2-4)\sqrt{4x^2+4y^2+1}\ dA$$

Now, we can switch to polar coordinates. Here, we have dA is in the xy-plane, meaning we will have to use $rdrd\theta$. Our integral becomes

$$\int_0^{2\pi}\int_0^2 r^2\sqrt{4r^2+1}rdrd\theta$$

The above integral is the same as the first one, and can be computed in a similar way. ☐

Remark 39.2. When converting to cylindrical coordinates above, we do not use $rdrd\theta$ and have to use $drd\theta$, as by the definition of the surface integral, there is no difference between dS and dA (both are in the same plane).

Example 39.3

Find the surface integral of the vector field of $x\mathbf{i} + y\mathbf{j} - z\mathbf{k}$ along the unit sphere. Use a parametrization whose normal vector points outwards.

Solution. Let $\vec{n}$ be the normal vector that we want to evaluate. We want to compute $\iint_{x^2+y^2+z^2=} (x\mathbf{i} + y\mathbf{j} - z\mathbf{k}) \cdot \vec{n}\ dS$. For our parametrization, let $X(\varphi, \theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$. Now, we have $X_r \times X_\theta = \rho\sin\varphi \cdot X$ $= \sin\varphi(\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$. Recall from last class we derived $\vec{n}dS = X_\varphi \times X_\theta = (X_s \times X_t)\ dA$. Finally, we have to check the orientation of the vector $\vec{n}$. However, as the problem explicitly stated that this orientation was outward, we only have to test that it is outward (if it was inward, we would simply use $-\vec{n}$). At $(0, 1, 0,$ we have $\varphi = \frac{\pi}{2}$ and $\theta = \frac{\pi}{2}$. As $X_\varphi \times X_\theta$ at $\varphi = \frac{\pi}{2}$ and $\theta = \frac{\pi}{2}$ points outwards (we can determine this by drawing the vector on the diagram), we have that its orientation is outward Thus, we have

$$
\begin{aligned}
&\iint_{x^2+y^2+z^2=1} (x\mathbf{i} + y\mathbf{j} - z\mathbf{k}) \cdot \vec{n}\ dS \\
&= \int_0^{2\pi}\int_0^{\pi} (\sin\varphi\cos\theta, \sin\varphi\sin\theta, -\cos\varphi) \cdot \sin\varphi(\sin\varphi\cos\theta, \sin\varphi, \sin\theta, \cos\varphi) d\varphi d\theta \\
&= \int_0^{2\pi}\int_0^{\pi} (\sin^3\varphi\cos^2\theta + \sin^3\varphi\sin^3\theta - \sin\varphi\cos^2\varphi)\ d\varphi d\theta \\
&= \int_0^{2\pi}\int_0^{2\pi} (\sin^3\varphi - \sin\varphi\cos^2\varphi) d\varphi d\theta
\end{aligned}
$$

The computation of this integral is left as an exercise; it should be simpler with the substitution $u = \cos\varphi$. ☐

We have a similar notion for reparametrization of curves to surfaces.

Definition 39.4. Let D_1, D_2 be two regions in $\mathbb{R}^2$ and $X_1 : D_1 \to \mathbb{R}^3$ and $X_2 : D_2 \to \mathbb{R}^3$ be two parametrized surfaces. We say X_2 is a **reparametrization** of X_1 if there is a bijection $H : D_2 \to D_1$ such that $X_2 = X_1 \circ H$ and if X_1 and X_2 are piecewise smooth and H and H^{-1} are continuously differentiable we say X_2 is a **smooth reparametrization** of X_1.

Example 39.5

Let $\sum$ be part of the cylinder $r = 1$ that lies between $z = 0$ and $z = 1$ along with the disks $x^2 + y^2 \leq 1$ in the planes $z = 0$ and $z = 1$ oriented outward from the cylinder. Evaluate $\iint\limits_{\sigma} (x^2\mathbf{i} + z\mathbf{j})\ d\mathbf{S}$.

Solution. Let $\sum_1$ be $x^2 + y^2 \leq 1$ and $z = 0$, $\sum_2$ be $x^2 + y^2 \leq 1$ and $z = 1$, and $\sum_3$ be $x^2 + y^2 = 1$ and $0 \leq x \leq 1$ We have $\sum = \sum_1 \cup \sum_2 \cup \sum_3$. We can find the surface integrals of each of these individual regions and add our results.

Let $\sum_1 : X(x, y, 0)$ with $x^2 + y^2 \leq 1$. Now, we have $X_x \times X_y = (0, 0, 1)$. Now, we have to check that this orientation matches the orientation that we want. This orientation is upward, and does not match the one we want; we must negate it to obtain $(0, 0, -1)$. Now, our integral becomes

$$\iint\limits_{\Sigma_1} (x^2\mathbf{i} + z\mathbf{j}) \cdot dS = \iint\limits_{x^2+y^2\leq 1} (x^2\mathbf{i} + 0\mathbf{j}) \cdot (0, 0, -1)\ dA$$
$$= 0$$

Note that no calculation was needed above, as the dot product is 0. Now, let's set up our second integral. First, we must parametrize it: let $\sum_2 : X(x, y) = (x, y, 1)$ with $x^2 + y^2 \leq 1$. Now, we can set up our integral:

$$\iint\limits_{x^2+y^2\leq 1} (x^2\mathbf{i} + 1\mathbf{j}) \cdot (0, 0, 1)\ dA = 0$$

Once again, no calculation was needed above, as the dot product is 0. Finally, we can set up our integral for $\sum_2$. Let our parametrization be $X(\theta, z) = (\cos\theta, \sin\theta, z)$ with $0 \leq \theta \leq 2\pi$ and $0 \leq z \leq 1$. We have $X_\theta \times X_z = (\cos\theta, \sin\theta, 0)$. Now, we must check the orientation of the vector; let's use $(0, 1, 0)$ and $\theta = \frac{\pi}{2}$. At this point, the vector is outward, which is what we want. Thus, we can compute

$$\int_0^{2\pi} \int_0 61(\cos^2\theta\mathbf{i} + z\mathbf{j}) \cdot (\cos\theta, \sin\theta, 0)\ dzd\theta$$
$$= 0$$

Computing the above integral is left as an exercise. $\square$

§40 Wednesday, December 7, 2022

Today is the last day we will cover new material; we'll discus **Stokes' Theorem** and **Gauss' Theorem**.

§40.1 Stokes' Theorem

Stokes' Theorem equates the surface integral of the curl of a C^1 vector field over a piecewise smooth, orientable surface with the line integral of the vector field along the boundary curve(s) of the surface. Since both vector line and surface integrals are examples of **oriented** integrals (i.e. they depend on the particular orientations chosen), we must comment on the way in which orientations need to b taken.

Definition 40.1. Let S be a bounded, piecewise smooth, oriented surface in $\mathbb{R}^3$ and C' be any simple, closed curve lying in S. Consider the unit normal vector $\mathbf{n}$ that indicates the orientation of S at any point inside C'. Use the right hand rule to orient C', e.g. if the thumb of your right hand points along $\mathbf{n}$, the fingers curl in the direction of the orientation of C'. We say that C' with the orientation just described is **orieented consistently** with S or that the orientation is the one **induced** from that of S. Now, suppose the boundary ∂S of S consists of finitely many piecewise C^1 simple, closed curves. Then, we say ∂S is **oriented consistently** (or that ∂S has its **orientation induced** from that of S) if each of its simple, closed pieces is oriented consistently with S.

Theorem 40.2 (Stokes' Theorem)
Let S be a bounded, piecewise smooth, oriented surface in $\mathbb{R}^3$.

Example 40.3
Let S be a part of the plane $x+2y+3z=6$ that lies in the first octant. Let C be the boundary of S oriented counterclockwise when viewed from above. Evaluate $\int\limits_C x^2\,dx + y^2z\,dz + z^2\,dz$.

Solution. We can evaluate this integral using Stokes' Theorem, but also with regular line integrals. We will do it with Stokes Theorem. After drawing a diagram of our region, we have that the orientation must be upwards; this was also given in the problem statement. We can calculate the curl of this region, which is $-y^2\mathbf{j}$. Now, we need a parametrization for this surface. □

Example 40.4
Let S be the surface formed by the cylinder $x^2+y^2=1$ with $0 \le z \le 2$ together with the disk $x^2+y^2 \le 1$ in the xy-plane. Consider the orientation of normal vectors outwards from the cylinder. Evaluate $\iint\limits_S \operatorname{curl}(-y\mathbf{i}+x\mathbf{j}+x^2\mathbf{k})\,dS$ using Stokes' Theorem.

Solution. □

§40.2 Gauss' Theorem

§41 Friday, December 9, 2022

Today is the second-to-last lecture and the last day we will cover new material. Monday will be a review day for the final exam. Fill these notes in ASAP.