

Multivariable Calculus

Azure

2022

Contents

1	Day 1: Vectors	3
1.1	Vectors	3
1.2	Dot Product	4
1.3	Cross Product	6
1.4	Distances and Planes	8
1.5	Matrices	10
1.6	Different Coordinate Systems	14
2	Day 2: Differentiation	15
2.1	Limits & Continuity	15
2.2	Differentiability	18
2.3	Differentiability in Higher Dimensions	19
2.4	Directional Derivative	23
3	Day 3: Extrema	25
3.1	Taylor's Theorem	25
3.2	Extrema	26
3.3	Lagrange Multipliers	27
4	Day 4: Integration	28
4.1	Double Integral	28
4.2	Triple Integral	30
4.3	Change of Variables	33
5	Day 5: Line Integrals	35
5.1	Parametrized Curves	35
5.2	Vector Fields	35
5.3	Line Integrals	36
5.4	Green's Theorem	37
5.5	Conservative Vector Fields	37
6	Day 6: Surface Integrals	38
6.1	Parametrized Surfaces	38
6.2	Surface Integrals	38
6.3	Stokes's & Gauss's Theorems	39

Definition 0.0.1: Common Notation

The notes will be using common mathematical notations.

\mathbb{Z}	The set of all integers
\mathbb{Q}	The set of all rational numbers
\mathbb{R}	The set of all real numbers
\mathbb{C}	The set of all complex numbers
$\{x_1, \dots, x_n\}$	Finite set with elements $x_1, x_2, x_3, \dots, x_n$
$x \in E$	x is an element of set E
$x \notin E$	x is not an element of set E
$A \subset B$	Set A is a subset set E
$A \not\subset B$	Set A is not a subset set E
$f: A \rightarrow B$	Function f maps set A into set B

To note, most theorems will include proofs, but some are beyond the scope of this course and thus, not included. However, none of the proofs are required in order for these theorems to be applied.

1 Vectors

1.1 Vectors

Definition 1.1.1: Vectors

A **scalar** c is a number in \mathbb{R} .

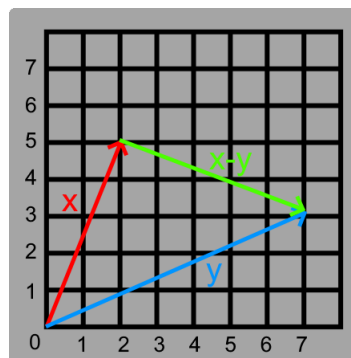
A **vector** $x \in \mathbb{R}^n$ is an ordered n -tuple of real numbers.

$$x = (x_1, \dots, x_n) = \langle x_1, \dots, x_n \rangle \quad \text{where each } x_i \in \mathbb{R}$$

Let the zero vector $0 = (0, \dots, 0)$.

If $x, y \in \mathbb{R}^n$ and c is a scalar:

Comparison:	$x = y$ if $x_i = y_i$ for $i = \{1, \dots, n\}$
Vector Addition:	$x + y = (x_1 + y_1, \dots, x_n + y_n)$
Scalar Multiplication:	$cx = (cx_1, \dots, cx_n)$

**Theorem 1.1.2: Vector Operations**

(a) $x + y = y + x$

Proof

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

(b) $x + (y + z) = (x + y) + z$

Proof

$$\begin{aligned} x + (y + z) &= (x_1, \dots, x_n) + (y_1 + z_1, \dots, y_n + z_n) = (x_1 + y_1 + z_1, \dots, x_n + y_n + z_n) \\ &= (x_1 + y_1, \dots, x_n + y_n) + (z_1, \dots, z_n) = (x + y) + z \end{aligned}$$

(c) $x+0 = x$

Proof

$$x+0 = (x_1 + 0, \dots, x_n + 0) = (x_1, \dots, x_n) = x$$

(d) $c(x+y) = cx + cy$

Proof

$$c(x+y) = (c(x_1 + y_1), \dots, c(x_n + y_n)) = (cx_1 + cy_1, \dots, cx_n + cy_n) = cx + cy$$

(e) $(c+k)v = cv + kv$

Proof

$$(c+k)v = ((c+k)v_1, \dots, (c+k)v_n) = (cv_1 + kv_1, \dots, cv_n + kv_n) = cv + kv$$

(f) $c(kx) = (ck)x = k(cx)$

Proof

$$\begin{aligned} c(kv) &= (c(kx_1), \dots, c(kx_n)) = (ckx_1, \dots, ckx_n) = (ck)x = (kcx_1, \dots, kcx_n) \\ &= (k(cx_1), \dots, k(cx_n)) = k(cx) \end{aligned}$$

Definition 1.1.3: Standard Basis Vectors

The **standard basis vectors** for \mathbb{R}^n are e_1, \dots, e_n where each $i = \{1, \dots, n\}$:

$$e_i = (0, 0, \dots, 0, 1, 0, \dots, 0, 0, \dots)$$

1 2 i-1 i i+1 n-1 n

Thus, for any $x \in \mathbb{R}^n$:

$$x = (x_1, \dots, x_n) = x_1 e_1 + \dots + x_n e_n$$

1.2 Dot Product**Definition 1.2.1: Dot Product, Norm, and Orthogonality**

The **dot product** of $x, y \in \mathbb{R}^n$ is the sum of the products of their components:

$$x \cdot y = x_1 y_1 + \dots + x_n y_n$$

The length of $x \in \mathbb{R}^n$ is the **norm**:

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{x \cdot x} \quad \Rightarrow \quad x \cdot x = \|x\|^2$$

Thus, $\|cx\| = \sqrt{(cx_1)^2 + \dots + (cx_n)^2} = |c| \sqrt{x_1^2 + \dots + x_n^2} = |c| \|x\|$.

Then, a **unit vector** (i.e. vector of length 1) in the direction of x is $\frac{x}{\|x\|}$.

$x, y \in \mathbb{R}^n$ are **orthogonal** (i.e. perpendicular) if:

$$x \cdot y = 0$$

Theorem 1.2.2: Properties of the Dot Product

(a) $x \cdot x \geq 0$

Proof

$$x \cdot x = x_1 x_1 + \dots + x_n x_n = x_1^2 + \dots + x_n^2 \geq 0 + \dots + 0 = 0$$

(b) $x \cdot x = 0$ if and only if $x = 0$

Proof

$$\begin{aligned} x \cdot x &= x_1 x_1 + \dots + x_n x_n = x_1^2 + \dots + x_n^2 \\ \text{Thus, } x \cdot x &= 0 \text{ if and only if each } x_i^2 = 0 \text{ so each } x_i = 0. \text{ Thus, } x = 0. \end{aligned}$$

(c) $x \cdot y = y \cdot x$

Proof

$$x \cdot y = x_1 y_1 + \dots + x_n y_n = y_1 x_1 + \dots + y_n x_n = y \cdot x$$

(d) $x \cdot (y + z) = x \cdot y + x \cdot z$

Proof

$$\begin{aligned} x \cdot (y + z) &= x_1(y_1 + z_1) + \dots + x_n(y_n + z_n) = (x_1y_1 + x_1z_1) + \dots + (x_ny_n + x_nz_n) \\ &= (x_1y_1 + \dots + x_ny_n) + (x_1z_1 + \dots + x_nz_n) = x \cdot y + x \cdot z \end{aligned}$$

(e) $(x + y) \cdot z = x \cdot z + y \cdot z$

Proof

$$\begin{aligned} (x + y) \cdot z &= (x_1 + y_1)z_1 + \dots + (x_n + y_n)z_n = (x_1z_1 + y_1z_1) + \dots + (x_nz_n + y_nz_n) \\ &= (x_1z_1 + \dots + x_nz_n) + (y_1z_1 + \dots + y_nz_n) = x \cdot z + y \cdot z \end{aligned}$$

(f) $cx \cdot y = c(x \cdot y) = x \cdot cy$

Proof

$$\begin{aligned} cx \cdot y &= (cx_1)y_1 + \dots + (cx_n)y_n = c(x_1y_1) + \dots + c(x_ny_n) = c(x \cdot y) \\ &= x_1(cy_1) + \dots + x_n(cy_n) = x \cdot cy \end{aligned}$$

Theorem 1.2.3: $x \cdot y = \|x\| \|y\| \cos(\theta)$

For $x, y \in \mathbb{R}^n$:

$$x \cdot y = \|x\| \|y\| \cos(\theta)$$

where $\theta \in [0, \pi]$ is the angle between x and y

Proof

Since x , y , and $x - y$ form a triangle, by the Law of Cosine:

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos(\theta)$$

where $\theta \in [0, \pi]$ is the angle between x and y . Since:

$$\|x - y\|^2 = (x - y) \cdot (x - y) = x \cdot x + y \cdot y - 2(x \cdot y) = \|x\|^2 + \|y\|^2 - 2(x \cdot y)$$

then $x \cdot y = \|x\| \|y\| \cos(\theta)$.

Theorem 1.2.4: Vector Projection

The projection of $x \in \mathbb{R}^n$ onto $y \in \mathbb{R}^n$ is the component of x parallel to y :

$$\text{proj}_y x = \frac{x \cdot y}{\|y\|^2} y$$

Proof

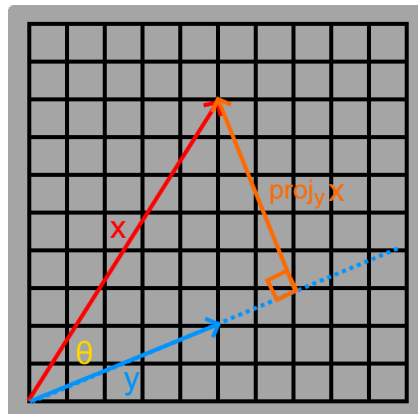
Since $\text{proj}_y x$ is parallel to y , let $\text{proj}_y x = cy$ for some constant $c \in \mathbb{R}$.

Let y^\perp be the orthogonal component of x to y . Thus, $x = \text{proj}_y x + y^\perp = cy + y^\perp$.

Since y^\perp is orthogonal to y , then:

$$x \cdot y = (cy + y^\perp) \cdot y = cy \cdot y + y^\perp \cdot y = cy \cdot y = c\|y\|^2$$

Thus, $c = \frac{x \cdot y}{\|y\|^2}$ so $\text{proj}_y x = cy = \frac{x \cdot y}{\|y\|^2} y$.



Theorem 1.2.5: Cauchy-Schwarz Inequality

For $x, y \in \mathbb{R}^n$, $|x \cdot y| \leq \|x\| \|y\|$

Proof

Let $y = \text{proj}_x y + x^\perp = cx + x^\perp$ where x^\perp is the orthogonal component of y to x and $\text{proj}_x y = cx$ is the parallel component to x for some $c \in \mathbb{R}$.

$$x \cdot y = x \cdot (cx + x^\perp) = c(x \cdot x) + x \cdot x^\perp = c\|x\|^2 + 0 = c\|x\|^2$$

Thus, $c = \frac{x \cdot y}{\|x\|^2}$. Then:

$$\begin{aligned} \|y\|^2 &= \|cx + x^\perp\|^2 = (cx + x^\perp) \cdot (cx + x^\perp) = cx \cdot cx + x^\perp \cdot x^\perp + 2(cx \cdot x^\perp) \\ &= c^2\|x\|^2 + \|x^\perp\|^2 = \left(\frac{x \cdot y}{\|x\|^2}\right)^2 \|x\|^2 + \|x^\perp\|^2 \end{aligned}$$

$$\|x\|^2 \|y\|^2 = \|x\|^2 \left(\frac{x \cdot y}{\|x\|^2} \right)^2 \|x\|^2 + \|x\|^2 \|x^\perp\|^2 = (x \cdot y)^2 + \|x\|^2 \|x^\perp\|^2$$

Since $\|x\|^2 \|x^\perp\|^2 \geq 0$, then $(x \cdot y)^2 \leq \|x\|^2 \|y\|^2$ so $|x \cdot y| \leq \|x\| \|y\|$.

Theorem 1.2.6: Triangle Inequality

For $x, y \in \mathbb{R}^n$, $\|x + y\| \leq \|x\| + \|y\|$

Proof

$$\begin{aligned} \|x + y\|^2 &= (x + y) \cdot (x + y) = x \cdot x + y \cdot y + 2(x \cdot y) = \|x\|^2 + \|y\|^2 + 2(x \cdot y) \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2 \end{aligned}$$

1.3 Cross Product**Definition 1.3.1: Cross Product**

The **cross product** of $x, y \in \mathbb{R}^3$ is the determinant of the standard basis, x, y :

$$x \times y = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} e_3$$

Theorem 1.3.2: Properties of the Cross Product

(a) $x \times y = -(y \times x)$

Proof

$$\begin{aligned} x \times y &= \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} e_3 \\ &= - \begin{vmatrix} y_2 & y_3 \\ x_2 & x_3 \end{vmatrix} e_1 + \begin{vmatrix} y_1 & y_3 \\ x_1 & x_3 \end{vmatrix} e_2 - \begin{vmatrix} y_1 & y_2 \\ x_1 & x_2 \end{vmatrix} e_3 = -(y \times x) \end{aligned}$$

(b) $x \times (y + z) = x \times y + x \times z$

Proof

$$\begin{aligned} x \times (y + z) &= \begin{vmatrix} x_2 & x_3 \\ y_2 + z_2 & y_3 + z_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 + z_1 & y_3 + z_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 + z_1 & y_2 + z_2 \end{vmatrix} e_3 \\ &= \left(\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ z_2 & z_3 \end{vmatrix} \right) e_1 - \left(\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \begin{vmatrix} x_1 & x_3 \\ z_1 & z_3 \end{vmatrix} \right) e_2 + \left(\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} \right) e_3 \\ &= \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} e_3 + \begin{vmatrix} x_2 & x_3 \\ z_2 & z_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ z_1 & z_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} e_3 \\ &= x \times y + x \times z \end{aligned}$$

(c) $(x + y) \times z = x \times z + y \times z$

Proof

$$\begin{aligned} (x + y) \times z &= -[z \times (x + y)] = -[z \times x + z \times y] \\ &= -[-(x \times z) + -(y \times z)] = x \times z + y \times z \end{aligned}$$

(d) $c(x \times y) = cx \times y = x \times cy$

Proof

$$\begin{aligned} c(x \times y) &= c \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} e_1 - c \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} e_2 + c \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} e_3 \\ &= \begin{vmatrix} cx_2 & cx_3 \\ y_2 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} cx_1 & cx_3 \\ y_1 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} cx_1 & cx_2 \\ y_1 & y_2 \end{vmatrix} e_3 = cx \times y \\ &= \begin{vmatrix} x_2 & x_3 \\ cy_2 & cy_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ cy_1 & cy_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ cy_1 & cy_2 \end{vmatrix} e_3 = x \times cy \end{aligned}$$

(e) $x \times x = 0$

Proof

$$\text{By part (a), } x \times x = -(x \times x) \quad \rightarrow \quad 2(x \times x) = 0 \quad \rightarrow \quad x \times x = 0$$

Theorem 1.3.3: Orthogonality of $x \times y$

$x \times y$ is orthogonal to x and y

Proof

$$\begin{aligned} x \times y \cdot x &= \left(\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}, -\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \right) \cdot (x_1, x_2, x_3) \\ &= (x_2y_3 - x_3y_2)x_1 - (x_1y_3 - x_3y_1)x_2 + (x_1y_2 - x_2y_1)x_3 \\ &= x_1x_2y_3 - x_1x_3y_2 - x_1x_2y_3 + x_2x_3y_1 + x_1x_3y_2 - x_2x_3y_1 = 0 \\ x \times y \cdot y &= \left(\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}, -\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \right) \cdot (y_1, y_2, y_3) \\ &= (x_2y_3 - x_3y_2)y_1 - (x_1y_3 - x_3y_1)y_2 + (x_1y_2 - x_2y_1)y_3 \\ &= x_2y_1y_3 - x_3y_1y_2 - x_1y_2y_3 + x_3y_1y_2 + x_1y_2y_3 - x_2y_1y_3 = 0 \end{aligned}$$

Theorem 1.3.4: $\|x \times y\| = \|x\| \|y\| \sin(\theta)$

For $x, y \in \mathbb{R}^3$:

$$\|x \times y\| = \|x\| \|y\| \sin(\theta)$$

where $\theta \in [0, \pi]$ is the angle between x and y

Proof

By **theorem 1.2.3**, $x \cdot y = \|x\| \|y\| \cos(\theta)$ where $\theta \in [0, \pi]$ is the angle between x, y .

$$\|x\|^2 \|y\|^2 - (x \cdot y)^2 = \|x\|^2 \|y\|^2 (1 - \cos^2(\theta)) = \|x\|^2 \|y\|^2 \sin^2(\theta)$$

Also:

$$\begin{aligned} \|x\|^2 \|y\|^2 - (x \cdot y)^2 &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\ &= (x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2 + x_1^2y_2^2 + x_1^2y_3^2 + x_2^2y_1^2 + x_2^2y_3^2 + x_3^2y_1^2 + x_3^2y_2^2 \\ &\quad - x_1^2y_1^2 - x_2^2y_2^2 - x_3^2y_3^2 - 2x_1x_2y_1y_2 - 2x_1x_3y_1y_3 - 2x_2x_3y_2y_3) \\ &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 = \|x \times y\|^2 \end{aligned}$$

Thus, $\|x \times y\| = \|x\| \|y\| \sin(\theta)$.

Theorem 1.3.5: Area of Parallelogram

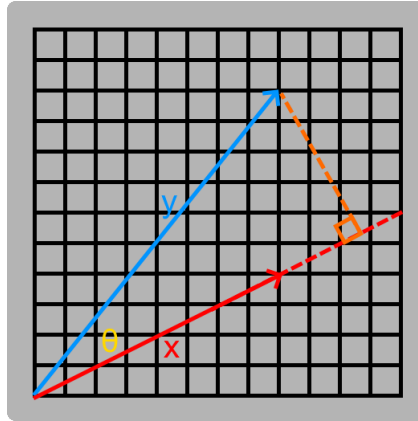
The area of a parallelogram P with sides $x, y \in \mathbb{R}^3$:

$$\text{Vol}_2(P(x, y)) = \|x \times y\|$$

Proof

Since parallelogram P with sides x and y is two triangles with sides x and y , then:

$$\begin{aligned} \text{Vol}_2(P(x, y)) &= 2 * \text{Vol}_2(\text{Triangle}(x, y)) \\ &= 2 * \frac{1}{2} (\text{base of triangle}) * (\text{height of triangle}) \\ &= \|x\| * (\|y\| \sin(\theta)) = \|x \times y\| \end{aligned}$$

**Theorem 1.3.6: Volume of Parallelepiped**

The volume of a parallelepiped P with sides $x, y, z \in \mathbb{R}^n$:

$$\text{Vol}_3(P(x, y, z)) = |(x \times y) \cdot z|$$

Proof

Let sides x and y form a base for P .

$$\text{Vol}_3(P(x, y, z)) = (\text{Area of base}) * (\text{height}) = \|x \times y\| * (\|z\| \cos(\theta))$$

where $\theta \in [0, \pi]$ is the angle between $x \times y$ and z . By **theorem 1.2.3**:

$$\text{Vol}_3(P(x, y, z)) = (x \times y) \cdot z$$

Since $-1 \leq \cos(\theta) \leq 1$ for $\theta \in [0, 2\pi]$, then $(x \times y) \cdot z$ can be negative. Thus:

$$\text{Vol}_3(P(x, y, z)) = |(x \times y) \cdot z|$$

1.4 Distances and Planes**Theorem 1.4.1: Equation of a Plane: Method #1: Point and Normal Vector**

A plane in \mathbb{R}^3 through a point $p = (p_x, p_y, p_z)$ and orthogonal to a vector called a normal vector $n = (a, b, c)$ has an equation of the form:

$$n \cdot [(x, y, z) - p] = a(x - p_x) + b(y - p_y) + c(z - p_z) = 0$$

Proof

Let (x, y, z) be any point in the plane. Then $(x, y, z) - p = (x - p_x, y - p_y, z - p_z)$ is a vector parallel to the plane. Since the plane is orthogonal to vector n , then any vector parallel to the plane is orthogonal to n . Thus:

$$n \cdot (x - p_x, y - p_y, z - p_z) = 0$$

$$a(x - p_x) + b(y - p_y) + c(z - p_z) = 0$$

Theorem 1.4.2: Equation of a Plane: Method #2: 3 Points

A plane in \mathbb{R}^3 through points $p_1 = (x_1, y_1, z_1)$, $p_2 = (x_2, y_2, z_2)$, and $p_3 = (x_3, y_3, z_3)$ has an equation of the form:

$$[(p_2 - p_1) \times (p_3 - p_1)] \cdot [(x, y, z) - p_1] = 0$$

Proof

Since p_1 , p_2 , and p_3 are on the plane, then $p_2 - p_1$ and $p_3 - p_1$ are vectors on the plane and thus, parallel to the plane. Since $(p_2 - p_1) \times (p_3 - p_1)$ is orthogonal to $(p_2 - p_1)$ and $(p_3 - p_1)$, then $(p_2 - p_1) \times (p_3 - p_1)$ is orthogonal to the plane and thus, a normal vector. By **theorem 1.4.1**, then:

$$[(p_2 - p_1) \times (p_3 - p_1)] \cdot [(x, y, z) - p_1] = 0$$

Theorem 1.4.3: Distance: Point + Line or 2 Parallel Lines

The distance from line $L(t) = tv + x_0$ to point $p \in \mathbb{R}^3$ where $t \in \mathbb{R}$, $v, x_0 \in \mathbb{R}^3$:

$$\frac{\|v \times (p - x_0)\|}{\|v\|}$$

If line $L_2(t)$ is parallel to $L(t)$, choose a point on $L_2(t)$ and apply formula above to get the distance between two parallel lines.

Proof

Since x_0 is a point on $L(t)$, then $p - x_0$ is a vector from line $L(t)$ to p .

Let θ be the angle between $p - x_0$ and $L(t)$. Thus:

$$\sin(\theta) = \frac{d}{\|p - x_0\|} \Rightarrow d = \|p - x_0\| \sin(\theta) = \frac{\|v\| \|p - x_0\| \sin(\theta)}{\|v\|} = \frac{\|v \times (p - x_0)\|}{\|v\|}$$

Theorem 1.4.4: Distance: Parallel Planes

The distance between parallel planes $P_1: a(x - x_1) + b(y - y_1) + c(z - z_1) = 0$ and $P_2: a(x - x_2) + b(y - y_2) + c(z - z_2) = 0$:

$$d = \frac{|(a,b,c) \cdot (x_2 - x_1, y_2 - y_1, z_2 - z_1)|}{\sqrt{a^2 + b^2 + c^2}}$$

Proof

Planes P_1 and P_2 are parallel since they both have the normal vector $n = (a, b, c)$. Since (x_1, y_1, z_1) is a point on P_1 and (x_2, y_2, z_2) is a point on P_2 , then $(x_2, y_2, z_2) - (x_1, y_1, z_1)$ is a vector from P_1 to P_2 .

Then the distance is the norm of the orthogonal component of $(x_2, y_2, z_2) - (x_1, y_1, z_1)$ to P_1, P_2 . Since normal vector n is orthogonal to both planes, then the orthogonal component of $(x_2, y_2, z_2) - (x_1, y_1, z_1)$ and n are parallel.

Thus, by **theorem 1.2.4**:

$$d = \|\text{proj}_n[(x_2, y_2, z_2) - (x_1, y_1, z_1)]\| = \left\| \frac{[(x_2, y_2, z_2) - (x_1, y_1, z_1)] \cdot (a, b, c)}{\|(a, b, c)\|^2} (a, b, c) \right\|$$

$$d = \frac{|(x_2 - x_1, y_2 - y_1, z_2 - z_1) \cdot (a, b, c)|}{\|(a, b, c)\|}$$

Theorem 1.4.5: Distance: Skew Lines

Lines $L_1, L_2 \in \mathbb{R}^3$ are skewed if they are neither parallel or intersecting.

Let $L_1(t) = tv_1 + x_1$ and $L_2(t) = tv_2 + x_2$. The distance between L_1 and L_2 :

$$d = \frac{|(v_2 \times v_1) \cdot (x_2 - x_1)|}{\|v_2 \times v_1\|}$$

Proof

Let L_1, L_2 be in two parallel planes. Note the distance between L_1 and L_2 is the distance between the two planes.

Since $v_2 \times v_1$ is orthogonal to v_1, v_2 and v_1, v_2 are vectors parallel to each plane, then $v_2 \times v_1$ is orthogonal to each plane and thus, a normal vector. By **theorem 1.4.4**:

$$d = \frac{|(v_2 \times v_1) \cdot (x_2 - x_1)|}{\|v_2 \times v_1\|}$$

1.5 Matrices

Definition 1.5.1: Matrix

A **m by n matrix** $M_{m \times n}(\mathbb{R})$:

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \quad \text{where each } a_{ij} \in \mathbb{R}$$

A **row vector** is a 1 by n matrix:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \end{bmatrix}$$

A **column vector** is a m by 1 matrix:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix}$$

A **zero matrix** $0 \in M_{m \times n}(\mathbb{R})$:

$$0 = \begin{bmatrix} 0_{11} & 0_{12} & a_{13} & \dots & 0_{1n} \\ 0_{21} & 0_{22} & a_{23} & \dots & 0_{2n} \\ 0_{31} & 0_{32} & a_{33} & \dots & 0_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{m1} & 0_{m2} & a_{m3} & \dots & 0_{mn} \end{bmatrix}$$

Theorem 1.5.2: Matrix Operations

(a) Addition

For $A, B \in M_{m \times n}(\mathbb{R})$, then $A+B \in M_{m \times n}(\mathbb{R})$ where each a_{ij}, b_{ij} :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

(b) Scalar Multiplication

For $A \in M_{m \times n}(\mathbb{R})$, then $cA \in M_{m \times n}(\mathbb{R})$ where each a_{ij}, b_{ij} :

$$c \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}$$

(c) Multiplication

For $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times k}(\mathbb{R})$, then $AB \in M_{m \times k}(\mathbb{R})$ where each a_{ij}, b_{ij} :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i}b_{i1} & \sum_{i=1}^n a_{1i}b_{i2} & \dots & \sum_{i=1}^n a_{1i}b_{ik} \\ \sum_{i=1}^n a_{2i}b_{i1} & \sum_{i=1}^n a_{2i}b_{i2} & \dots & \sum_{i=1}^n a_{2i}b_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{mi}b_{i1} & \sum_{i=1}^n a_{mi}b_{i2} & \dots & \sum_{i=1}^n a_{mi}b_{ik} \end{bmatrix}$$

Theorem 1.5.3: Properties of Matrix Operations

(a) $A+B = B+A$

Proof

$$[A+B]_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = [B+A]_{ij}$$

(b) $A+(B+C) = (A+B)+C$

Proof

$$[A+(B+C)]_{ij} = a_{ij} + (b_{ij} + c_{ij}) = (a_{ij} + b_{ij}) + c_{ij} = [(A+B)+C]_{ij}$$

(c) $A+0 = A$

Proof

$$[A+0]_{ij} = a_{ij} + 0_{ij} = a_{ij} = [A]_{ij}$$

(d) $(c+k)A = cA + kA$

Proof

$$[(c+k)A]_{ij} = (c+k)a_{ij} = ca_{ij} + ka_{ij} = [cA]_{ij} + [kA]_{ij} = [cA + kA]_{ij}$$

(e) $c(A+B) = cA + cB$

Proof

$$[c(A+B)]_{ij} = c(a_{ij} + b_{ij}) = ca_{ij} + cb_{ij} = [cA]_{ij} + [cB]_{ij} = [cA + cB]_{ij}$$

(f) $c(kA) = (ck)A = k(cA)$

Proof

$$[c(kA)]_{ij} = c(ka_{ij}) = (ck)a_{ij} = [(ck)A]_{ij} = k(ca_{ij}) = [k(cA)]_{ij}$$

(g) $A(BC) = (AB)C$

Proof

Let $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times k}(\mathbb{R})$, and $C \in M_{k \times p}(\mathbb{R})$.
 For $u \in \{1, \dots, n\}$ and $v \in \{1, \dots, p\}$, then $[BC]_{uv} = \sum_{s=1}^k b_{us}c_{sv}$.
 Thus, for $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, p\}$:

$$\begin{aligned}
 [A(BC)]_{ij} &= \sum_{t=1}^n a_{it}[BC]_{tj} = \sum_{t=1}^n [a_{it} \sum_{s=1}^k b_{ts}c_{sj}] = \sum_{t=1}^n \sum_{s=1}^k a_{it}b_{ts}c_{sj} \\
 &= \sum_{s=1}^k \sum_{t=1}^n a_{it}b_{ts}c_{sj} = \sum_{s=1}^k [\sum_{t=1}^n a_{it}b_{ts}]c_{sj} = \sum_{s=1}^k [AB]_{is}c_{sj} = [(AB)C]_{ij}
 \end{aligned}$$

(h) $c(AB) = (cA)B = A(cB)$

Proof

Let $A \in M_{m \times n}(\mathbb{R})$ and $B \in M_{n \times k}(\mathbb{R})$. For $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, k\}$:

$$\begin{aligned}
 [c(AB)]_{ij} &= c \sum_{t=1}^n a_{it}b_{tj} = \sum_{t=1}^n ca_{it}b_{tj} = \sum_{t=1}^n (ca_{it})b_{tj} = [(cA)B]_{ij} \\
 [c(AB)]_{ij} &= c \sum_{t=1}^n a_{it}b_{tj} = \sum_{t=1}^n a_{it}cb_{tj} = \sum_{t=1}^n a_{it}(cb_{tj}) = [A(cB)]_{ij}
 \end{aligned}$$

(i) $A(B+C) = AB + AC$

Proof

Let $A \in M_{m \times n}(\mathbb{R})$ and $B, C \in M_{n \times k}(\mathbb{R})$. For $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, k\}$:

$$\begin{aligned}
 [A(B+C)]_{ij} &= \sum_{t=1}^n a_{it}[B+C]_{tj} = \sum_{t=1}^n a_{it}(b_{tj} + c_{tj}) = \sum_{t=1}^n a_{it}b_{tj} + a_{it}c_{tj} \\
 &= \sum_{t=1}^n a_{it}b_{tj} + \sum_{t=1}^n a_{it}c_{tj} = [AB]_{ij} + [AC]_{ij} = [AB + AC]_{ij}
 \end{aligned}$$

(j) $(A+B)C = AC + BC$

Proof

Let $A, B \in M_{m \times n}(\mathbb{R})$ and $C \in M_{n \times k}(\mathbb{R})$.
 For $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, k\}$, the ij -th entry for $(A+B)C$:

$$\begin{aligned}
 [(A+B)C]_{ij} &= \sum_{t=1}^n [A+B]_{it}c_{tj} = \sum_{t=1}^n (a_{it} + b_{it})c_{tj} = \sum_{t=1}^n a_{it}c_{tj} + b_{it}c_{tj} \\
 &= \sum_{t=1}^n a_{it}c_{tj} + \sum_{t=1}^n b_{it}c_{tj} = [AC]_{ij} + [BC]_{ij} = [AC + BC]_{ij}
 \end{aligned}$$

Definition 1.5.4: Transpose

For matrix $A \in M_{m \times n}(\mathbb{R})$:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

then the **transpose**, $A^T \in M_{n \times m}(\mathbb{R})$:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \dots & a_{m3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{mn} \end{bmatrix}$$

Theorem 1.5.5: Properties of the Transpose

(a) $(A^T)^T = A$

Proof

$$[(A^T)^T]_{ij} = [A^T]_{ji} = [A]_{ij}$$

(b) $(AB)^T = B^T A^T$

Proof

$$\text{Let } A \in M_{m \times n}(\mathbb{R}) \text{ and } B \in M_{n \times k}(\mathbb{R}). \text{ For } i = \{1, \dots, k\} \text{ and } j = \{1, \dots, m\}: \\ [(AB)^T]_{ij} = [AB]_{ji} = \sum_{t=1}^n a_{jt} b_{ti} = \sum_{t=1}^n b_{ti} a_{jt} = \sum_{t=1}^n b_{it}^T a_{tj}^T = [B^T A^T]_{ij}$$

(c) $x \cdot y = x^T y$

Proof

$$x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = [x_1 \ x_2 \ \dots \ x_n] y = x^T y$$

Definition 1.5.6: Determinant

For $A \in M_{n \times n}(\mathbb{R})$, let $\text{prod}(A) = a_{1,j_1} * a_{2,j_2} * \dots * a_{n,j_n}$ such that for any two a_{k,j_k}, a_{p,j_p} where $k < p$, then $j_k \neq j_p$. Let $\text{prod}(A)$ be unique in the sense that no two $\text{prod}(A)$ have exactly the same $\{a_{1,j_1}, a_{2,j_2}, \dots, a_{n,j_n}\}$.

Also, for any two such a_{k,j_k}, a_{p,j_p} , let an inversion be 1 if $j_k < j_p$ and 0 if $j_k > j_p$. Then for any $\text{prod}(A)$, associate a $\text{sign}(A) = (-1)^{\text{total number of inversions in prod}(A)}$.

Then the **determinant** of A :

$$\det(A) = \sum_{\text{all prod}(A)} \text{prod}(A) * \text{sign}(A)$$

Example

$$\text{Let } A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 1 \\ 5 & -2 & 3 \end{bmatrix}.$$

$$\det(A) = (1*1*3)(-1)^0 + (1*-2*1)(-1)^1 + (-1*2*3)(-1)^1 + (-1*-2*3)(-1)^2 \\ + (5*2*1)(-1)^2 + (5*1*3)(-1)^3 = 12$$

Theorem 1.5.7: Cofactor Expansion

Let $A \in M_{n \times n}(\mathbb{R})$. Let A_{ij} be A , but the i -th row and j -th column removed.

Then for a fixed $i \in \{1, \dots, n\}$:

$$\det(A) = (-1)^{i+1}a_{i1}\det(A_{i1}) + (-1)^{i+2}a_{i2}\det(A_{i2}) + \dots + (-1)^{i+n}a_{in}\det(A_{in})$$

Or for a fixed $j \in \{1, \dots, n\}$:

$$\det(A) = (-1)^{1+j}a_{1j}\det(A_{1j}) + (-1)^{2+j}a_{2j}\det(A_{2j}) + \dots + (-1)^{n+j}a_{nj}\det(A_{nj})$$

Proof

For any n by n matrix A , each $\text{prod}(A)$ must contain n a_{ij} where each a_{ij} 's i, j is different from another a_{ij} 's i, j . Thus, each $\text{prod}(A)$ must contain only one a_{ij} in each row and column.

There are n possible a_{ij} choices in the first column and by choosing any such one, then that row is eliminated for choice in the following columns. Thus, there are $n-1$ possible a_{ij} choices in the second column and by choosing any such one, then that row is also eliminated for choice in the following columns. Repeating the pattern, then there are $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 1 = n!$ total unique $\text{prod}(A)$ combinations. In the cofactor expansion, let choose a fixed i . The case for a fixed j is analogous. For a fixed i , the cofactor expansion iterates through each of the n columns in row i so there are n unique a_{ij} . For each a_{ij} , the A_{ij} has the i -th row and j -th column removed so A_{ij} is a $(n-1)$ by $(n-1)$ matrix and thus, there are $(n-1)!$ unique $\text{prod}(A_{ij})$ combinations as proved earlier. Since each A_{ij} removes a different j -th column, then each $\text{prod}(A_{ij})$ from different columns are unique. Thus, the n unique a_{ij} has $(n-1)!$ unique $\text{prod}(A_{ij})$ combinations so there are $n \cdot (n-1)! = n!$ unique $\text{prod}(A)$ combinations. Thus, the $\text{prod}(A)$ combinations in the cofactor expansion must be equivalent to the $\text{prod}(A)$ combinations in the original determinant.

For the fixed i , let fixed $j \in \{1, \dots, n\}$:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \dots & a_{1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & a_{i-1,2} & a_{i-1,3} & \dots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i,1} & a_{i,2} & a_{i,3} & \dots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \dots & a_{i,n} \\ a_{i+1,1} & a_{i+1,2} & a_{i+1,3} & \dots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \dots & a_{n,n} \end{bmatrix}$$

In the original determinant, each $\text{prod}(A)$ associates $\text{sign}(A) = (-1)^{\# \text{ inversions in } \text{prod}(A)}$.

As proven earlier, each $\text{prod}(A)$ is expressed in the cofactor expansion. So for any $\text{prod}(A)$ that contains a_{ij} with the fixed i, j , then from the $a_{ij}\det(A_{ij})$ in the cofactor expansion, the $\det(A_{ij})$ consists of the other a_{ij} in the $\text{prod}(A)$ since none of the other a_{ij} can exist in row i or column j by definition of the determinant and thus, $\det(A_{ij})$ must account for all the inversions exclusively between the other a_{ij} . To account for the inversions between the other a_{ij} and the fixed a_{ij} , refer to the matrix above. The only a_{ij} which contributes an inversion with the fixed a_{ij} must be in the lower left and upper right of the matrix by definition of the determinant. Let $A = \#a_{ij}$ in upper left, $B = \#a_{ij}$ in upper right, $C = \#a_{ij}$ in lower left, and $D = \#a_{ij}$ in lower right.

Since each $\text{prod}(A)$ must have a a_{ij} in each row and column, then:

$$A+B = i-1 \quad A+C = j-1 \quad \Rightarrow \quad B+C = i+j-2-2A$$

Thus, $\text{sign}(A) = (-1)^{B+C} = (-1)^{i+j-2-2A} = (-1)^{i+j}(-1)^{-2}(-1)^{-2A} = (-1)^{i+j}$ which is the coefficient in the cofactor expansion and thus, the cofactor expansion is calculated in the same way as the original determinant and thus, have the same value.

1.6 Different Coordinate Systems

Definition 1.6.1: Polar Coordinates

Thus far, all vectors has been in the Cartesian (i.e. rectangular (x,y)) System. However, vectors can also be expressed in the Polar (i.e. circular) System.

For any point (x,y), a right triangle can be drawn by adding a perpendicular line from the x-axis to (x,y). Thus:

$$r = \sqrt{x^2 + y^2} \quad x = r \cos(\theta) \quad y = r \sin(\theta)$$

Thus, the **polar coordinates** can express points as (r, θ) .

To convert from polar to rectangular:

$$x = r \cos(\theta) \quad y = r \sin(\theta)$$

To convert from rectangular to polar:

$$r^2 = x^2 + y^2 \quad \tan(\theta) = \frac{y}{x}$$

Definition 1.6.2: Cylindrical Coordinates

While polar coordinates are the circular equivalent to \mathbb{R}^2 , cylindrical coordinates are the circular equivalent to \mathbb{R}^3 .

Cylindrical coordinates are expressed as (r, θ, z) where:

$$x = r \cos(\theta) \quad y = r \sin(\theta) \quad z = z$$

The standard basis vectors for cylindrical coordinates:

$$e_r = \frac{xe_1 + ye_2}{\sqrt{x^2 + y^2}} = \cos(\theta)e_1 + \sin(\theta)e_2$$

$$e_z = e_3$$

$$e_\theta = e_z \times e_r = -\sin(\theta)e_1 + \cos(\theta)e_2$$

Definition 1.6.3: Spherical Coordinates

Although way to express coordinates in \mathbb{R}^3 is spherical coordinates.

Spherical coordinates are expressed as (p, θ, ϕ) where:

$$x = p \sin(\phi) \cos(\theta) \quad y = p \sin(\phi) \sin(\theta) \quad z = p \cos(\phi)$$

To convert from rectangular to spherical:

$$p^2 = x^2 + y^2 + z^2 \quad \tan(\phi) = \frac{\sqrt{x^2 + y^2}}{z} \quad \tan(\theta) = \frac{y}{x}$$

To convert from cylindrical to spherical:

$$p^2 = r^2 + z^2 \quad \tan(\phi) = \frac{r}{z} \quad \theta = \theta$$

The standard basis vectors for spherical coordinates:

$$e_p = \frac{xe_1 + ye_2 + ze_3}{\sqrt{x^2 + y^2 + z^2}} = \sin(\phi) \cos(\theta)e_1 + \sin(\phi) \sin(\theta)e_2 + \cos(\phi)e_3$$

$$e_\theta = -\sin(\theta)e_1 + \cos(\theta)e_2$$

$$e_\phi = e_\theta \times e_p = \cos(\phi) \cos(\theta)e_1 + \cos(\phi) \sin(\theta)e_2 - \sin(\phi)e_3$$

2 Differentiation

2.1 Limits & Continuity

Definition 2.1.1: Limit

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, let $a \in X$.

If for every $\epsilon > 0$, there is a $\delta > 0$ such that for all $x \in X$ where $\|x - a\| < \delta$:

$$\|f(x) - L\| < \epsilon$$

Then the **limit** of $f(x)$ as x approaches a is $\lim_{x \rightarrow a} f(x) = L$.

Example

Let $f(x,y) = 2x^2 + xy$. Find $f(x,y)$ as $(x,y) \rightarrow (-1,1)$.

$$\begin{aligned} L &= f(-1,1) = 1. \text{ Let } \sqrt{(x+1)^2 + (y-1)^2} < \delta \text{ so } |x+1| < \delta \text{ and } |y-1| < \delta. \text{ Thus:} \\ |f(x,y) - L| &= |2x^2 + xy - 1| = |2x^2 - 2 + xy + 1| \\ &= |2(x+1)(x-1) + (x+1)(y+1) - (x+1+y-1)| \\ &\leq 2|x+1| * |x-1| + |x+1| * |y+1| + |x+1| + |y-1| \\ &< 2\delta(\delta+2) + \delta(\delta+2) + 2\delta = 3\delta^2 + 8\delta \\ \text{Since } \min(3\delta^2 + 8\delta) &= \frac{-16}{3} < 0, \text{ then for any } \epsilon > 0, \text{ there is a } \delta \text{ where } 3\delta^2 + 8\delta < \epsilon. \\ \text{Thus, } |f(x,y) - L| &< 3\delta^2 + 8\delta < \epsilon. \end{aligned}$$

Theorem 2.1.2: Limits are Unique

If $\lim_{x \rightarrow a} f(x) = L_1$ and $\lim_{x \rightarrow a} f(x) = L_2$, then $L_1 = L_2$.

Proof

$$\begin{aligned} \text{Since } \lim_{x \rightarrow a} f(x) &= L_1, \text{ there is a } \delta_1 \text{ where for } \|x - a\| < \delta_1, \text{ then } \|f(x) - L_1\| < \frac{\epsilon}{2}. \\ \text{Since } \lim_{x \rightarrow a} f(x) &= L_2, \text{ there is a } \delta_2 \text{ where for } \|x - a\| < \delta_2, \text{ then } \|f(x) - L_2\| < \frac{\epsilon}{2}. \\ \text{Let } \delta &= \min(\delta_1, \delta_2). \text{ Then for } \|x - a\| < \delta: \\ \|L_1 - L_2\| &\leq \|L_1 - f(x)\| + \|f(x) - L_2\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Theorem 2.1.3: Properties of the Limit

(a) For $f, g: \mathbb{R}^n \rightarrow \mathbb{R}^m$, if $\lim_{x \rightarrow a} f(x) = A$ and $\lim_{x \rightarrow a} g(x) = B$, then:

$$\lim_{x \rightarrow a} (f+g)(x) = A+B$$

Proof

$$\begin{aligned} \text{Since } \lim_{x \rightarrow a} f(x) &= A, \text{ there is a } \delta_1 \text{ where for } \|x - a\| < \delta_1, \text{ then:} \\ \|f(x) - A\| &< \frac{\epsilon}{2} \\ \text{Since } \lim_{x \rightarrow a} g(x) &= B, \text{ there is a } \delta_2 \text{ where for } \|x - a\| < \delta_2, \text{ then:} \\ \|g(x) - B\| &< \frac{\epsilon}{2} \\ \text{Let } \delta &= \min(\delta_1, \delta_2). \text{ Then for } \|x - a\| < \delta: \\ \|(f+g)(x) - (A+B)\| &= \|f(x) + g(x) - A - B\| \\ &\leq \|f(x) - A\| + \|g(x) - B\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

(b) For $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, if $\lim_{x \rightarrow a} f(x) = A$ and scalar $c \in \mathbb{R}$, then:

$$\lim_{x \rightarrow a} cf(x) = cA$$

Proof

$$\begin{aligned} \text{Since } \lim_{x \rightarrow a} f(x) &= A, \text{ there is a } \delta \text{ where for } \|x - a\| < \delta, \text{ then:} \\ \|f(x) - A\| &< \frac{\epsilon}{c} \\ \text{Then, } \|cf(x) - cA\| &= c\|f(x) - A\| < c\frac{\epsilon}{c} = \epsilon. \end{aligned}$$

- (c) For $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$, if $\lim_{x \rightarrow a} f(x) = A$ and $\lim_{x \rightarrow a} g(x) = B$, then:
 $\lim_{x \rightarrow a} (fg)(x) = AB$

Proof

Note $4fg = (f + g)^2 - (f - g)^2$.

By part (a), there is a δ where for $\|x - a\| < \delta$:

$$|(f + g)(x) - (A + B)| < \epsilon$$

Then as $x \rightarrow a$:

$$\begin{aligned} |[(f + g)(x)]^2 - [A + B]^2| &= |[(f + g)(x) - (A + B)][(f + g)(x) + (A + B)]| \\ &= |(f + g)(x) - (A + B)| * |(f + g)(x) + (A + B)| = \epsilon(2(A + B)) \end{aligned}$$

Thus, $\lim_{x \rightarrow a} (f + g)^2(x) = (A + B)^2$.

The proof for $\lim_{x \rightarrow a} (f - g)^2(x) = (A - B)^2$ is analogous. Thus:

$$\begin{aligned} \lim_{x \rightarrow a} (fg)(x) &= \lim_{x \rightarrow a} \frac{1}{4}[(f + g)^2(x) - (f - g)^2(x)] \\ &= \frac{1}{4}[(A + B)^2 - (A - B)^2] = \frac{1}{4}4AB = AB \end{aligned}$$

- (d) For $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$, if $\lim_{x \rightarrow a} f(x) = A$ and $\lim_{x \rightarrow a} g(x) = B \neq 0$, then:
 $\lim_{x \rightarrow a} \left(\frac{f}{g}\right)(x) = \frac{A}{B}$

Proof

Since $\lim_{x \rightarrow a} g(x) = B$, then there is a δ where for $\|x - a\| < \delta$:

$$|g(x) - B| < \epsilon$$

Thus, as $x \rightarrow a$:

$$\left|\frac{1}{g(x)} - \frac{1}{B}\right| = \left|\frac{B - g(x)}{Bg(x)}\right| = |B - g(x)| * \left|\frac{1}{Bg(x)}\right| < \epsilon \frac{1}{B^2}$$

Thus, $\lim_{x \rightarrow a} \frac{1}{g(x)} = \frac{1}{B}$. By part (c), then $\lim_{x \rightarrow a} \left(\frac{f}{g}\right)(x) = \frac{A}{B}$.

Theorem 2.1.4: Components of Limits

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, let $f(x) = (f_1(x), \dots, f_m(x))$. Then for $i = \{1, \dots, m\}$:
 $\lim_{x \rightarrow a} f(x) = L = (L_1, \dots, L_m)$ if and only if each $\lim_{x \rightarrow a} f_i(x) = L_i$

Proof

If $\lim_{x \rightarrow a} f(x) = L = (L_1, \dots, L_m)$, then there is a δ such that for $\|x - a\| < \delta$:

$$\|f(x) - L\| < \epsilon$$

$$\|(f_1(x), \dots, f_m(x)) - (L_1, \dots, L_m)\| = \sqrt{(f_1(x) - L_1)^2 + \dots + (f_m(x) - L_m)^2} < \epsilon$$

Thus, each $|f_i(x) - L_i| < \epsilon$ for $\|x - a\| < \delta$ so $\lim_{x \rightarrow a} f_i(x) = L_i$.

If each $\lim_{x \rightarrow a} f_i(x) = L_i$, then there are δ_i such that for $\|x - a\| < \delta_i$:

$$|f_i(x) - L_i| < \frac{\epsilon}{\sqrt{m}}$$

Let $\delta = \min(\delta_1, \dots, \delta_m)$. Then for $\|x - a\| < \delta$:

$$\begin{aligned} \|f(x) - L\| &= \|(f_1(x), \dots, f_m(x)) - (L_1, \dots, L_m)\| \\ &= \sqrt{(f_1(x) - L_1)^2 + \dots + (f_m(x) - L_m)^2} < \sqrt{\sum_{i=1}^m \left(\frac{\epsilon}{\sqrt{m}}\right)^2} = \sqrt{\epsilon^2} = \epsilon \end{aligned}$$

Definition 2.1.5: Continuity

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, let $a \in X$.

Then f is **continuous** at a if $\lim_{x \rightarrow a} f(x) = f(a)$.

If f is continuous at every $x \in X$, then f is continuous on X .

Theorem 2.1.6: Properties of Continuity

- (a) If
- $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$
- are continuous at
- $a \in X$
- , then
- $f+g$
- is continuous at
- a

Proof

Since $\lim_{x \rightarrow a} f(x) = f(a)$ and $\lim_{x \rightarrow a} g(x) = g(a)$, by **theorem 2.1.3(a)**, then $A = f(a)$ and $B = g(a)$. Thus, $\lim_{x \rightarrow a} (f+g)(x) = f(a)+g(a)$.

- (b) If
- $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$
- is continuous at
- $a \in X$
- and scalar
- $c \in \mathbb{R}$
- , then
- cf
- is continuous at
- a

Proof

Since $\lim_{x \rightarrow a} f(x) = f(a)$, by **theorem 2.1.3(b)**, then $A = f(a)$.
Thus, $\lim_{x \rightarrow a} cf(x) = cf(a)$.

- (c) If
- $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$
- are continuous at
- $a \in X$
- , then
- fg
- is continuous at
- a

Proof

Since $\lim_{x \rightarrow a} f(x) = f(a)$ and $\lim_{x \rightarrow a} g(x) = g(a)$, by **theorem 2.1.3(c)**, then $A = f(a)$ and $B = g(a)$. Thus, $\lim_{x \rightarrow a} (fg)(x) = f(a)g(a)$.

- (d) If
- $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$
- are continuous at
- $a \in X$
- where
- $g(x) \neq 0$
- , then
- $\frac{f}{g}$
- is continuous at
- a

Proof

Since $\lim_{x \rightarrow a} f(x) = f(a)$ and $\lim_{x \rightarrow a} g(x) = g(a)$, by **theorem 2.1.3(d)**, then $A = f(a)$ and $B = g(a) \neq 0$. Thus, $\lim_{x \rightarrow a} \left(\frac{f}{g}\right)(x) = \frac{f(a)}{g(a)}$.

Theorem 2.1.7: Components of Continuity

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, let $f(x) = (f_1(x), \dots, f_m(x))$. Then for $i = \{1, \dots, m\}$:
 f is continuous at $a \in X$ if and only if each f_i is continuous at a

Proof

If f is continuous at a , then $\lim_{x \rightarrow a} f(x) = f(a) = (f_1(a), \dots, f_m(a))$. By **theorem 2.1.4**, then $L = (f_1(a), \dots, f_m(a))$ so each $L_i = f_i(a)$. Thus, for each $i = \{1, \dots, m\}$:

$$\lim_{x \rightarrow a} f_i(x) = L_i = f_i(a)$$

If each f_i is continuous at a , then for $i = \{1, \dots, m\}$, $\lim_{x \rightarrow a} f_i(x) = f_i(a)$. By **theorem 2.1.4**, then $L = (f_1(a), \dots, f_m(a))$. Thus:

$$\lim_{x \rightarrow a} f(x) = L = (f_1(a), \dots, f_m(a)) = f(a)$$

Theorem 2.1.8: Composite of Continuous functions are Continuous

If $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g: Y \subset \mathbb{R}^m \rightarrow \mathbb{R}^k$ are continuous where $f(X) \subset Y$,
then $g \circ f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ is continuous

Proof

For any $a \in X$ and any $\delta > 0$, there is a $\eta > 0$ such that for $\|x - a\| < \eta$:

$$\|f(x) - f(a)\| < \delta$$

Since $f(X) \subset Y$, then for any $x \in X$, then $f(x) \in Y$.

For any $f(a) \in Y$ and any $\epsilon > 0$, there is a $\delta > 0$ such that for $\|y - f(a)\| < \delta$:

$$\|g(y) - g(f(a))\| < \epsilon$$

Thus, for $\|x - a\| < \eta$, then $\|g(f(x)) - g(f(a))\| < \epsilon$.

2.2 Differentiability

Definition 2.2.1: Partial Derivative

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, let $x = (x_1, \dots, x_n) \in X$.

For $i = \{1, \dots, n\}$, the **partial derivative** of f with respect to x_i :

$$D_i f = \frac{\partial f}{\partial x_i} = f_{x_i}(x) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_n)}{h}$$

Theorem 2.2.2: Tangent Plane

For $f: X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, let $z = f(x, y)$.

The **tangent plane** at $(a, b, f(a, b))$ has an equation of the form:

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

Proof

Since $f_x(a, b)$ which is the change in z for every change in x is a tangent vector to f in direction of x at (a, b) , then $(1, 0, f_x(a, b))$ is parallel to the tangent plane. Similarly, $(0, 1, f_y(a, b))$ is parallel to the tangent plane.

Thus, $(1, 0, f_x(a, b)) \times (0, 1, f_y(a, b)) = (-f_x(a, b), -f_y(a, b), 1)$ is orthogonal to the tangent plane. Thus, for any (x, y, z) in the plane:

$$\begin{aligned} (-f_x(a, b), -f_y(a, b), 1) \cdot [(x, y, z) - (a, b, f(a, b))] &= 0 \\ -f_x(a, b)(x - a) - f_y(a, b)(y - b) + z - f(a, b) &= 0 \\ z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \end{aligned}$$

Definition 2.2.3: Differentiability in $\mathbb{R}^2 \rightarrow \mathbb{R}$

$f: X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is **differentiable** at $x \in X$ if there is an $A \in M_{1 \times 2}(\mathbb{R})$ such that for $h \in X$:

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} = 0$$

Then, the **derivative** of f at x is $Df(x) = A = \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{bmatrix}$.

If f is differentiable at every $x \in X$, then f is differentiable on X .

Theorem 2.2.4: Continuous partials imply Differentiability

If $f: X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ has continuous partial derivatives at (a, b) , then f is differentiable at (a, b) .

Proof

Since $f_x(x, y), f_y(x, y)$ is continuous at (a, b) , then for $\epsilon > 0$, there is a $\delta > 0$ where for $\|(x, y) - (a, b)\| < \delta$:

$$|f_x(x, y) - f_x(a, b)| < \epsilon \quad |f_y(x, y) - f_y(a, b)| < \epsilon$$

Then for $h = h_1 e_1 + h_2 e_2$:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{|f(a+h_1, b+h_2) - f(a, b) - [f_x(a, b)h_1 + f_y(a, b)h_2]|}{\|h\|} \\ = \lim_{h \rightarrow 0} \frac{|f(a+h_1, b+h_2) - f(a+h_1, b) + f(a+h_1, b) - f(a, b) - [f_x(a, b)h_1 + f_y(a, b)h_2]|}{\|h\|} \end{aligned}$$

Since $f_x(x, y), f_y(x, y)$ exist, then by the Mean Value Theorem, there are $t_1 \in (0, h_1)$ and $t_2 \in (0, h_2)$ such that:

$$\begin{aligned} f(a+h_1, b) - f(a, b) &= h_1 * f_x(a+t_1, b) \\ f(a+h_1, b+h_2) - f(a+h_1, b) &= h_2 * f_y(a+h_1, b+t_2) \end{aligned}$$

Thus, for $\|h - (a, b)\| < \delta$:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{|f(a+h_1, b+h_2) - f(a, b) - [f_x(a, b)h_1 + f_y(a, b)h_2]|}{\|h\|} \\ = \lim_{h \rightarrow 0} \frac{|h_2 * f_y(a+h_1, b+t_2) + h_1 * f_x(a+t_1, b) - [f_x(a, b)h_1 + f_y(a, b)h_2]|}{\|h\|} \\ = \lim_{h \rightarrow 0} \frac{|h_2 * [f_y(a+h_1, b+t_2) - f_y(a, b)] + h_1 * [f_x(a+t_1, b) - f_x(a, b)]|}{\|h\|} < \lim_{h \rightarrow 0} \frac{\|h\|\epsilon + \|h\|\epsilon}{\|h\|} = 2\epsilon \end{aligned}$$

Theorem 2.2.5: Differentiability implies Continuity

If $f: X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is differentiable at (a,b) , then f is continuous at (a,b)

Proof

If f is differentiable at (a,b) , then $\lim_{h \rightarrow 0} \frac{|f((a,b)+h) - f(a,b) - Ah|}{\|h\|} = 0$.

Thus, as $h \rightarrow 0$, then $A = \frac{f((a,b)+h) - f(a,b)}{\|h\|}$. So:

$$f((a,b)+h) - f(a,b) = [f((a,b)+h) - f(a,b)] \frac{\|h\|}{\|h\|} = A\|h\| \rightarrow 0$$

Thus, f is continuous at (a,b) .

2.3 Differentiability in Higher Dimensions**Definition 2.3.1: Differentiability in $\mathbb{R}^n \rightarrow \mathbb{R}$**

Differentiability can be extended for \mathbb{R}^n .

$f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable at $x \in X$ if there is an $A \in M_{1 \times n}(\mathbb{R})$ such that for $h \in X$:

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} = 0$$

Then, the derivative of f at x is $Df(x) = A = \left[\frac{\partial f}{\partial x_1}(x) \quad \frac{\partial f}{\partial x_2}(x) \quad \dots \quad \frac{\partial f}{\partial x_n}(x) \right]$.

If f is differentiable at every $x \in X$, then f is differentiable on X .

The **gradient** of f :

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) = [Df(x)]^T$$

Definition 2.3.2: Differentiability in $\mathbb{R}^n \rightarrow \mathbb{R}^m$

Differentiability can be extended into \mathbb{R}^m .

$f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ where $f = (f_1, \dots, f_m)$ is differentiable at $x \in X$ if there is an $A \in M_{m \times n}(\mathbb{R})$ such that for $h \in X$:

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} = 0$$

Then, the derivative of f at x is $Df(x) = A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$.

If f is differentiable at every $x \in X$, then f is differentiable on X .

Theorem 2.3.3: Differentiability implies Continuity in Higher Dimensions

If $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable at $a \in X$, then f is continuous at a

Proof

Analogous to **theorem 2.2.5**. Replace (a,b) with $a = (a_1, \dots, a_n)$.

Theorem 2.3.4: Continuous partials imply differentiability in Higher Dimensions

If $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ has continuous partial derivatives, $\frac{\partial f_i}{\partial x_j}$, at $a \in X$ for $j = \{1, \dots, n\}$ and $i = \{1, \dots, m\}$, then f is differentiable at a

Proof

Analogous to **theorem 2.2.4**. Instead, $h = h_1 e_1 + \dots + h_n e_n$ where:

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} = \lim_{h \rightarrow 0} \sum_{i=1}^m \frac{|f_i(x+h) - f_i(x) - [\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x) h_j]|}{\|h\|}$$

and add each $f_i(x + h_1 e_1 + \dots + h_k e_k)$ and apply Mean Value Theorem and continuity of partial derivatives analogously as performed in **theorem 2.2.4**.

Theorem 2.3.5: Components of Differentiability

$f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ where $f = (f_1, \dots, f_m)$ is differentiable at $a \in X$ if and only if each f_i is differentiable at a for $i = \{1, \dots, m\}$

Proof

$$\text{Note } \lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} = \lim_{h \rightarrow 0} \sum_{i=1}^m \frac{|f_i(x+h) - f_i(x) - [\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x) h_j]|}{\|h\|}.$$

If f is differentiable at a , then for any $\epsilon > 0$:

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} < \epsilon$$

So $\lim_{h \rightarrow 0} \frac{|f_i(x+h) - f_i(x) - [\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x) h_j]|}{\|h\|} < \epsilon$ for each $i = \{1, \dots, m\}$ and thus, each f_i is differentiable at a .

If each $\lim_{h \rightarrow 0} \frac{|f_i(x+h) - f_i(x) - [\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x) h_j]|}{\|h\|} < \frac{\epsilon}{m}$ for $i = \{1, \dots, m\}$, then:

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{\|h\|} < \lim_{h \rightarrow 0} \sum_{i=1}^m \frac{\epsilon}{m} = \epsilon$$

Thus, f is differentiable at a .

Theorem 2.3.6: Properties of Differentiability

- (a) For $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, if f, g are differentiable at $a \in X$, then:
 $f+g$ is differentiable at a where $D(f+g)(a) = Df(a) + Dg(a)$

Proof

Since f, g are differentiable at $a \in X$, by **theorem 2.3.5**, then for $i = \{1, \dots, m\}$:

$$\begin{aligned} D(f+g)_i(a) &= \begin{bmatrix} D_1(f_i + g_i)(a) & D_2(f_i + g_i)(a) & \dots & D_n(f_i + g_i)(a) \end{bmatrix} \\ &= \begin{bmatrix} D_1 f_i(a) & D_2 f_i(a) & \dots & D_n f_i(a) \end{bmatrix} + \begin{bmatrix} D_1 g_i(a) & D_2 g_i(a) & \dots & D_n g_i(a) \end{bmatrix} \\ &= Df_i(a) + Dg_i(a) \end{aligned}$$

- (b) For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, if f is differentiable at $a \in X$ and scalar $c \in \mathbb{R}$, then:
 cf is differentiable at a where $D(cf)(a) = cDf(a)$

Proof

Since f is differentiable at $a \in X$, by **theorem 2.3.5**, then for $i = \{1, \dots, m\}$:

$$\begin{aligned} D(cf)_i(a) &= \begin{bmatrix} D_1(cf_i)(a) & D_2(cf_i)(a) & \dots & D_n(cf_i)(a) \end{bmatrix} \\ &= \begin{bmatrix} cD_1 f_i(a) & cD_2 f_i(a) & \dots & cD_n f_i(a) \end{bmatrix} = cDf_i(a) \end{aligned}$$

- (c) For $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, if f, g are differentiable at $a \in X$, then:
 fg is differentiable at a where $D(fg)(a) = Df(a)g(a) + f(a)Dg(a)$

Proof

Since f, g are differentiable at $a \in X$:

$$\begin{aligned} D(fg)(a) &= \begin{bmatrix} D_1(fg)(a) & D_2(fg)(a) & \dots & D_n(fg)(a) \end{bmatrix} \\ &= \begin{bmatrix} D_1f(a)g(a)+f(a)D_1g(a) & D_2f(a)g(a)+f(a)D_2g(a) & \dots & D_nf(a)g(a)+f(a)D_ng(a) \end{bmatrix} \\ &= \begin{bmatrix} D_1f(a)g(a) & D_2f(a)g(a) & \dots & D_nf(a)g(a) \end{bmatrix} + \begin{bmatrix} f(a)D_1g(a) & f(a)D_2g(a) & \dots & f(a)D_ng(a) \end{bmatrix} \\ &= Df(a)g(a) + f(a)Dg(a) \end{aligned}$$

- (d) For $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, if f, g are differentiable at $a \in X$ where $g(a) \neq 0$, then:
 $\frac{f}{g}$ is differentiable at a where $D(\frac{f}{g})(a) = \frac{Df(a)g(a)-f(a)Dg(a)}{[g(a)]^2}$

Proof

Since f, g are differentiable at $a \in X$:

$$\begin{aligned} D(\frac{f}{g})(a) &= \begin{bmatrix} D_1(\frac{f}{g})(a) & D_2(\frac{f}{g})(a) & \dots & D_n(\frac{f}{g})(a) \end{bmatrix} \\ &= \begin{bmatrix} \frac{D_1f(a)g(a)-f(a)D_1g(a)}{[g(a)]^2} & \frac{D_2f(a)g(a)-f(a)D_2g(a)}{[g(a)]^2} & \dots & \frac{D_nf(a)g(a)-f(a)D_ng(a)}{[g(a)]^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{D_1f(a)g(a)}{[g(a)]^2} & \frac{D_2f(a)g(a)}{[g(a)]^2} & \dots & \frac{D_nf(a)g(a)}{[g(a)]^2} \end{bmatrix} - \begin{bmatrix} \frac{f(a)D_1g(a)}{[g(a)]^2} & \frac{f(a)D_2g(a)}{[g(a)]^2} & \dots & \frac{f(a)D_ng(a)}{[g(a)]^2} \end{bmatrix} \\ &= Df(a)\frac{g(a)}{[g(a)]^2} - \frac{f(a)}{[g(a)]^2}Dg(a) = \frac{Df(a)g(a)-f(a)Dg(a)}{[g(a)]^2} \end{aligned}$$

Definition 2.3.7: Partial Derivatives of Higher Orders

The **second order partial derivative** of f in respect to x_i :

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right) = f_{x_i x_i}(x) = \lim_{h \rightarrow 0} \frac{f_{x_i}(x_1, \dots, x_i+h, \dots, x_n) - f_{x_i}(x_1, \dots, x_n)}{h}$$

The **mixed partial derivative** of f in respect to first x_i , then x_j :

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) = f_{x_i x_j}(x) = \lim_{h \rightarrow 0} \frac{f_{x_i}(x_1, \dots, x_j+h, \dots, x_n) - f_{x_i}(x_1, \dots, x_n)}{h}$$

In general, for $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, the k -th order partial derivative of f in respect to x_{i_1}, \dots, x_{i_k} in such order for $k = \{1, \dots, n\}$:

$$\begin{aligned} \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} &= \frac{\partial}{\partial x_{i_k}} \dots \frac{\partial f}{\partial x_{i_1}} = f_{x_{i_1} \dots x_{i_k}}(x) \\ &= \lim_{h \rightarrow 0} \frac{f_{x_{i_1} \dots x_{i_{k-1}}}(x_1, \dots, x_{i_k}+h, \dots, x_n) - f_{x_{i_1} \dots x_{i_{k-1}}}(x_1, \dots, x_n)}{h} \end{aligned}$$

Definition 2.3.8: Smoothness

For $k = \{1, \dots, n\}$, $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is C^k if all partial derivatives of order 1 to k exist and are continuous on X .

If f has continuous partial derivatives of all order, then f is **smooth** (i.e C^∞).

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ where $f = (f_1, \dots, f_m)$, then f is C^k if each f_i is C^k for $i = \{1, \dots, m\}$.

Theorem 2.3.9: Clairaut's Theorem

If $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is C^k , then:

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{j_1} \dots \partial x_{j_k}}$$

Proof

If the claim holds true for C^2 , then replace f with $f_{x_{i_p}}$ for any $p = \{1, \dots, k\}$ and since f is C^k , then $f_{x_{i_p}}$ is C^{k-1} and apply the theorem again. Repeating the process k times, the result holds true by induction. Now the proof for C^2 :

Since f is C^2 , then f_x, f_y, f_{xy}, f_{yx} exist and are continuous.

Let $d(x, y) = f(x + h_1, y + h_2) - f(x + h_1, y) - f(x, y + h_2) + f(x, y)$.

Since f_x exist, then by the Mean Value Theorem, there is a $t_1 \in (0, h_1)$ where:

$$d(x, y) = h_1 * (f_x(x + t_1, y + h_2) - f_x(x + t_1, y))$$

Since f_y exist, then by the Mean Value Theorem, there is a $t_2 \in (0, h_2)$ where:

$$d(x, y) = h_1 * h_2 * f_{xy}(x + t_1, y + t_2)$$

Since f_{xy} is continuous, then since $(t_1, t_2) \rightarrow (0, 0)$ as $(h_1, h_2) \rightarrow (0, 0)$:

$$\begin{aligned} f_{xy}(x, y) &= \lim_{(h_1, h_2) \rightarrow (0, 0)} f_{xy}(x + h_1, x + h_2) \\ &= \lim_{(h_1, h_2) \rightarrow (0, 0)} f_{xy}(x + t_1, x + t_2) = \lim_{(h_1, h_2) \rightarrow (0, 0)} \frac{d(x, y)}{h_1 h_2} \end{aligned}$$

Rearrange $d(x, y) = f(x + h_1, y + h_2) - f(x + h_1, y) - f(x, y + h_2) + f(x, y)$.

Since f_y exist, by the Mean Value Theorem, there is a $s_2 \in (0, h_2)$ where:

$$d(x, y) = h_2 * (f_y(x + h_1, y + s_2) - f_y(x, y + s_2))$$

Since f_x exist, by the Mean Value Theorem, there is a $s_1 \in (0, h_1)$ where:

$$d(x, y) = h_2 * h_1 * f_{yx}(x + s_1, y + s_2)$$

Since f_{yx} is continuous, then since $(s_1, s_2) \rightarrow (0, 0)$ as $(h_1, h_2) \rightarrow (0, 0)$:

$$\begin{aligned} f_{yx}(x, y) &= \lim_{(h_1, h_2) \rightarrow (0, 0)} f_{yx}(x + h_1, x + h_2) \\ &= \lim_{(h_1, h_2) \rightarrow (0, 0)} f_{xy}(x + s_1, x + s_2) = \lim_{(h_1, h_2) \rightarrow (0, 0)} \frac{d(x, y)}{h_1 h_2} \end{aligned}$$

Thus, $f_{xy}(x, y) = f_{yx}(x, y)$.

Theorem 2.3.10: Chain Rule

Let $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable at $x_0 \in X$ and $g: f(X) \subset Y \subset \mathbb{R}^m \rightarrow \mathbb{R}^k$ be differentiable at $f(x_0)$.

Then $g \circ f = g(f): X \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ is differentiable at x_0 such that:

$$D[g(f(x_0))] = Dg(f(x_0)) Df(x_0)$$

Proof

Since f is differentiable at x_0 and g is differentiable at $f(x_0)$, then there is a $A = Df(x_0)$ and $B = Dg(f(x_0))$ such that:

$$f(x_0 + h) - f(x_0) = Ah + r_A(h) \quad \text{where } \lim_{h \rightarrow 0} \frac{|r_A(h)|}{|h|} = 0$$

$$g(f(x_0) + k) - g(f(x_0)) = Bk + r_B(k) \quad \text{where } \lim_{k \rightarrow 0} \frac{|r_B(k)|}{|k|} = 0$$

Let $k = f(x_0 + h) - f(x_0)$. Thus:

$$\begin{aligned} g(f(x_0 + h)) - g(f(x_0)) &= BAh \\ &= g(f(x_0) + k) - g(f(x_0)) - BAh = Bk + r_B(k) - BAh = B(k - Ah) + r_B(k) \\ &= B(f(x_0 + h) - f(x_0) - Ah) + r_B(k) = Br_A(h) + r_B(k) \end{aligned}$$

Since f is differentiable at x_0 , then f is continuous at x_0 and thus, $\lim_{h \rightarrow 0} k = 0$.

Since $\lim_{h \rightarrow 0} \frac{|r_A(h)|}{|h|} = 0$ and $\lim_{k \rightarrow 0} \frac{|r_B(k)|}{|k|} = 0$, then:

$$\lim_{h \rightarrow 0} \frac{|g(f(x_0 + h)) - g(f(x_0)) - BAh|}{|h|} \leq \lim_{h \rightarrow 0} (|B| \frac{|r_A(h)|}{|h|} + \frac{|r_B(k)|}{|h|}) = 0 + 0 = 0$$

Thus, $D[g(f(x_0))] = BA = Dg(f(x_0)) Df(x_0)$.

Theorem 2.3.11: Relationship between rectangular and polar partials

For $(x, y) = (r \cos(\theta), r \sin(\theta))$:

$$\frac{\partial}{\partial r} = \cos(\theta) \frac{\partial}{\partial x} + \sin(\theta) \frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial \theta} = -r \sin(\theta) \frac{\partial}{\partial x} + r \cos(\theta) \frac{\partial}{\partial y}$$

Thus:

$$\frac{\partial}{\partial x} = \cos(\theta) \frac{\partial}{\partial r} - \frac{\sin(\theta)}{r} \frac{\partial}{\partial \theta}$$

$$\frac{\partial}{\partial y} = \sin(\theta) \frac{\partial}{\partial r} + \frac{\cos(\theta)}{r} \frac{\partial}{\partial \theta}$$

Proof

Let $z = g(r, \theta)$. Then let $z = f(x, y)$ such that $(r \cos(\theta), r \sin(\theta)) = (x, y)$.

By **theorem 2.3.10**:

$$D[g(r, \theta)] = D[f(x, y)] D(x(r, \theta), y(r, \theta))$$

$$\begin{bmatrix} \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) \frac{\partial f}{\partial x} + \sin(\theta) \frac{\partial f}{\partial y} \\ -r \sin(\theta) \frac{\partial f}{\partial x} + r \cos(\theta) \frac{\partial f}{\partial y} \end{bmatrix}$$

Thus:

$$\frac{\partial}{\partial r} = \cos(\theta) \frac{\partial}{\partial x} + \sin(\theta) \frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial \theta} = -r \sin(\theta) \frac{\partial}{\partial x} + r \cos(\theta) \frac{\partial}{\partial y}$$

Then:

$$-r \cos(\theta) \frac{\partial}{\partial r} + \sin(\theta) \frac{\partial}{\partial \theta} = -r \cos^2(\theta) \frac{\partial}{\partial x} - r \sin^2(\theta) \frac{\partial}{\partial x} = -r \frac{\partial}{\partial x}$$

$$r \sin(\theta) \frac{\partial}{\partial r} + \cos(\theta) \frac{\partial}{\partial \theta} = r \sin^2(\theta) \frac{\partial}{\partial y} + r \cos^2(\theta) \frac{\partial}{\partial y} = r \frac{\partial}{\partial y}$$

2.4 Directional Derivative**Definition 2.4.1: Directional Derivative**

Let $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable at $a \in X$. Then the **directional derivative** of f at a in the direction of vector $v \in \mathbb{R}^n$:

$$D_v f(a) = \lim_{h \rightarrow 0} \frac{f(a+hv) - f(a)}{\|hv\|}$$

Theorem 2.4.2: Relationship between Directional Derivative and Gradient

Let $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable at $a \in X$. Then the directional derivative of f at a in the direction of vector $v \in \mathbb{R}^n$:

$$D_v f(a) = \nabla f(a) \cdot \frac{v}{\|v\|}$$

If v is a unit vector, then $D_v f(a) = \nabla f(a) \cdot v$.

Proof

Let $y(t) = a + tv$ for $t \in (-\infty, \infty)$. Then by **theorem 2.3.10**:

$$\begin{aligned} D_v f(a) &= \lim_{h \rightarrow 0} \frac{f(a+hv) - f(a)}{\|hv\|} = \lim_{h \rightarrow 0} \frac{f(y(h)) - f(y(0))}{|h|} \frac{1}{\|v\|} = Df(y(0)) Dy(0) \frac{1}{\|v\|} \\ &= Df(a) \frac{v}{\|v\|} = [Df(a)]^T \cdot \frac{v}{\|v\|} = \nabla f(a) \cdot \frac{v}{\|v\|} \end{aligned}$$

Theorem 2.4.3: Direction of Steepest Ascent

The directional derivative $D_v f(a) = \nabla f(a) \cdot \frac{v}{\|v\|}$ is:

Maximized when v is in the same direction as $\nabla f(a)$ with value $\|\nabla f(a)\|$

Minimized when v is in the opposite direction of $\nabla f(a)$ with value $-\|\nabla f(a)\|$

Proof

By **theorem 1.2.3**, $D_v f(a) = \nabla f(a) \cdot \frac{v}{\|v\|} = \|\nabla f(a)\| \left\| \frac{v}{\|v\|} \right\| \cos(\theta) = \|\nabla f(a)\| \cos(\theta)$. where $\theta \in [0, \pi]$ is the angle between $\nabla f(a)$ and $\left| \frac{v}{\|v\|} \right|$. Since $D_v f(a)$ is maximized at $\|\nabla f(a)\|$ when $\theta = 0$, then $\nabla f(a)$ and v points in the same direction. Also, $D_v f(a)$ is minimized at $-\|\nabla f(a)\|$ when $\theta = \pi$, then $\nabla f(a)$ and v points in opposite directions.

Theorem 2.4.4: Gradient is orthogonal to the surface

If $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is C^1 , then for any x_0 where $f(x_0) = c$ for constant $c \in \mathbb{R}$, then $\nabla f(x_0)$ is orthogonal to the surface $f(x) = c$ at x_0 .

Proof

For surface $f(x) = c$, let curve $C(t) = (x_1(t), \dots, x_n(t))$ where $C(0) = x_0$ be defined such that $f(C(t)) = c$. Thus:

$$\frac{d}{dt} f(C(t)) = \frac{d}{dt} c = 0$$
Let v be the tangent vector to $C(t)$ at x_0 . Then by **theorem 2.3.10**, for $t = 0$:

$$\frac{d}{dt} f(C(0)) = Df(C(0)) C'(0) = \nabla f(C(0)) \cdot C'(0) = \nabla f(x_0) \cdot v$$
Since $\nabla f(x_0) \cdot v = 0$ where v is tangent to $C(t)$ which lies on surface $f(x) = c$ and thus, is tangent is $f(x) = c$, then $\nabla f(x_0)$ is orthogonal to $f(x) = c$ at x_0 .

3 Extrema

3.1 Taylor's Theorem

Theorem 3.1.1: Taylor's Theorem for one variable

Let $f: X \subset \mathbb{R} \rightarrow \mathbb{R}$ be C^{k+1} .

Then for $a \in X$, the k -th order Taylor polynomial of f centered at a :

$$p_k(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(k)}(a)}{k!}(x-a)^k$$

Then for $h > 0$, there is a $x \in [a, a+h]$ such that:

$$f(a+h) = p_k(a+h) + \frac{f^{(k+1)}(x)}{(k+1)!}h^{k+1}$$

Proof

Let M be defined such that $f(a+h) = p_k(a+h) + M(a+h-a)^{k+1}$.

Let $g(t) = f(t) - p_k(t) - M(t-a)^{k+1}$ for $t \in [a, a+h]$. Then:

$$g^{(k+1)}(t) = f^{(k+1)}(t) - (k+1)!M$$

Since $p_k^{(i)}(a) = f^{(i)}(a)$ for $i = \{0, \dots, k\}$, then:

$$g^{(i)}(a) = f^{(i)}(a) - p_k^{(i)}(a) - \frac{(k+1)!}{(k+1-i)!}M(a-a)^{k+1-i} = 0$$

Since $g(a+h) = 0$ where f is differentiable so g is differentiable, then by the Mean Value Theorem, there is a $t_1 \in (0, h)$:

$$0 - 0 = g(a+h) - g(a) = h * g'(a+t_1) \Rightarrow g'(a+t_1) = 0$$

Since g' is differentiable, by the Mean Value Theorem, there is a $t_2 \in (0, t_1)$:

$$0 - 0 = g'(a+t_1) - g'(a) = t_1 * g''(a+t_2) \Rightarrow g''(a+t_2) = 0$$

Repeating the process $k+1$ times, there is a $t_{k+1} \in (0, t_k)$:

$$0 - 0 = g^{(k)}(a+t_k) - g^{(k)}(a) = t_k * g^{(k+1)}(a+t_{k+1}) \Rightarrow g^{(k+1)}(a+t_{k+1}) = 0$$

Since $t_{k+1} < t_k < \dots < t_1 < h$, then:

$$0 = g^{(k+1)}(a+t_{k+1}) = f^{(k+1)}(a+t_{k+1}) - (k+1)!M$$

Thus, $M = \frac{f^{(k+1)}(a+t_{k+1})}{(k+1)!}$ where $a+t_{k+1} \in [a, a+h]$.

Theorem 3.1.2: Taylor's Theorem for multiple variables

Let $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be C^2 . Then for $a \in X$:

$$p_k(x) = f(a) + \sum_{i=1}^n f_{x_i}(a)(x_i - a_i) + \frac{1}{2} \sum_{i,j=1}^n f_{x_i x_j}(a)(x_i - a_i)(x_j - a_j) + \dots + \frac{1}{k!} \sum_{i_1, \dots, i_k=1}^n f_{x_{i_1} \dots x_{i_k}}(a)(x_{i_1} - a_{i_1}) \dots (x_{i_k} - a_{i_k})$$

Then:

$$f(x) = p_k(x) + R(a) \quad \text{where } \lim_{x \rightarrow a} \frac{R(a)}{\|x-a\|} = 0$$

3.2 Extrema

Definition 3.2.1: Local Extrema

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, let $a \in X$.

f has a **local minimum** at a if there is a $\delta > 0$ such that:

$$f(x) \geq f(a) \text{ for all } x \in X \text{ where } \|x - a\| < \delta$$

f has a **local maximum** at a if there is a $\delta > 0$ such that:

$$f(x) \leq f(a) \text{ for all } x \in X \text{ where } \|x - a\| < \delta$$

Theorem 3.2.2: Local extrema have a derivative of 0

For differentiable $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, let $a \in X$.

If f has a local extrema at a , then $Df(a) = 0$.

Proof

Let f have a local maximum at a . The proof for local minimum is analogous.
 For any $i \in \{1, \dots, n\}$, let $F(t) = f(a + te_i)$. Since $F: \mathbb{R} \rightarrow \mathbb{R}$ has a local maximum at $F(0)$, then $0 = F'(0) = \frac{\partial f}{\partial x_i}(a)$.
 Since $Df(a) = \left[\frac{\partial f}{\partial x_1}(a) \quad \frac{\partial f}{\partial x_2}(a) \quad \dots \quad \frac{\partial f}{\partial x_n}(a) \right]$, then $Df(a) = 0$.

Definition 3.2.3: Critical Point and Saddle Point

For $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, let $a \in X$.

If $Df(a) = 0$ or undefined, then a is a **critical point**.

Even if $Df(a) = 0$, $f(a)$ might not be a local extrema since it might be a local maximum along some paths and a local minimum along other paths. For this type of ambiguity of extrema, a is defined as a **saddle point**.

Theorem 3.2.4: Second Derivative Test

Let $f: X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ be C^2 .

If $a \in X$ is a critical point of f , then:

- (a) If $f_{xx}(a)f_{yy}(a) - [f_{xy}(a)]^2 > 0$ and $f_{xx}(a) > 0$, then a is a local minimum
- (b) If $f_{xx}(a)f_{yy}(a) - [f_{xy}(a)]^2 > 0$ and $f_{xx}(a) < 0$, then a is a local maximum
- (c) If $f_{xx}(a)f_{yy}(a) - [f_{xy}(a)]^2 < 0$, then a is saddle point

Definition 3.2.5: Compactness

Let set $X \subset \mathbb{R}^n$.

A point $x \in X$ is a **limit point** if for any $\delta > 0$:

The set of all p where $\|p - x\| < \delta$ contain a $p \in X$ where $p \neq x$.

X is closed if all limit point of X are in X .

X is **bounded** if there is a $M \in \mathbb{R}$ such that for all $x \in X$:

$$\|x\| < M$$

Then, X is **compact** if X is closed and bounded.

Theorem 3.2.6: Extreme Value Theorem

If $X \subset \mathbb{R}^n$ is compact and $f: X \rightarrow \mathbb{R}$ is continuous, then there are

$x_{\min}, x_{\max} \in X$ such that for all $x \in X$:

$$f(x_{\min}) \leq f(x) \leq f(x_{\max})$$

3.3 Lagrange Multipliers for Constrained Extrema

Theorem 3.3.1: Lagrange Multiplier: Optimization for one constraint

Let $f, g: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be C^1 and $g(x) = c$ for constant $c \in \mathbb{R}$.

If $f(x)$ has an extrema at x_0 where $g(x_0) = c$ and $\nabla g(x_0) \neq 0$, then there is a scalar $\lambda \in \mathbb{R}$ such that:

$$\nabla f(x_0) = \lambda \nabla g(x_0)$$

Theorem 3.3.2: Lagrange Multiplier: Optimization for multiple constraints

Let $f, g_1, \dots, g_k: X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be C^1 where $k < n$ and each $g_i(x) = c_i$ for constants $c_i \in \mathbb{R}$ for $i = \{1, \dots, k\}$.

If $f(x)$ has an extrema at x_0 where each $g_i(x_0) = c_i$ and $\nabla g_i(x_0) \neq 0$, then there are scalars $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that:

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \lambda_2 \nabla g_2(x_0) + \dots + \lambda_k \nabla g_k(x_0)$$

4 Integration

4.1 Double Integral

Definition 4.1.1: Riemann Sum in \mathbb{R}^2

For region $R = [a,b] \times [c,d]$, a **partition**, P , consist:

$\{x_0, \dots, x_n\} \in [a,b]$ such that each $x_i < x_{i+1}$ and $\Delta x_i = x_i - x_{i-1}$

$\{y_0, \dots, y_n\} \in [c,d]$ such that each $y_i < y_{i+1}$ and $\Delta y_i = y_i - y_{i-1}$

Let $\|P\| = \max(\Delta x_i, \Delta y_j)$ for $i,j = \{1, \dots, n\}$. Then for each $[x_{i-1}, x_i] \times [y_{j-1}, y_j]$, choose one point c_{ij} . Let **sample points**, C , consist of all c_{ij} .

Then the **Riemann sum** of $f(x,y)$ for partition P , sample points C :

$$R(f,P,C) = \sum_{i,j=1}^n f(c_{ij}) \Delta x_i \Delta y_j$$

Definition 4.1.2: Double Integral

$f: R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is **integrable** if:

$$\int \int_R f(x,y) \, dA = \int \int_R f(x,y) \, dx dy = \lim_{\|P\| \rightarrow 0} \sum_{i,j=1}^n f(c_{ij}) \Delta x_i \Delta y_j$$

exist. Then, $\int \int_R f(x,y) \, dA$ is the **double integral** of f on R .

Theorem 4.1.3: Continuous functions in \mathbb{R}^2 are integrable

If $f: R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous on region R , then $\int \int_R f(x,y) \, dA$ exists.

Theorem 4.1.4: Piecewise continuous functions in \mathbb{R}^2 are integrable

Set $X \subset \mathbb{R}^2$ has **measure zero** if for any $\epsilon > 0$ and every $x_i \in X$, there are $[a_i, b_i] \times [c_i, d_i]$ where $x_i \in [a_i, b_i] \times [c_i, d_i]$ such that:

$$\sum_{x_i \in X} (b_i - a_i) * (d_i - c_i) < \epsilon$$

If $f: R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is bounded and the set S of discontinuities of f on R has measure zero, then $\int \int_R f(x,y) \, dA$ exists.

Theorem 4.1.5: Fubini's Theorem in \mathbb{R}^2 : Double Integrals \Rightarrow Iterated Integrals

Let $f: R = [a,b] \times [c,d] \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ be bounded and the set S of discontinuities of f on R have measure zero. If every line parallel to $x=0$ or $y=0$ contains finitely many points on S , then:

$$\int \int_R f(x,y) \, dA = \int_c^d \int_a^b f(x,y) \, dx dy = \int_a^b \int_c^d f(x,y) \, dy dx$$

Theorem 4.1.6: Properties of the Integral in \mathbb{R}^2

(a) If f, g are integrable on R , then $f+g$ is integrable on R where:

$$\int \int_R f+g \, dA = \int \int_R f \, dA + \int \int_R g \, dA$$

(b) If f is integrable on R and scalar $c \in \mathbb{R}$, then cf is integrable on R where:

$$\int \int_R cf \, dA = c \int \int_R f \, dA$$

(c) If f, g are integrable on R where $f(x,y) \leq g(x,y)$ for all $(x,y) \in R$:

$$\int \int_R f \, dA \leq \int \int_R g \, dA$$

(d) If f is integrable on R , then $|f|$ is integrable on R where:

$$|\int \int_R f \, dA| \leq \int \int_R |f| \, dA$$

Definition 4.1.7: Elementary Region and the Extension of f in \mathbb{R}^2

Set $D \subset \mathbb{R}^2$ is an **elementary region** if:

(a) **Type 1**

For $x \in [a,b]$, there are continuous $y_1(x), y_2(x)$ where $y_1(x) \leq y \leq y_2(x)$.

$$D = \{ (x,y) \mid x \in [a,b], y \in [y_1(x), y_2(x)] \}$$

(b) **Type 2**

For $y \in [c,d]$, there are continuous $x_1(y), x_2(y)$ where $x_1(y) \leq x \leq x_2(y)$.

$$D = \{ (x,y) \mid x \in [x_1(y), x_2(y)], y \in [c,d] \}$$

(c) **Type 3**

Both Type 1 and Type 2

If $f: D \rightarrow \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous and D is an elementary region, then the extension of f , f^{ext} :

$$f^{ext}(x, y) = \begin{cases} f(x, y) & (x, y) \in D \\ 0 & (x, y) \notin D \end{cases}$$

Since f is continuous and thus, integrable and the discontinuities of f^{ext} exist only at the boundary of D which is a curve and thus, have measure zero, then f^{ext} is integrable.

Theorem 4.1.8: Double Integral over a General Region

If $f: D \rightarrow \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous and D is an elementary region, then for any region $R = [a,b] \times [c,d]$ such that $D \subset R$:

$$\int \int_D f(x,y) \, dA = \int \int_R f^{ext}(x, y) \, dA$$

Then if:

(a) D is a Type 1 elementary region

$$\int \int_D f(x,y) \, dA = \int_a^b \int_{y_1(x)}^{y_2(x)} f(x,y) \, dy \, dx$$

(b) D is a Type 2 elementary region

$$\int \int_D f(x,y) \, dA = \int_c^d \int_{x_1(y)}^{x_2(y)} f(x,y) \, dx \, dy$$

Proof

Suppose D is a Type 1 elementary region.

Let region $R = [a,b] \times [c,d]$ be such that $c \leq y_1(x) \leq y \leq y_2(x) \leq d$.

Since $f^{ext}(x, y) = 0$ for any $(x,y) \notin D$, then by **theorem 4.1.5**:

$$\int \int_D f(x,y) \, dA = \int \int_R f^{ext}(x, y) \, dA = \int_a^b \int_c^d f^{ext}(x, y) \, dy \, dx = \int_a^b \int_{y_1(x)}^{y_2(x)} f(x,y) \, dy \, dx$$

Suppose D is a Type 2 elementary region.

Let region $R = [a,b] \times [c,d]$ be such that $a \leq x_1(y) \leq x \leq x_2(y) \leq b$.

Since $f^{ext}(x, y) = 0$ for any $(x,y) \notin D$, then by **theorem 4.1.5**:

$$\int \int_D f(x,y) \, dA = \int \int_R f^{ext}(x, y) \, dA = \int_c^d \int_a^b f^{ext}(x, y) \, dx \, dy = \int_c^d \int_{x_1(y)}^{x_2(y)} f(x,y) \, dx \, dy$$

Definition 4.1.9: Area

The **area** of a region $R \subset \mathbb{R}^2$:

$$A = \int \int_R 1 \, dA$$

4.2 Triple Integral

Definition 4.2.1: Riemann Sum in \mathbb{R}^3

For region $R = [a,b] \times [c,d] \times [e,f]$, a **partition**, P , consist:

$$\{x_0, \dots, x_n\} \in [a,b] \text{ such that each } x_i < x_{i+1} \text{ and } \Delta x_i = x_i - x_{i-1}$$

$$\{y_0, \dots, y_n\} \in [c,d] \text{ such that each } y_i < y_{i+1} \text{ and } \Delta y_i = y_i - y_{i-1}$$

$$\{z_0, \dots, z_n\} \in [e,f] \text{ such that each } z_i < z_{i+1} \text{ and } \Delta z_i = z_i - z_{i-1}$$

Let $\|P\| = \max(\Delta x_i, \Delta y_j, \Delta z_k)$ for $i,j,k = \{1, \dots, n\}$. Then for each $[x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$, choose one point c_{ijk} . Let **sample points**, C , consist of all c_{ijk} . Then the **Riemann sum** of $f(x,y,z)$ for partition P , sample points C :

$$R(f,P,C) = \sum_{i,j,k=1}^n f(c_{ijk}) \Delta x_i \Delta y_j \Delta z_k$$

Definition 4.2.2: Triple Integral

$f: R \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ is **integrable** if:

$$\begin{aligned} \iiint_R f(x,y,z) dV &= \iiint_R f(x,y,z) dx dy dz \\ &= \lim_{\|P\| \rightarrow 0} \sum_{i,j,k=1}^n f(c_{ijk}) \Delta x_i \Delta y_j \Delta z_k \end{aligned}$$

exist. Then, $\iiint_R f(x,y,z) dV$ is the **triple integral** of f on R .

Theorem 4.2.3: Continuous functions in \mathbb{R}^3 are integrable

If $f: R \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ is continuous on region R , then $\iiint_R f(x,y,z) dV$ exists.

Theorem 4.2.4: Piecewise continuous functions in \mathbb{R}^3 are integrable

Set $X \subset \mathbb{R}^3$ has measure zero if for any $\epsilon > 0$ and every $x_i \in X$, there are $[a_i, b_i] \times [c_i, d_i] \times [e_i, f_i]$ where $x_i \in [a_i, b_i] \times [c_i, d_i] \times [e_i, f_i]$ such that:

$$\sum_{x_i \in X} (b_i - a_i) * (d_i - c_i) * (f_i - e_i) < \epsilon$$

If $f: R \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ is bounded and the set S of discontinuities of f on R has measure zero, then $\iiint_R f(x,y,z) dV$ exists.

Theorem 4.2.5: Fubini's Theorem in \mathbb{R}^3 : Triple Integrals \Rightarrow Iterated Integrals

Let $f: R = [a,b] \times [c,d] \times [e,f] \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ be bounded and the set S of discontinuities of f on R have measure zero. If every line parallel to $x=0$, $y=0$, or $z=0$ contains finitely many points on S , then:

$$\begin{aligned} \iiint_R f(x,y,z) dV &= \int_e^f \int_c^d \int_a^b f(x,y,z) dx dy dz = \int_c^d \int_e^f \int_a^b f(x,y,z) dx dz dy \\ &= \int_e^f \int_a^b \int_c^d f(x,y,z) dy dx dz = \int_a^b \int_e^f \int_c^d f(x,y,z) dy dz dx \\ &= \int_c^d \int_a^b \int_e^f f(x,y,z) dz dx dy = \int_a^b \int_c^d \int_e^f f(x,y,z) dz dy dx \end{aligned}$$

Theorem 4.2.6: Properties of the Integral in \mathbb{R}^3

(a) If f, g are integrable on R , then $f+g$ is integrable on R where:

$$\iiint_R f+g dV = \iiint_R f dV + \iiint_R g dV$$

(b) If f is integrable on R and scalar $c \in \mathbb{R}$, then cf is integrable on R where:

$$\iiint_R cf dV = c \iiint_R f dV$$

(c) If f, g are integrable on R where $f(x,y) \leq g(x,y)$ for all $(x,y) \in R$:

$$\iiint_R f dV \leq \iiint_R g dV$$

(d) If f is integrable on R , then $|f|$ is integrable on R where:

$$|\iiint_R f dV| \leq \iiint_R |f| dV$$

Definition 4.2.7: Elementary Region and the Extension of f in \mathbb{R}^3

Set $D \subset \mathbb{R}^3$ is an elementary region if:

(a) Type 1

For $x \in [a, b]$, there are continuous $y_1(x), y_2(x)$ where $y_1(x) \leq y \leq y_2(x)$ and continuous $z_1(x, y), z_2(x, y)$ where $z_1(x, y) \leq z \leq z_2(x, y)$.

$$D = \{ (x, y, z) \mid x \in [a, b], y \in [y_1(x), y_2(x)], z \in [z_1(x, y), z_2(x, y)] \}$$

OR

For $x \in [a, b]$, there are continuous $z_1(x), z_2(x)$ where $z_1(x) \leq z \leq z_2(x)$ and continuous $y_1(x, z), y_2(x, z)$ where $y_1(x, z) \leq y \leq y_2(x, z)$.

$$D = \{ (x, y, z) \mid x \in [a, b], y \in [y_1(x, z), y_2(x, z)], z \in [z_1(x), z_2(x)] \}$$

(b) Type 2

For $y \in [c, d]$, there are continuous $x_1(y), x_2(y)$ where $x_1(y) \leq x \leq x_2(y)$ and continuous $z_1(x, y), z_2(x, y)$ where $z_1(x, y) \leq z \leq z_2(x, y)$.

$$D = \{ (x, y, z) \mid x \in [x_1(y), x_2(y)], y \in [c, d], z \in [z_1(x, y), z_2(x, y)] \}$$

OR

For $y \in [c, d]$, there are continuous $z_1(y), z_2(y)$ where $z_1(y) \leq z \leq z_2(y)$ and continuous $x_1(y, z), x_2(y, z)$ where $x_1(y, z) \leq x \leq x_2(y, z)$.

$$D = \{ (x, y, z) \mid x \in [x_1(y, z), x_2(y, z)], y \in [c, d], z \in [z_1(y), z_2(y)] \}$$

(c) Type 3

For $z \in [e, f]$, there are continuous $x_1(z), x_2(z)$ where $x_1(z) \leq x \leq x_2(z)$ and continuous $y_1(x, z), y_2(x, z)$ where $y_1(x, z) \leq y \leq y_2(x, z)$.

$$D = \{ (x, y, z) \mid x \in [x_1(z), x_2(z)], y \in [y_1(x, z), y_2(x, z)], z \in [e, f] \}$$

OR

For $z \in [e, f]$, there are continuous $y_1(z), y_2(z)$ where $y_1(z) \leq y \leq y_2(z)$ and continuous $x_1(y, z), x_2(y, z)$ where $x_1(y, z) \leq x \leq x_2(y, z)$.

$$D = \{ (x, y, z) \mid x \in [x_1(y, z), x_2(y, z)], y \in [y_1(z), y_2(z)], z \in [e, f] \}$$

(d) Type 4

All Type 1, 2, and 3

If $f: D \rightarrow \mathbb{R}^3 \rightarrow \mathbb{R}$ is continuous and D is an elementary region, then the extension of f , f^{ext} :

$$f^{ext}(x, y, z) = \begin{cases} f(x, y, z) & (x, y, z) \in D \\ 0 & (x, y, z) \notin D \end{cases}$$

Since f is continuous and thus, integrable and the discontinuities of f^{ext} exist only at the boundary of D which is a 2d surface and thus, have measure zero, then f^{ext} is integrable.

Theorem 4.2.8: Triple Integral over a General Region

If $f: D \rightarrow \mathbb{R}^3 \rightarrow \mathbb{R}$ is continuous and D is an elementary region, then for any region $R = [a,b] \times [c,d] \times [e,f]$ such that $D \subset R$:

$$\iiint_D f(x,y,z) \, dV = \iiint_R f^{ext}(x,y,z) \, dV$$

Then if:

(a) D is a Type 1 elementary region

$$\begin{aligned} \iiint_D f(x,y,z) \, dV &= \int_a^b \int_{y_1(x)}^{y_2(x)} \int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z) \, dz dy dx \\ &= \int_a^b \int_{z_1(x)}^{z_2(x)} \int_{y_1(x,z)}^{y_2(x,z)} f(x,y,z) \, dy dz dx \end{aligned}$$

(b) D is a Type 2 elementary region

$$\begin{aligned} \iiint_D f(x,y,z) \, dV &= \int_c^d \int_{x_1(y)}^{x_2(y)} \int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z) \, dz dx dy \\ &= \int_c^d \int_{z_1(y)}^{z_2(y)} \int_{x_1(y,z)}^{x_2(y,z)} f(x,y,z) \, dx dz dy \end{aligned}$$

(c) D is a Type 3 elementary region

$$\begin{aligned} \iiint_D f(x,y,z) \, dV &= \int_e^f \int_{x_1(z)}^{x_2(z)} \int_{y_1(x,z)}^{y_2(x,z)} f(x,y,z) \, dy dx dz \\ &= \int_e^f \int_{y_1(z)}^{y_2(z)} \int_{x_1(y,z)}^{x_2(y,z)} f(x,y,z) \, dx dy dz \end{aligned}$$

Proof

Suppose D is a Type 1 elementary region.

Let region $R = [a,b] \times [c,d] \times [e,f]$ be such that $c \leq y_1(x) \leq y \leq y_2(x) \leq d$ and $e \leq z_1(x,y) \leq z \leq z_2(x,y) \leq f$.

Since $f^{ext}(x,y,z) = 0$ for any $(x,y,z) \notin D$, then by **theorem 4.1.5**:

$$\begin{aligned} \iiint_D f(x,y,z) \, dV &= \iiint_R f^{ext}(x,y,z) \, dV = \int_a^b \int_c^d \int_e^f f^{ext}(x,y,z) \, dz dy dx \\ &= \int_a^b \int_c^d \int_{z_1(x,y)}^{z_2(x,y)} f^{ext}(x,y,z) \, dz dy dx \\ &= \int_a^b \int_{y_1(x)}^{y_2(x)} \int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z) \, dz dy dx \end{aligned}$$

Proof is analogous for the other five cases.

Definition 4.2.9: Volume

The **volume** of a region $R \subset \mathbb{R}^3$:

$$V = \iiint_R 1 \, dV$$

4.3 Change of Variables

Theorem 4.3.1: Linear Transformations Scaling Factor

Let $A \in M_{2 \times 2}(\mathbb{R})$ where $\det(A) \neq 0$ such that $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$:

$$(x,y) = T(u,v) = A \begin{bmatrix} u \\ v \end{bmatrix}$$

Transformation T is 1-1 and onto where if D is a parallelogram in the uv -plane, then $T(D)$ is a parallelogram in the xy -plane where:

$$\text{Vol}_2(T(D)) = |\det(A)| * \text{Vol}_2(D)$$

Proof

The proof that T is 1-1 and onto uses linear algebra, but it will be made simpler here. Note however if linear algebra is used in full effect, the given $\det(A) \neq 0$ satisfies the condition for 1-1 and onto.

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Suppose $(a,c) = k(b,d)$ for some scalar $k \in \mathbb{R}$.

$$A \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} au + bv \\ cu + dv \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} u + \begin{bmatrix} b \\ d \end{bmatrix} v = k \begin{bmatrix} b \\ d \end{bmatrix} u + \begin{bmatrix} b \\ d \end{bmatrix} v = \begin{bmatrix} b \\ d \end{bmatrix} (ku + v)$$

Let $(u,v) = (1,-k)$ and $(u,v) = (2,-2k)$. In both cases, $A \begin{bmatrix} u \\ v \end{bmatrix} = 0$ so T is not 1-1 if

$(a,c) = k(b,d)$. But, if there is a k such that $(a,c) = k(b,d)$, then $\det(A) = ad - bc = (kb)d - b(kd) = 0$ which contradicts that $\det(A) \neq 0$. Also, note if there is a k such that $(a,c) = k(b,d)$, then $A(u,v)$ consist of only multiples of (b,d) so T is not onto since vectors that are not multiples of (b,d) such as $(b,d+1)$ cannot be reached. Now suppose there isn't a k such that $(a,c) = k(b,d)$. Then $A(u,v)$ cannot have two different $(u_1, v_1), (u_2, v_2)$ such that $A(u_1, v_1) = A(u_2, v_2)$ since

$$A(u_1, v_1) - A(u_2, v_2) = (a,c)(u_1 - u_2) + (b,d)(v_1 - v_2) = 0$$

so $(a,c) = -\frac{v_1 - v_2}{u_1 - u_2}(b,d)$ and thus, there is a $k = -\frac{v_1 - v_2}{u_1 - u_2}$ such that $(a,c) = k(b,d)$ which is a contradiction. Thus, if there isn't a k such that $(a,c) = k(b,d)$, then for two different (u_1, v_1) and (u_2, v_2) , then $A(u_1, v_1) \neq A(u_2, v_2)$ so T is 1-1. Also, note since there isn't a k such that $(a,c) = k(b,d)$, then any (x,y) can be reached by solving the system of equations from $A(u,v) = (x,y)$ for u,v and thus, T is onto.

If there isn't a k such that $(a,b) = k(c,d)$, then $\det(A) \neq 0$ since if $\det(A) = ad - bc = 0$, then $ad = bc$ so $\frac{a}{b} = \frac{c}{d}$. Note $\frac{a}{b} = \frac{kc}{kd} = \frac{c}{d}$ if there is a k such that $(a,b) = k(c,d)$ which is a contradiction. Thus, by the restriction $\det(A) \neq 0$, then T is 1-1 and onto.

For any two $(u_1, v_1), (u_2, v_2)$ where $(u_1, v_1) \neq k(u_1, v_1)$ for any $k \in \mathbb{R}$, a parallelogram with sides $(u_1, v_1), (u_2, v_2)$ can be formed in uv -plane. Then T maps $(u_1, v_1), (u_2, v_2)$ into two different $A(u_1, v_1), A(u_2, v_2)$ since T is 1-1 so $A(u_1, v_1), A(u_2, v_2)$ are sides of a parallelogram in the xy -plane. Thus:

$$\begin{aligned} \text{Vol}_2(T(D)) &= \|A(u_1, v_1) \times A(u_2, v_2)\| = \|[(a, c)u_1 + (b, d)v_1] \times [(a, c)u_2 + (b, d)v_2]\| \\ &= \|(a, c)u_1 \times (a, c)u_2 + (a, c)u_1 \times (b, d)v_2 + (b, d)v_1 \times (a, c)u_2 + (b, d)v_1 \times (b, d)v_2\| \\ &= \| [u_1v_2 - v_1u_2][(a, c) \times (b, d)] \| = \|(a, c) \times (b, d)\| [u_1v_2 - v_1u_2] \\ &= |ad - bc| \|(u_1, v_1) \times (u_2, v_2)\| = |\det(A)| * \text{Vol}_2(D) \end{aligned}$$

Theorem 4.3.2: Integral: Change of Variables

The **Jacobian** of $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ where $T(u,v) = (x,y)$:

$$\frac{\partial(x,y)}{\partial(u,v)} = \det(DT(u,v)) = \det\left(\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}\right)$$

Let elementary region D be in uv -plane. If transformation T is C^1 (i.e. partial derivatives above are continuous), then for integrable $f: T(D) \rightarrow \mathbb{R}$ where $x = x(u,v)$ and $y = y(u,v)$:

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

The **Jacobian** of $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ where $T(u,v,w) = (x,y,z)$:

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \det(DT(u,v)) = \det\left(\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix}\right)$$

Let elementary region D be in uvw -space. If transformation T is C^1 , then for integrable $f: T(D) \rightarrow \mathbb{R}$ where $x = x(u,v,w)$, $y = y(u,v,w)$, and $z = z(u,v,w)$:

$$\begin{aligned} \int \int \int_{T(D)} f(x,y,z) \, dV(x,y,z) \\ = \int \int \int_D f(x(u,v,w), y(u,v,w), z(u,v,w)) \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| \, dV(u,v,w) \end{aligned}$$

Proof

Take the case for $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$. The case for \mathbb{R}^3 is analogous.

Since T is C^1 , then $DT(u,v)$ exist and since $T(u,v) = A(u,v) = (x(u,v), y(u,v))$, then:

$$DT(u,v) = A' = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \left| \frac{\partial(x,y)}{\partial(u,v)} \right|$$

Thus, by **theorem 4.3.1**:

$$\begin{aligned} \int \int_{T(D)} f(x,y) \, dA(x,y) \\ &= \lim_{\|P\| \rightarrow 0} \sum_{i,j=1}^n f(x_{ij}, y_{ij}) \Delta x_i \Delta y_j \\ &= \lim_{\|P\| \rightarrow 0} \sum_{i,j=1}^n f(x_{ij}, y_{ij}) \text{Vol}_2([\Delta x_i, \Delta y_j]) \\ &= \lim_{\|P\| \rightarrow 0} \sum_{i,j=1}^n f(x_{ij}(u,v), y_{ij}(u,v)) |DT(u,v)| \text{Vol}_2([\Delta u_i, \Delta v_j]) \\ &= \lim_{\|P\| \rightarrow 0} \sum_{i,j=1}^n f(x_{ij}(u,v), y_{ij}(u,v)) |DT(u,v)| \Delta u_i \Delta v_j \\ &= \int \int_D f(x(u,v), y(u,v)) \left| \frac{\partial(x,y)}{\partial(u,v)} \right| \, dA(u,v) \end{aligned}$$

5 Line Integrals

5.1 Parametrized Curves

Definition 5.1.1: Path

A **path** in \mathbb{R}^n is a continuous $C(t): [a, b] \rightarrow \mathbb{R}^n$.

If $C(t)$ is twice-differentiable, then the **velocity** of $C(t)$ at $t_0 \in [a, b]$:

$$v(t_0) = x'(t_0)$$

Also, the acceleration of $C(t)$ at $t_0 \in [a, b]$:

$$a(t_0) = x''(t_0)$$

Definition 5.1.2: Arclength

The length of a C^1 path $C(t): [a, b] \rightarrow \mathbb{R}^n$:

$$L = \int_a^b \|C'(t)\| dt$$

Proof

Choose $\{x_1, \dots, x_n\} \in [a, b]$ such that each $x_i < x_{i+1}$. Then the length of $C(t)$:

$$L = \lim_{n \rightarrow \infty} \sum_{i=1}^n \|C(x_i) - C(x_{i-1})\|$$

Let each $C(x_i) = (C_1(x_i), \dots, C_n(x_i))$. Thus:

$$\|C(x_i) - C(x_{i-1})\| = \sqrt{(C_1(x_i) - C_1(x_{i-1}))^2 + \dots + (C_n(x_i) - C_n(x_{i-1}))^2}$$

Since $C(t)$ is C^1 , by the Mean Value Theorem, there is a $t_{i_k} \in [x_{i+1}, x_i]$ such that:

$$C_k(x_i) - C_k(x_{i-1}) = (x_i - x_{i-1})C'_k(t_{i_k})$$

Thus:

$$\|C(x_i) - C(x_{i-1})\| = \sqrt{(x_i - x_{i-1})^2 [C'_1(t_{i_1})]^2 + \dots + (x_i - x_{i-1})^2 [C'_n(t_{i_n})]^2}$$

$$L = \lim_{n \rightarrow \infty} \sum_{i=1}^n \sqrt{[C'_1(t_{i_1})]^2 + \dots + [C'_n(t_{i_n})]^2} (x_i - x_{i-1})$$

$$= \int_a^b \sqrt{[C'_1(t)]^2 + \dots + [C'_n(t)]^2} dt = \int_a^b \|C'(t)\| dt$$

5.2 Vector Fields

Definition 5.2.1: Vector Field and Flow Lines

A **vector field** on \mathbb{R}^n is $F: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$

A **flow line** of vector field F is a differentiable path $C(t): [a, b] \rightarrow \mathbb{R}^n$ such that:

$$C'(t) = F(C(t))$$

Definition 5.2.2: Del Operator

The **Del Operator** on \mathbb{R}^n :

$$\nabla = \sum_{i=1}^n \frac{\partial}{\partial x_i} e_i = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right)$$

Definition 5.2.3: Divergence

For differentiable vector field $F: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, let $F = (F_1, \dots, F_n)$.

Then the **divergence**, $\text{div}: \mathbb{R}^n \rightarrow \mathbb{R}$, of F :

$$\text{div}(F) = \nabla \cdot F = \frac{\partial F_1}{\partial x_1} + \dots + \frac{\partial F_n}{\partial x_n}$$

If $\text{div}(F) = 0$ everywhere, then F is incompressible.

Definition 5.2.4: Curl

For differentiable vector field $F: X \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$, let $F = (F_1, F_2, F_3)$.

Then the **curl**, $\text{curl}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, of F :

$$\text{curl}(F) = \nabla \times F = \begin{bmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{bmatrix} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right)$$

If $\text{curl}(F) = 0$ everywhere, then F is irrotational.

Theorem 5.2.5: Vector fields from Gradients are irrotational

If $f: X \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ is C^2 , then $\text{curl}(\nabla f) = 0$

Proof

Since $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$, then by **theorem 2.3.9**:

$$\text{curl}(\nabla f) = \begin{bmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} = \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y}, \frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z}, \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) = 0$$

Theorem 5.2.6: The curl is incompressible

If $F: X \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is C^2 , then $\text{div}(\text{curl}(F)) = 0$

Proof

Since $\text{curl}(F) = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right)$, then by **theorem 2.3.9**:

$$\text{div}(\text{curl}(F)) = \frac{\partial}{\partial x} \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) = 0$$
5.3 Line Integrals**Theorem 5.3.1: Scalar Line Integral**

Let C^1 path $C(t): [a, b] \rightarrow \mathbb{R}^n$ and $f: C([a, b]) \subset X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be continuous.

Then the **scalar line integral** of f along $C(t)$:

$$\int_{C(t)} f \, ds = \int_a^b f(C(t)) \|C'(t)\| \, dt$$

Proof

Let partition $P = \{t_0, \dots, t_n\} \subset [a, b]$ with sample points $C = \{t_1^*, \dots, t_n^*\}$, the Riemann sum of f along $C(t)$ with partition P and sample points C :

$$\sum_{i=1}^n f(C(t_i^*)) \Delta s_i \quad \text{where } \Delta s_i = \int_{t_{i-1}}^{t_i} \|C'(t)\| \, dt$$

Since $C(t)$ is C^1 so Δs_i is differentiable, then by the Mean Value Theorem, there is a $t_i^{**} \in [t_{i-1}, t_i]$:

$$\Delta s_i = \int_{t_{i-1}}^{t_i} \|C'(t)\| \, dt = (t_i - t_{i-1}) \|C'(t_i^{**})\| = \|C'(t_i^{**})\| \Delta t_i$$

Thus:

$$\lim_{\|P\| \rightarrow 0} \sum_{i=1}^n f(C(t_i^*)) \Delta s_i = \lim_{\|P\| \rightarrow 0} \sum_{i=1}^n f(C(t_i^*)) \|C'(t_i^{**})\| \Delta t_i = \int_a^b f(C(t)) \|C'(t)\| \, dt$$

Theorem 5.3.2: Vector Line Integral

Let C^1 path $C(t): [a,b] \rightarrow \mathbb{R}^n$ and $F: C([a,b]) \subset X \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be continuous. Then the **vector line integral** of F along $C(t)$:

$$\int_{C(t)} F \cdot ds = \int_a^b F(C(t)) \cdot C'(t) dt$$

If $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, let $F(x,y,z) = (M(x,y,z), N(x,y,z), P(x,y,z))$ and $C(t) = (x(t), y(t), z(t))$, then:

$$\int_{C(t)} F \cdot ds = \int_a^b M(x,y,z) dx + N(x,y,z) dy + P(x,y,z) dz$$

The vector line integral is also called the **circulation of F along $C(t)$** .

5.4 Green's Theorem**Theorem 5.4.1: Green's Theorem**

Let D be a closed, bounded region in \mathbb{R}^2 where the boundary of D , C consist of finitely many simple, closed, piecewise C^1 curves C_i . For each C_i , let parametrization $C_i(t)$ be such that as t increases, D is at the left of $C_i(t)$.

Then for C^1 vector field $F(x,y) = (M(x,y), N(x,y))$:

$$\oint_C F \cdot ds = \int \int_D -\frac{\partial M}{\partial y} + \frac{\partial N}{\partial x} dA = \int \int_D (\nabla \times F) \cdot e_3 dA$$

Theorem 5.4.2: Gauss's Theorem: Divergence Theorem in \mathbb{R}^2

If Green's Theorem applies for region D and n is a outward normal unit vector to D , then for C^1 vector field $F(x,y) = (M(x,y), N(x,y))$:

$$\oint_C F \cdot n ds = \int \int_D \nabla \cdot F dA$$

5.5 Conservative Vector Fields**Definition 5.5.1: Path Independence**

A continuous vector field F has **path independent** line integrals if for any two simple, piecewise C^1 , oriented curves C_1, C_2 on $[a,b]$ where $C_1(a) = C_2(a)$ and $C_1(b) = C_2(b)$:

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

Theorem 5.5.2: Path Independence $\Leftrightarrow \int_{C_1} F \cdot ds = 0$

For continuous vector field F , F has path independent line integrals if and only if $\int_C F \cdot ds = 0$ for all simple, piecewise C^1 , closed curves.

Definition 5.5.3: Potential of F

Let continuous vector field $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ have a C^1 function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ such that $\nabla f = F$. Then f is the **potential** of F and F is a **conservative vector field**.

Theorem 5.5.4: Conservative Vector Fields \Leftrightarrow Path independence

Let continuous vector field $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Then $F = \nabla f$ if and only if F has path independent line integrals.

If $F = \nabla f$ for some C^1 function f , then for any piecewise C^1 , oriented curve C with initial point A and terminal point B :

$$\int_C F \cdot ds = f(B) - f(A)$$

6 Surface Integrals

6.1 Parametrized Surfaces

Definition 6.1.1: Parametrized Surfaces

A **parametrized surface** is a continuous function $X: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$.

Let surface S be parametrized by differentiable $X(s,t) = (x(s,t), y(s,t), z(s,t))$.

Then the tangent vectors:

$$T_s(s,t) = \frac{\partial X}{\partial s}(s,t) = \left(\frac{\partial x}{\partial s}(s,t), \frac{\partial y}{\partial s}(s,t), \frac{\partial z}{\partial s}(s,t) \right)$$

$$T_t(s,t) = \frac{\partial X}{\partial t}(s,t) = \left(\frac{\partial x}{\partial t}(s,t), \frac{\partial y}{\partial t}(s,t), \frac{\partial z}{\partial t}(s,t) \right)$$

S is **smooth** if X is C^1 and the **standard normal vector** of X :

$$N_X(s,t) = T_s(s,t) \times T_t(s,t) \neq 0.$$

S is **piecewise smooth** if $X = \{X_1, \dots, X_m\}$ where each parametrized surface X_i is C^1 and $N_{X_i}(s,t) \neq 0$

Theorem 6.1.2: Surface Area

Let surface S be parametrized by C^1 function $X: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$.

Then the surface area of S :

$$\iint_D \|N_X(s,t)\| \, dA = \iint_D \sqrt{\left(\frac{\partial(x,y)}{\partial(s,t)}\right)^2 + \left(\frac{\partial(x,z)}{\partial(s,t)}\right)^2 + \left(\frac{\partial(y,z)}{\partial(s,t)}\right)^2} \, dA$$

If the surface S is $z = f(x,y)$, then the surface area of S :

$$\iint_D \sqrt{f_x^2 + f_y^2 + 1} \, dA$$

6.2 Surface Integrals

Theorem 6.2.1: Scalar Surface Integrals

Let surface S be parametrized by smooth $X(s,t): D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ and $f(x,y,z): S \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ be continuous.

Then the **scalar surface integral** of f along $X(s,t)$:

$$\iint_{X(s,t)} f \, dS = \iint_D f(X(s,t)) \|N_X(s,t)\| \, dA$$

Theorem 6.2.2: Vector Surface Integrals

Let surface S be parametrized by smooth $X(s,t): D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ and vector field $F(x,y,z): S \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be continuous.

Then the **vector surface integral** of F along $X(s,t)$:

$$\iint_{X(s,t)} F \cdot dS = \iint_D F(X(s,t)) \cdot N_X(s,t) \, dA$$

The vector surface integral is also called the **flux of F across $S = X(s,t)$** .

6.3 Stokes's & Gauss's Theorems

Definition 6.3.1: Induced Orientation

Let S be a bounded, piecewise smooth surface in \mathbb{R}^3 with boundary defined by a simple, closed curve C . Let unit normal vector \mathbf{n} be orthogonal to S at a point so surface S has an orientation of \mathbf{n} .

Then for points on S , let there be a spinning disc with a direction such that the curl is in the same direction as \mathbf{n} at such points.

Then C has an **induced orientation** from S if the orientation of curve C is in the same direction as the spinning discs of points near C .

Theorem 6.3.2: Stokes's Theorem

Let S be a bounded, piecewise smooth, oriented surface in \mathbb{R}^3 with boundary defined by finitely many simple, closed, piecewise C^1 curves $C = \{C_1, \dots, C_m\}$ which have an induced orientation from S .

Then for C^1 vector field $F(x, y, z): S \subset \mathbb{R} \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$\iint_S \nabla \times F \cdot d\mathbf{S} = \oint_C F \cdot d\mathbf{s}$$

Theorem 6.3.3: Gauss's Theorem: Divergence Theorem in \mathbb{R}^3

Let D be a bounded region in \mathbb{R}^3 with boundary defined by finitely many closed, piecewise smooth surfaces $S = \{S_1, \dots, S_m\}$ oriented with unit normal vectors that points away from D .

Then for C^1 vector field $F(x, y, z): D \subset \mathbb{R} \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$\oiint_S F \cdot d\mathbf{S} = \iiint_D \nabla \cdot F \, dV$$

References

- [1] Susan Jane Colley, *Vector Calculus (4th Edition)*, ISBN-13: 978-0321780652