Mathematical Methods for Bioengineering

Biomedic Engineering Degree

Javier Quintero

Slides – Open Educational Material Rey Juan Carlos University
Avaible at Institutional Open Archive Rey Juan Carlos University (BURJC Digital)
`https://burjcdigital.urjc.es/home`

2024 – 2025 v1.02

Contents

§I. Geometry in the Euclidean Space

- Vectors in Several Dimensions
- Equations for Planes. Distances. Normal Vectors
- New Coordinate Systems
- Neighbourhoods. Open and Closed Sets.

We use the following notation for vectors

1. The real number line is denoted as $\mathbb{R}$
2. The set of all ordered pairs (x, y) of real numbers is denoted as $\mathbb{R}^2$ and represents the points in the plane.
3. The set of all ordered pairs (x, y, z) of real numbers is denoted as $\mathbb{R}^3$ and represents the points in the three-dimensional space.
4. The set of all ordered pairs $(x_1, x_2, \ldots, x_n)$ of real numbers is denoted as $\mathbb{R}^n$ and represents the points in the n-dimensional space.

Two vectors are equal if and only if their respective components are equal. In the text we reserve the boldfont to represent vectors, for example $\mathbf{u} \in \mathbb{R}^n$ is a vector and $u \in \mathbb{R}$ is a scalar.

Definition (Addition and scalar multiplication)

We define addition between two vectors $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, u_n)$ and multiplication by a scalar $k \in \mathbb{R}$ as

1. $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n)$
2. $\alpha\mathbf{u} = (\alpha u_1, \alpha u_2, \dots, \alpha u_n)$

These operations are performed component-wise.

Example

- $(0,1,3) + (7,-2,10) = (7,-1,13)$ in $\mathbb{R}^3$
- $(1,1) + (\pi, \sqrt{2}) = (1+\pi, 1+\sqrt{2})$ in $\mathbb{R}^2$

  If $\mathbf{a} = (2,0,\sqrt{2})$ and $k = 7$ then $k\mathbf{a} = (14,0,7\sqrt{2})$.

Proposition (Properties of Vector Addition)

1. **Commutativity**: $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ for all $\mathbf{a},\mathbf{b}\in\mathbb{R}^n$
2. **Associativity**: $\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$ for all $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{R}^n$
3. **Zero Vector** or neutral element: a special vector $\mathbf{0}=(0,0,0)$ with the property that $\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}$ for all $\mathbf{a}\in\mathbb{R}^3$. In $\mathbb{R}^2$ we have $\mathbf{0}=(0,0)$

Proposition (Properties of Scalar Multiplication)

For all vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$ and scalars k and l in $\mathbb{R}$, we have

1. $(k+l)\mathbf{a}=k\mathbf{a}+l\mathbf{a}$ (**distributivity**)
2. $k(\mathbf{a}+\mathbf{b})=k\mathbf{a}+k\mathbf{b}$ (**distributivity**)
3. $k(l\mathbf{a})=(kl)\mathbf{a}=l(k\mathbf{a})$

Definition (Canonical base)

The n canonical vectors e_k for $k \in \{1, \ldots, n\}$ are defined as

$$e_k = (0, \ldots, 1, \ldots, 0)$$

where the non-zero entry is in the kth position

Example

The canonical vectors in $\mathbb{R}^3$ are

$$\mathbf{i} = (1,0,0) \qquad \mathbf{j} = (0,1,0) \qquad \mathbf{k} = (0,0,1)$$

When working on $\mathbb{R}^2$, we only consider $\mathbf{i} = (1,0)$ and $\mathbf{j} = (0,1)$.

Note that we can use the canonical vectors to enable an alternative notation, for example, the vector (a,b,c) can be written as $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$.

Definition (Dot Product)

For two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ we define the dot product (or inner product) as

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

Dot product transforms two vectors into a scalar.

Definition (Norm or length of a vector)

The Euclidean norm of a vector $\mathbf{v}$ is defined as

$$\|\mathbf{v} \cdot \mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

Definition (Angle between vectors)

The angle θ between two vectors $\mathbf{v}$ and $\mathbf{w}$ is defined as

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{w}}{\|\mathbf{u}\| \|\mathbf{w}\|}$$

Proposition (Properties of Dot Products)

If $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are any vectors in $\mathbb{R}^n$, and $k \in \mathbb{R}$ is any scalar:

- $\mathbf{a} \cdot \mathbf{a} \geq 0$, and $\mathbf{a} \cdot \mathbf{a} = 0$ if and only if $\mathbf{a} = \mathbf{0}$.
- $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ (**commutative** property)
- $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$ (**distributive** property)
- $(k\mathbf{a}) \cdot \mathbf{b} = k(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (k\mathbf{b})$

Theorem (Cauchy-Schwarz inequality)

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, *then*

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$$

Corollary (Triangle inequality)

Under the previous conditions

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Definition (Distance between two points)

The Euclidean distance between two points $\mathbf{P} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{Q} = (y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$ is given by:

$$d(\mathbf{P}, \mathbf{Q}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = ||\mathbf{P} - \mathbf{Q}||$$

This is the generalization of the Pythagorean theorem to n-dimensional space.

Definition (Orthogonal projection)

The Orthogonal Projection of the vector $\mathbf{v}$ on the vector $\mathbf{a}$ is the vector

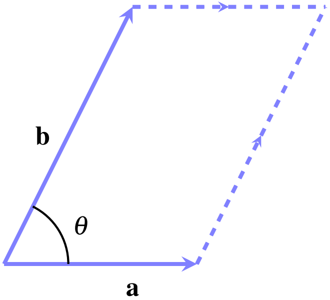
$$\text{proj}_a \mathbf{v} = \frac{\mathbf{a} \cdot \mathbf{v}}{||\mathbf{a}||^2} \mathbf{a}$$

Definition (Cross Product)

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. The cross product or vector product of **a** and **b** is denoted as $a \times b$ and defined as:

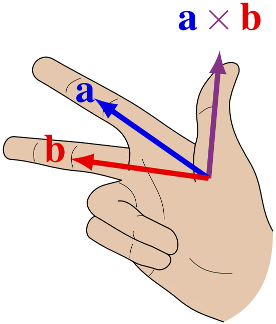
Direction: a vector orthogonal to **a** and **b** as the same time following the right-hand rule

Magnitude: $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\| \sin\theta$, where θ is the angle between **a** and **b**. This magnitude is the area of the parallelogram spanned by **a** and **b**.



Remark

The direction of the cross product is taken such that the system $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a} \times \mathbf{b}$ is a right-handed set of vectors.



Definition (Parametric Equation of a Line)

The vector parametric equation for the line through the point $P_0 = (b_1, b_2, b_3)$, whose position vector is

$$\overrightarrow{OP_0} = \mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$$

and parallel to

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

is

$$\mathbf{r}(t) = \mathbf{b} + t\mathbf{a}$$

This can be written as

$$\begin{cases} x = a_1 t + b_1 \\ y = a_2 t + b_2 \qquad t \in \mathbb{R} \\ z = a_3 t + b_3 \end{cases}$$

Remark

This definition can be extended to n dimensions by adapting the number of components.

Example

Find the parametric equations of the line through $(1,-2,3)$ and parallel to the vector $\pi\mathbf{i}-3\mathbf{j}+\mathbf{k}$

$\mathbf{a}=\pi\mathbf{i}-3\mathbf{j}+\mathbf{k}$

$\mathbf{b}=\mathbf{i}-2\mathbf{j}+3\mathbf{k}$

$\mathbf{r}(t)=\mathbf{i}-2\mathbf{j}+3\mathbf{k}+t(\pi\mathbf{i}-3\mathbf{j}+\mathbf{k})=(1+\pi t)\mathbf{i}+(-2-3t)\mathbf{j}+(3+t)\mathbf{k}$

The parametric equations may be read as

$$\begin{cases} x=\pi t+1 \\ y=-3t-2 \qquad t\in\mathbb{R} \\ z=t+3 \end{cases}$$

Equations for Planes. Distances. Normal Vectors

In the study of geometry and vector algebra, planes play a crucial role. This section explores how to represent planes in three-dimensional space and calculate distances between points and planes.

Definition (Equation of a Plane)

The general form of a plane Π in three-dimensional space is given by the equation:

$$\Pi : Ax + By + Cz = D$$

where A, B, and C are constants that define the orientation of the plane, and D is a constant that defines its position.

Normal Vector. The vector $\mathbf{n} = (A, B, C)$ is normal (perpendicular) to the plane. Therefore the equation of the plane can also be derived from the dot product of this normal vector with a vector contained in the plane. Such vector is defined as the difference $(x - x_0, y - y_0, z - z_0)$, where the point $P : (x_0, y_0, z_0) \in \Pi$. Thus, the plane is also defined as the point $\mathbf{r} = (x, y, z)$ such that

$$\begin{aligned} \mathbf{n} \cdot (\mathbf{r} - P) &= 0 \\ \mathbf{n} \cdot \mathbf{r} &= \mathbf{n} \cdot P \\ Ax + By + Cz &= D \end{aligned}$$

This formulation highlights the fact that the plane is perpendicular to the vector $\mathbf{n}$.
The distance from a point to a plane is defined as the length of the smallest segment joining the point and the plane.

URJC

Proposition (Distance from a Point to a Plane)

Given a point $P_0(x_0, y_0, z_0)$ and a plane defined by $Ax + By + Cz = D$, the shortest distance d from the point to the plane can be calculated using the formula:

$$d = \frac{|Ax_0 + By_0 + Cz_0 - D|}{\sqrt{A^2 + B^2 + C^2}}$$

This equation represents the perpendicular distance from the point to the plane.

Example

Consider the plane $2x - 3y + 4z = 12$ and the point $P_0(1, 2, 3)$. The distance from a point $P_0(x_0, y_0, z_0)$ to a plane $Ax + By + Cz = D$ is given by:

$$d = \frac{|Ax_0 + By_0 + Cz_0 - D|}{\sqrt{A^2 + B^2 + C^2}}$$

Substituting $A = 2$, $B = -3$, $C = 4$, $D = 12$, and $P_0(1, 2, 3)$, we get:

$$d = \frac{|2(1) - 3(2) + 4(3) - 12|}{\sqrt{2^2 + (-3)^2 + 4^2}} = \frac{|2 - 6 + 12 - 12|}{\sqrt{4 + 9 + 16}} = \frac{4}{\sqrt{29}} \approx 0.743$$

New Coordinate Systems

In many applications, the standard Cartesian coordinate system is not the most convenient. Alternate coordinate systems such as polar, cylindrical, and spherical coordinates are often more suited to problems with specific symmetries or geometries.

Definition (Polar Coordinates)

Polar coordinates are useful in two-dimensional problems where circular symmetry is present. A point in the plane is described by two coordinates (r, θ), where:

- r is the radial distance from the origin, $r \geq 0$
- θ is the angle measured counterclockwise from the positive x-axis, $0 \leq \theta < 2\pi$

The relationship between polar coordinates and Cartesian coordinates (x, y) is:

$$x = r\cos\theta, \quad y = r\sin\theta$$

The relationship between Cartesian coordinates and polar coordinates is:

$$r = \sqrt{x^2 + y^2} \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$

Note: To find the angle, we must take into account the quadrant.

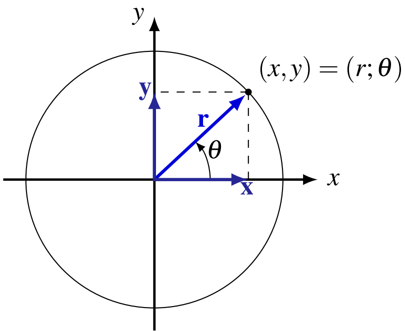


Figure: Polar Coordinates in the Cartesian plane

Example

Given the point in polar coordinates $(r, \theta) = (4, \frac{\pi}{4})$, convert to Cartesian coordinates.

$$x = r\cos\theta = 4\cos\frac{\pi}{4} = 4 \times \frac{\sqrt{2}}{2} = 2\sqrt{2}$$
$$y = r\sin\theta = 4\sin\frac{\pi}{4} = 4 \times \frac{\sqrt{2}}{2} = 2\sqrt{2}$$

Thus, the Cartesian coordinates are $(x, y) = (2\sqrt{2}, 2\sqrt{2})$.

Definition (Cylindrical Coordinates)

Cylindrical coordinates extend polar coordinates to three dimensions by adding a height component z. A point in space is described by three coordinates (r, θ, z), where:

- r is the radial distance in the xy-plane
- θ is the angle in the xy-plane, measured from the positive x-axis
- z is the vertical distance from the xy-plane

The relationship between cylindrical and Cartesian coordinates (x, y, z) is:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z$$

The inverse formulas coincide with the polar formulas.

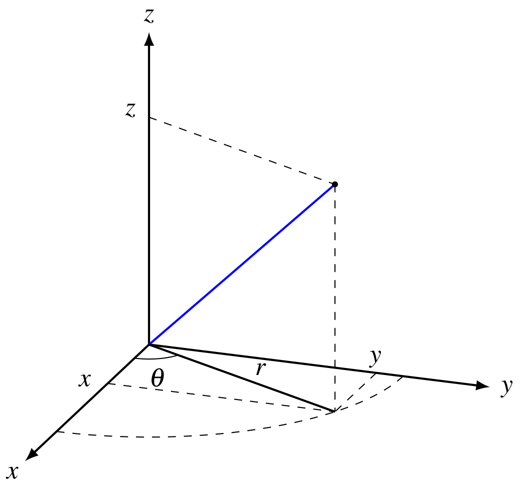


Figure: Cylindrical Coordinates in the Cartesian space

Example

Convert the Cartesian point $(x,y,z) = (3,3,5)$ to cylindrical coordinates.

$$r = \sqrt{x^2+y^2} = \sqrt{3^2+3^2} = \sqrt{18} = 3\sqrt{2}$$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}\left(\frac{3}{3}\right) = \frac{\pi}{4}$$

$$z = 5$$

Thus, the cylindrical coordinates are $(r,\theta,z) = (3\sqrt{2}, \frac{\pi}{4}, 5)$.

Definition (Spherical Coordinates)

Spherical coordinates are useful in three-dimensional problems with spherical symmetry. A point is described by three coordinates (ρ, θ, ϕ), where:

- ρ is the radial distance from the origin, $\rho \geq 0$.
- θ is the angle in the xy-plane (similar to cylindrical coordinates), $0 \leq \theta < 2\pi$.
- ϕ is the angle from the positive z-axis (the polar angle), $0 \leq \phi \leq \pi$

The relationship between spherical and Cartesian coordinates (x, y, z) is:

$$x = \rho \sin\phi \cos\theta, \quad y = \rho \sin\phi \sin\theta, \quad z = \rho \cos\phi$$

The relationship between Cartesian and spherical coordinates (x, y, z) is:

$$\rho = \sqrt{x^2 + y^2 + z^2} \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right) \qquad \phi = \cos^{-1}\left(\frac{z}{\rho}\right)$$

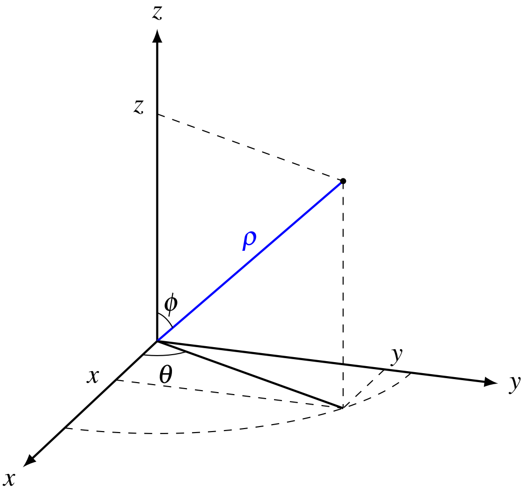


Figure: Spherical Coordinates in the Cartesian space

Example

Convert the Cartesian point $(x, y, z) = (1, 1, \sqrt{2})$ to spherical coordinates.

$$\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{1^2 + 1^2 + (\sqrt{2})^2} = \sqrt{4} = 2$$
$$\theta = \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}(1) = \frac{\pi}{4}$$
$$\phi = \cos^{-1}\left(\frac{z}{\rho}\right) = \cos^{-1}\left(\frac{\sqrt{2}}{2}\right) = \frac{\pi}{4}$$

Thus, the spherical coordinates are $(\rho, \theta, \phi) = \left(2, \frac{\pi}{4}, \frac{\pi}{4}\right)$.

Definition (Disk (balls))

Let $x_0 \in \mathbb{R}^n$ and let r a positive real number. We define the open disk of radius r and center $\mathbf{x}_0$ is defined as the set of all the points x such that $\|x - x_0\| < r$ and it is denoted as $D_r(x_0)$.

Example (Disks in usual spaces)

Let $r > 0$

- $(n = 1)$ Let $x_0 \in \mathbb{R}$, then $D_r(x_0)$ is the interval $(x_0 - r, x_0 + r) = \{x \in \mathbb{R} : |x - x_0|\}$
- $(n = 2)$ Let $x_0 \in \mathbb{R}^2$, then $D_r(x_0)$ is the circle $\{x \in \mathbb{R}^2 : \|x - x_0\| < r\}$

Definition (Open Set)

Let $U \subset \mathbb{R}^n$, we say that U is an open set if for each point $x_0 \in U$ there is some $r > 0$ such that $D_r(x_0) \in U$, that is we can find a disk centred at x_0 inside U.

Every disk is an open subset. We use the convention that the empty subset is an open subset.

Example

Prove that $A = \{(x,y) \in \mathbb{R}^2 : x > 0\}$ is an open set.
We must show that $\forall (x,y) \in A \quad \exists r > 0$ such that $D_r(x_0) \in A$. Set $r = x$.
If $(x_1, y_1) \in D_r(x,y)$ then

$$|x - x_1| = \sqrt{(x_1 - x)^2} \leq \sqrt{(x_1 - x)^2 + (y_1 - y)^2} < r = x$$

Thus, we have $|x_1 - x| < x$, this means $x_1 - x < x$ and $x - x_1 < x$. Therefore, $x_1 > 0$, that is $(x_1, y_1) \in A$. Thus, $D_r(x,y) \subset A$ which implies A is open.

Definition (Neighbourhood)

We define neighbourhood of a point $x_0 \in \mathbb{R}^n$ to an open subset containing the point. For example any disk centred at x_0 is a neighbourhood of x_0 for all $r > 0$.

Definition (Bounded Set)

A subset $A \subset \mathbb{R}^n$ is called **bounded** if there exists a real number $M > 0$ such that for all points $x \in A$, the norm $||x||$ satisfies:

$$||x|| \leq M$$

In other words, the set is bounded if all its elements lie within some finite distance from the origin, i.e., the set can be contained in a ball of some finite radius in $\mathbb{R}^n$.

URJC

Example

1. **A finite set**: Any finite set of points, such as $\{(1,0),(0,1),(1,1)\}$, is bounded. There is a finite distance between all points in the set, so it fits within some ball of radius M.
2. **A closed interval**: The set $A=[0,1]\subset\mathbb{R}$ is bounded because all points in this interval have a norm less than or equal to 1.
3. **The closed unit ball**: The set $B=\{x\in\mathbb{R}^n \mid \|x\|\leq 1\}$ is bounded because all points in the set satisfy $\|x\|\leq 1$.
4. **The set of points inside a circle**: The set $C=\{x\in\mathbb{R}^2 \mid \|x\|\leq 2\}$ is bounded because all points within the circle are within a finite distance of the origin.
5. **The set of rational points in a bounded region**: For example, the set of points $A=\{(x,y)\in\mathbb{Q}^2 \mid 0\leq x\leq 1, 0\leq y\leq 1\}$ is bounded since it lies within the unit square in $\mathbb{R}^2$.

Definition (Boundary points)

Let $A \subset \mathbb{R}^n$. A point x is called boundary point of A if every neighbourhood of x contains at least one point on A and one point not in A.
We call the boundary of a set A and it is denoted as ∂A to the set of all the boundary points of A.

Example

1. Let $A = (a,b]$, an interval. We have $\partial A = \{a,b\}$. Note that the boundary points can be either in or out of the original set.
2. Let $A = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < r^2\} = D_0(r)$, a circle of radius r. We have $\partial A = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = r^2\}$
3. Let the discrete set $A = \{1,2,3,4\}$, we have $\partial A = A$.

Definition (Closed set)

We say B is closed if and only if B^c is open, i.e., it has open complement. In $\mathbb{R}^n$ a closed subset is a set that contains all its boundary points $\partial B \subset B$

Example

1. **The entire space** $\mathbb{R}^n$: The whole space is trivially closed because it contains all of its boundary points (there are no points "outside" $\mathbb{R}^n$).
2. **The set of points satisfying a continuous equation**: For example, the set $A = \{x \in \mathbb{R}^n \mid x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}$ (the unit sphere boundary in $\mathbb{R}^n$) is closed beacuse $\partial A = A$.
3. **A bounded set**: For example, the set $A = \{x \in \mathbb{R}^n \mid x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1\}$ (the unit sphere in $\mathbb{R}^n$) is closed beacuse $\partial A \subset A$.
4. Let $A = (a, b]$, an interval. We have $\partial A = \{a, b\}$. This is not a closed set, because $a \notin A$.
5. **Closed intervals**: Any closed interval, such as $[a, b] \subset \mathbb{R}$, is closed in $\mathbb{R}^n$. It contains all of its boundary points.

Definition

Compact A subset $K \subset \mathbb{R}^n$ is called **compact** if it is both **closed** and **bounded**.

Note: This characterization only applies to $\mathbb{R}^n$ thanks to the Heine-Borel theorem.

Example

1. **Closed and bounded intervals**: A closed interval $[a,b] \subset \mathbb{R}$ is compact because it is both closed and bounded.
2. **The closed unit ball**: The set $B = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ is compact. It is closed because it includes all points on the boundary, and it is bounded because all points in the set satisfy $\|x\| \leq 1$.
3. **A finite set**: Any finite set of points in $\mathbb{R}^n$, such as $\{(1,0),(0,1),(1,1)\}$, is compact. It is closed and bounded, and hence compact.
4. **The Cantor set in $\mathbb{R}$**: The Cantor set is a classic example of a compact set. It is closed (contains all its limit points) and bounded (it lies within the interval $[0,1]$).

A **closed set** need not be bounded, while a **compact set** must be both closed and bounded. Every compact set is closed and bounded, but not every closed set is compact. For example, the set $\{x \in \mathbb{R} \mid x \geq 0\}$ is closed but not compact because it is unbounded.

§II. Differentiation in Several Variables

- Functions of Several Variables
  - Domain and Range of Functions of Several Variables
  - Limits of Functions of Several Variables
  - Continuity of Functions of Several Variables
  - Level Curves and Surfaces

Functions of Several Variables

In many real-world scenarios, we encounter functions that depend on more than one variable. These functions are essential in multivariable calculus, physics, economics, and engineering. This section introduces functions of several variables, their domains, and ranges, along with important properties and examples.

Definition (Scalar function)

A **scalar function of several variables** is a rule that assigns a real number $f(x_1, x_2, \ldots, x_n)$ to each element $(x_1, x_2, \ldots, x_n)$ in a subset D of $\mathbb{R}^n$. The set D is called the **domain** of the function, and the set of all possible values of f is called the **range**.

$$f : D \subset \mathbb{R}^n \to R \subset \mathbb{R}$$

Definition (Vectorial function)

A **vectorial function of several variables** is a rule that assigns a vector $f(x_1, x_2, \ldots, x_n) \in \mathbb{R}^m$ to each element $(x_1, x_2, \ldots, x_n)$ in a subset D of $\mathbb{R}^n$. The set D is called the **domain** of the function, and the set of all possible values of f is called the **range**.

$$f : D \subset \mathbb{R}^n \to R \subset \mathbb{R}^m$$

Example

For example, $f(x,y) = x^2 + y^2$ is a function of two variables, and $f(x,y,z) = x + 2y + 3z$ is a function of three variables, these are scalar functions.
On the other hand, the function

$$f(x,y) = (x,y,x+y)$$

is a vector function $f : \mathbb{R}^2 \to \mathbb{R}^3$.

In this text, we primarily consider scalar functions of two or three variables $f(x,y)$ or $f(x,y,z)$.

Definition (Domain)

The **domain** of a function $f(x_1, x_2, \ldots, x_n)$ is the set of all possible points $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ for which the function is defined.

Example

Consider the function $f(x,y) = \sqrt{1 - x^2 - y^2}$. The function is defined only when $1 - x^2 - y^2 \geq 0$, or equivalently when $x^2 + y^2 \leq 1$. Hence, the domain of this function is the unit disk:

$$D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$$

Example

For the function $f(x,y,z) = \frac{1}{x^2+y^2+z^2-1}$, the function is undefined when $x^2 + y^2 + z^2 = 1$. Therefore, the domain is:

$$D = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \neq 1\}$$

Definition (Range)

The **range** of a function $f(x_1, x_2, \ldots, x_n)$ is the set of all possible values that the function can take, i.e., the set:

$$R = \{f(x_1, x_2, \ldots, x_n) \mid (x_1, x_2, \ldots, x_n) \in D\}$$

Example

For the function $f(x,y) = x^2 + y^2$, the range is all non-negative real numbers, since the sum of squares is always non-negative. Therefore, the range is:

$$R = [0, \infty)$$

Example

For the function $f(x,y,z) = x + y + z$, the range is all real numbers, since for any real number r, we can find values of x, y, z such that $f(x,y,z) = r$. Thus, the range is:

$$R = \mathbb{R}$$

Proposition

The sum, difference, product, and quotient of two functions of several variables are also functions of several variables, provided that the operations are defined within their domains.

Proof. Let $f(x_1,x_2,\dots,x_n)$ and $g(x_1,x_2,\dots,x_n)$ be functions of several variables. Then:

$$(f+g)(x_1,x_2,\dots,x_n) = f(x_1,x_2,\dots,x_n) + g(x_1,x_2,\dots,x_n)$$

$$(f\cdot g)(x_1,x_2,\dots,x_n) = f(x_1,x_2,\dots,x_n)\cdot g(x_1,x_2,\dots,x_n)$$

and so on, are also functions of several variables. The quotient $\frac{f(x_1,x_2,\dots,x_n)}{g(x_1,x_2,\dots,x_n)}$ is defined when $g(x_1,x_2,\dots,x_n)\neq 0$. □

Definition (Composition of functions)

If g maps A to B and f maps B to C, the composition of g with f, or of f on g, denoted by $h = f \circ g$, is a function that maps A to C by sending x to $f(g(x))$,

$$h = f \circ g : A \to C,$$
$$h(x) = f(g(x))$$

Example (Composition)

1. Let $f(x) = \sin x$ and $g(x) = x^2$, then $f \circ g = f(g(x)) = \sin x^2$.
2. Let $x(s,t) = (s+t^2, t)$ and $z(x,y) = (x+y, x-y^2)$, then $z(s,t) = (s+t^2+t, s)$

Remark

As we can see on the last example, the range of the inner function must coincide with the domain of the outer function, in terms of elements and dimensions.

URJC

Definition (Limit)

Let $f(x,y)$ be a function defined on a subset of $\mathbb{R}^2$, except possibly at the point (a,b). We say that the **limit** of $f(x,y)$ as (x,y) approaches (a,b) is L, and write:

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

if, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $0 < \sqrt{(x-a)^2+(y-b)^2} < \delta$, then:

$$|f(x,y) - L| < \varepsilon$$

This definition extends to functions of more than two variables in a similar way, replacing $(x,y) \in \mathbb{R}^2$ with $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Proposition (Properties of Limits of Several Variables)

If $\lim_{(x,y)\to(a,b)} f(x,y) = L$ and $\lim_{(x,y)\to(a,b)} g(x,y) = M$, then the following properties hold:

1. **Sum:** $\lim_{(x,y)\to(a,b)} [f(x,y) + g(x,y)] = L + M$
2. **Difference:** $\lim_{(x,y)\to(a,b)} [f(x,y) - g(x,y)] = L - M$
3. **Product:** $\lim_{(x,y)\to(a,b)} [f(x,y) \cdot g(x,y)] = L \cdot M$
4. **Quotient:** If $M \neq 0$, then $\lim_{(x,y)\to(a,b)} \dfrac{f(x,y)}{g(x,y)} = \dfrac{L}{M}$

These properties are analogous to those for limits of single-variable functions and make it easier to compute limits of functions of several variables.

Example

Consider the function $f(x,y) = x^2 + y^2$. We want to compute the limit as $(x,y) \to (0,0)$:

$$\lim_{(x,y)\to(0,0)} (x^2 + y^2)$$

As (x,y) approaches $(0,0)$, both x^2 and y^2 approach 0. Therefore using the sum property we have

$$\lim_{(x,y)\to(0,0)} (x^2 + y^2) = 0$$

Remark

In one variable x, to approach a point x_0 we use only two trajectories, $x \to x_0^+$ and $x \to x_0^-$. These paths define the notion of lateral limits. We said that the limit exist, if and only if both lateral limits coincide. Now, in the plane, to reach a point, we have infinite paths, and the equivalent notion if that, the limit exists, if an only if, the limit trough all paths coincide. In practice, we can not check all the paths (they are infinite!), however, this give us a criteria to prove that a limit does not exist: finding two trajectories on which the limit differs.

Example

Consider the function $f(x,y) = \frac{xy}{x^2+y^2}$. We want to compute the limit as $(x,y) \to (0,0)$. We try approaching $(0,0)$ along different paths:

- Along the line $y = 0$, the function becomes:

$$f(x,0) = \frac{0}{x^2} = 0$$

- Along the line $x = 0$, the function becomes:

$$f(0,y) = \frac{0}{y^2} = 0$$

- Along the line $y = x$, the function becomes:

$$f(x,x) = \frac{x^2}{x^2+x^2} = \frac{1}{2}$$

Since the limit depends on the path taken, the limit does not exist.

Theorem (Squeeze Theorem for Several Variables)

Let $f(x,y)$, $g(x,y)$, and $h(x,y)$ be functions defined on a domain D, except possibly at (a,b). If $f(x,y) \leq g(x,y) \leq h(x,y)$ for all (x,y) near (a,b), and:

$$\lim_{(x,y)\to(a,b)} f(x,y) = \lim_{(x,y)\to(a,b)} h(x,y) = L$$

then:

$$\lim_{(x,y)\to(a,b)} g(x,y) = L$$

Proof. Since $f(x,y) \leq g(x,y) \leq h(x,y)$, for any $\varepsilon > 0$, there exists $\delta > 0$ such that:

$$|f(x,y) - L| < \varepsilon \quad \text{and} \quad |h(x,y) - L| < \varepsilon$$

for all (x,y) with $0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta$. By the inequality, it follows that $|g(x,y) - L| < \varepsilon$, which proves the result. □

Theorem

Let $f, g : \mathbb{R}^n \to \mathbb{R}$, if $\lim_{x \to x_0} f = 0$, that is f is infinitesimal, and $\|g\| \leq M$, for some positive constant M in a set closed A such that $x_0 \in A$, that is, g is bounded, then

$$\lim_{x \to x_0} f \cdot g = 0$$

Example

Prove that

$$\lim_{(x,y) \to (0,0)} \frac{x^2 y}{x^2 + y^2} = 0$$

Note that the function $g(x,y) = \frac{x^2}{x^2+y^2} \leq \frac{x^2}{x^2} = 1$, thus is bounded for all $(x,y) \in \mathbb{R}^2$. Also we have that $\lim_{y \to 0} y = 0$, thus $f(x,y) = y$ is infinitesimal. Then it holds that

$$\lim_{(x,y) \to (0,0)} \frac{x^2 y}{x^2 + y^2} = 0$$

Definition (Continuous function)

A function $f(x,y)$ is **continuous** at (a,b) if:

$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$$

If $f(x,y)$ is continuous at every point in its domain, it is said to be continuous on the domain.

Example

The function $f(x,y) = x^2 + y^2$ is continuous everywhere, since the limit of $f(x,y)$ as $(x,y) \to (a,b)$ equals $f(a,b) = a^2 + b^2$.

Example

The function $f(x,y) = \frac{xy}{x^2+y^2}$ is not continuous at $(0,0)$, since the limit does not exist at that point, as demonstrated earlier.

Proposition (Operations on Continuous Functions)

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be continuous functions, where $D \subset \mathbb{R}^n$ is an open subset. Then the following functions are continuous on D:

1. **Sum:** $f+g$ defined by $(f+g)(x) = f(x)+g(x)$, for all $x \in D$.
2. **Difference:** $f-g$ defined by $(f-g)(x) = f(x)-g(x)$, for all $x \in D$.
3. **Product:** $f \cdot g$ defined by $(f \cdot g)(x) = f(x) \cdot g(x)$, for all $x \in D$.
4. **Quotient:** $\frac{f}{g}$ defined by $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$, for all $x \in D$ where $g(x) \neq 0$.
5. **Composition:** $f \circ g$ defined by $(f \circ g)(x) = f(g(x))$, for all $x \in D$ if $g(x) \in \mathrm{dom}(f)$, where $\mathrm{dom}(f)$ denotes the domain of f.
6. **Scalar multiplication:** $c \cdot f$, for any constant $c \in \mathbb{R}$, defined by $(c \cdot f)(x) = c \cdot f(x)$, for all $x \in D$.

Definition (Level Curves)

For functions of two variables, the level curves of $f(x,y)$ are curves in the xy-plane where $f(x,y)$ takes on a constant value. For functions of three variables, the level surfaces are surfaces in $\mathbb{R}^3$ where the function takes on a constant value.

Example

Consider $f(x,y) = x^2 + y^2$. The level curves are given by:

$$x^2 + y^2 = c$$

which are circles with radius $\sqrt{c}$, for $c \geq 0$.

Example

Consider $f(x,y,z) = x^2 + y^2 + z^2$. The level surfaces are spheres centered at the origin with radius $\sqrt{c}$, where:

$$x^2 + y^2 + z^2 = c$$

Example

Consider the function $f(x,y) = x^2 + y^2$. The domain is $\mathbb{R}^2$, and the range is $[0,\infty)$. The level curves are circles centered at the origin.

Example

Consider $f(x,y,z) = \frac{xy}{z}$, with the domain $\{(x,y,z) \in \mathbb{R}^3 \mid z \neq 0\}$. The range is $\mathbb{R}$, excluding the points where $z = 0$.

§II. Differentiation in Several Variables

- Differentiation in Several Variables
  - Partial Derivatives and Differentiability
  - Higher-Order Partial Derivatives
  - Differentiation of Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$
  - Directional Derivatives and the Gradient

Definition (Partial Derivatives)

Let $f(x,y,z)$ be a function of multiple variables. The partial derivatives of f with respect to x, y, and z are defined as follows:

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y, z) - f(x,y,z)}{\Delta x},$$

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y, z) - f(x,y,z)}{\Delta y},$$

$$\frac{\partial f}{\partial z} = \lim_{\Delta z \to 0} \frac{f(x, y, z+\Delta z) - f(x,y,z)}{\Delta z}.$$

These derivatives measure how the function f changes as we vary one variable while keeping the others fixed. In practice to compute a partial derivative, we treat the function as one variable function by considering the remaining variables as constant.

Example

Let $f(x,y) = x^2 + y^3$. The partial derivatives are $\dfrac{\partial f}{\partial x} = 2x, \quad \dfrac{\partial f}{\partial y} = 3y^2$.

Example

For the function $f(x,y,z) = xyz$, the partial derivatives are:

$$\frac{\partial f}{\partial x} = yz, \quad \frac{\partial f}{\partial y} = xz, \quad \frac{\partial f}{\partial z} = xy.$$

Definition (Gradient)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. The **gradient** of f, denoted by ∇f, is the vector of all its first-order partial derivatives:

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right).$$

Example

The gradient of the function $f(x,y,z) = xyz$ is:

$$\nabla f(x,y,z) = (yz, xz, xy).$$

Definition (Differentiability of Functions)

A function $f(x_1, x_2, \dots, x_n)$ is said to be *differentiable* at a point $\mathbf{a} = (a_1, a_2, \dots, a_n)$ if it can be well-approximated by a linear function near that point. More precisely, f is differentiable at $\mathbf{a}$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}$ such that:

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

In this case, the map L is given by the gradient of f at $\mathbf{a}$, i.e.,

$$L(\mathbf{h}) = \nabla f(\mathbf{a}) \cdot \mathbf{h}.$$

Theorem (Differentiability Implies Continuity)

If a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point $\mathbf{a} \in \mathbb{R}^n$, then f is continuous at $\mathbf{a}$.

Proof. By the definition of differentiability, we have:

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

Since $L(\mathbf{h})$ is a linear map, $\|L(\mathbf{h})\| \to 0$ as $\|\mathbf{h}\| \to 0$, implying that:

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}).$$

Hence, f is continuous at $\mathbf{a}$. □

The partial derivatives of a function are necessary but not sufficient for differentiability. For a function $f(x_1, x_2, \ldots, x_n)$ to be differentiable at $\mathbf{a}$, the following conditions are typically required:

- The partial derivatives of f exist in a neighborhood of $\mathbf{a}$.
- The partial derivatives are continuous at $\mathbf{a}$ (a condition known as *continuous partial derivatives*).

Example (A Function with Non-continuous Partial Derivatives)

Consider the function:

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

The partial derivatives of f exist at $(0,0)$, but they are not continuous at that point. Hence, f is not differentiable at $(0,0)$.

Proposition (Properties of Differentiable Functions)

If a function $f(x_1, x_2, \ldots, x_n)$ is differentiable, it satisfies the following properties:

- **Linearity of Differentiation:** For any scalar functions f and g, and any constants α and β,

$$\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g.$$

- **Product Rule:** If f and g are differentiable functions, then

$$\nabla(fg) = f\nabla g + g\nabla f.$$

Definition (Tangent hyper-plane)

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ an explicit function $y = f(x_1, x_2, \cdots, x_n)$ and $F(x_1, x_2, \ldots, x_n) = 0$ an implict function with domain D. Assume $F, f \in C^1(X)$. Then the tangent hyper-plane at $\mathbf{x_0} \in D$ can be found using the following formulas

1. Tangent hyper-plane to f: $\nabla f(\mathbf{x_0}) \cdot (\mathbf{x} - \mathbf{x_0}) = y - f(\mathbf{x_0})$
2. Tangent hyper-plane to F: $\nabla F(\mathbf{x_0}) \cdot (\mathbf{x} - \mathbf{x_0}) = 0$

Remark

The term hyper-plane is the generalization of a linear variety for several dimensions. For example, we know that in $\mathbb{R}^2$ the tangent hyper-plane corresponds to a tangent line, and in the space $\mathbb{R}^3$ the tangent variety is a tangent plane. In this text we focus in the tangent plane for surfaces.

Tangency for scalar-valued functions of two variables

We have the normal vector and a point of the tangent plane.

$$\mathbf{n} = \mathbf{u}\times\mathbf{v} = -f_x(a,b)\mathbf{i} - f_y(a,b)\mathbf{j} + \mathbf{k}; \quad P = (a,b,f(a,b))$$

So, the **equation for the tangent plane** through $(a,b,f(a,b))$ with normal $\mathbf{n}$ is

$$(-f_x(a,b), -f_y(a,b), 1)\cdot(x-a, y-b, z-f(a,b)) = 0$$

Proposition (Properties of the tangent plane)

- If the graph of $z = f(x,y)$ has a tangent plane at $(a,b,f(a,b))$ then that tangent plane has equation

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

- Let us define the function $h(x,y)$ to be

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

- Then h has the following properties
  1. $h(a,b) = f(a,b)$ The values of h and f are the same at (a,b)
  2. $\frac{\partial h}{\partial x}(a,b) = \frac{\partial f}{\partial x}(a,b)$ and $\frac{\partial h}{\partial y}(a,b) = \frac{\partial f}{\partial y}(a,b)$ Partial derivatives of h and f are the same at (a,b)

Definition (Differentiability)

Let X be open in $\mathbb{R}^2$ and let $f : X \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a scalar-valued function of two variables. We say that f is **differentiable** at $(a,b) \in X$ if

- The partial derivatives $f_x(a,b)$ and $f_y(a,b)$ exist , and
- The function

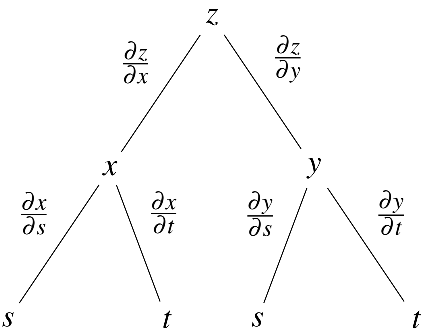
$$h(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$

is a good linear approximation to f near (a,b):

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y) - h(x,y)}{\|(x,y) - (a,b)\|} = 0$$

The chain rule allows us to compute the derivative of a composite function, where each variable depends on one or more independent variables. For a function $z = f(x, y)$ where both x and y depend on other variables variable t and s, the chain rule is given by:

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$



The chain rule is the result of covering all the trajectories in the tree. $\frac{\partial z}{\partial t}$ means that we must cover all the paths from z to t. The subsequent branches are multiplied, different branches are added up. For example

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial s}$$

Example (Chain Rule in Two Variables)

Suppose $z = x^2 + y^2$, where $x = t^2$ and $y = \sin(t)$. Applying the chain rule, we find:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Here,

$$\frac{\partial z}{\partial x} = 2x, \quad \frac{\partial z}{\partial y} = 2y,$$

$$\frac{dx}{dt} = 2t, \quad \frac{dy}{dt} = \cos(t).$$

Thus,

$$\frac{dz}{dt} = 2x \cdot 2t + 2y \cdot \cos(t) = 4t^3 + 2\sin(t)\cos(t).$$

The Chain Rule for Multivariable Functions

If f is a function of multiple variables $x_1, x_2, \ldots, x_n$, where each x_i is a function of t, the chain rule is generalized as:

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}.$$

Example

Let $f(x,y,z) = x + y^2 + z^3$, where $x = t^2$, $y = e^t$, and $z = \cos(t)$. Applying the chain rule,

we compute:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}.$$

Since:

$$\frac{\partial f}{\partial x} = 1, \quad \frac{\partial f}{\partial y} = 2y, \quad \frac{\partial f}{\partial z} = 3z^2,$$

and

$$\frac{dx}{dt} = 2t, \quad \frac{dy}{dt} = e^t, \quad \frac{dz}{dt} = -\sin(t),$$

we have:

$$\frac{df}{dt} = 1 \cdot 2t + 2e^t \cdot e^t + 3\cos^2(t) \cdot (-\sin(t)),$$

which simplifies to:

$$\frac{df}{dt} = 2t + 2e^{2t} - 3\cos^2(t)\sin(t).$$

Example

Consider a function $f(x,y) = x^2 + y^2$, where $x = g(t)$ and $y = h(t)$, with $g(t) = t^2$ and $h(t) = \sin(t)$. Applying the chain rule iteratively, we first differentiate $f(x,y)$ with respect to t:

$$\frac{df}{dt} = 2x\frac{dx}{dt} + 2y\frac{dy}{dt}.$$

Substituting $x = g(t)$ and $y = h(t)$ yields:

$$\frac{df}{dt} = 2(t^2) \cdot (2t) + 2\sin(t) \cdot \cos(t),$$

which simplifies to:

$$\frac{df}{dt} = 4t^3 + 2\sin(t)\cos(t).$$

The gradient of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ is a vector that points in the direction of the greatest rate of increase of f. It is a key concept and plays an important role in optimization and vector calculus.
The gradient can be thought of geometrically as the direction of steepest ascent of the function at a point.

Theorem (Geometric Interpretation of the Gradient)

The gradient $\nabla f(\mathbf{a})$ of a function f at a point $\mathbf{a} \in \mathbb{R}^n$ points in the direction of the maximum rate of increase of f at $\mathbf{a}$. Moreover, the magnitude of the gradient, $\|\nabla f(\mathbf{a})\|$, gives the rate of increase in that direction.

Proof. The directional derivative $D_{\mathbf{u}} f(\mathbf{a})$ gives the rate of change of f in the direction of a unit vector $\mathbf{u}$. By the formula $D_{\mathbf{u}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{u}$, the directional derivative is maximized when $\mathbf{u}$ points in the same direction as $\nabla f(\mathbf{a})$, i.e., when $\mathbf{u} = \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|}$. In this case, the maximum value of the directional derivative is $\|\nabla f(\mathbf{a})\|$. □

URJC

Example

Let $f(x,y,z) = x^2 + y^2 + z^2$. The gradient of f is:

$$\nabla f(x,y,z) = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}.$$

At the point $(1,1,1)$, the gradient is:

$$\nabla f(1,1,1) = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

This indicates that the direction of steepest ascent of f at $(1,1,1)$ is along the vector $(2,2,2)$, and the rate of change in this direction is $\|\nabla f(1,1,1)\| = 2\sqrt{3}$.

Directional Derivative

The directional derivative of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ measures the rate of change of f in a specific direction, rather than along one of the coordinate axes.

Definition (Directional Derivatives)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function at $\mathbf{a} \in \mathbb{R}^n$, and let $\mathbf{u} \in \mathbb{R}^n$ be a unit vector (i.e., $||\mathbf{u}|| = 1$). The **directional derivative** of f at $\mathbf{a}$ in the direction of $\mathbf{u}$ is defined as:

$$D_{\mathbf{u}} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{u}) - f(\mathbf{a})}{h}.$$

If the function f is differentiable at $\mathbf{a}$, the directional derivative can be computed using the gradient of f as follows:

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\mathbf{a}$, and let $\mathbf{u} \in \mathbb{R}^n$ be a unit vector. Then the directional derivative of f at $\mathbf{a}$ in the direction of $\mathbf{u}$ is given by:

$$D_{\mathbf{u}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{u},$$

where $\nabla f(\mathbf{a})$ is the gradient of f at $\mathbf{a}$, and $\cdot$ denotes the dot product.

Proof. The directional derivative is defined as:

$$D_{\mathbf{u}} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{u}) - f(\mathbf{a})}{h}.$$

Since f is differentiable at $\mathbf{a}$, the linear approximation of $f(\mathbf{a} + h\mathbf{u})$ is given by:

$$f(\mathbf{a} + h\mathbf{u}) \approx f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (h\mathbf{u}),$$

so:

$$D_{\mathbf{u}} f(\mathbf{a}) = \lim_{h \to 0} \frac{\nabla f(\mathbf{a}) \cdot h\mathbf{u}}{h} = \nabla f(\mathbf{a}) \cdot \mathbf{u}.$$

□

Example

Let $f(x,y) = x^2 + y^2$. To find the directional derivative of f at the point $\mathbf{a} = (1,1)$ in the direction of the unit vector $\mathbf{u} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, we first compute the gradient of f:

$$\nabla f(x,y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}.$$

At $(1,1)$, the gradient is:

$$\nabla f(1,1) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

The directional derivative is:

$$D_{\mathbf{u}} f(1,1) = \nabla f(1,1) \cdot \mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = 2\sqrt{2}.$$

Proposition (Properties of the Gradient and Directional Derivatives)

Linearity of Directional Derivatives: The directional derivative is linear in both the function and the direction vector:

$$D_{\mathbf{u}}(f+g) = D_{\mathbf{u}}f + D_{\mathbf{u}}g, \quad D_{\mathbf{u}}(cf) = cD_{\mathbf{u}}f, \quad c \in \mathbb{R}.$$

Relationship to Level Sets: The gradient of a function f at a point $\mathbf{a}$ is perpendicular to the level set of f at $\mathbf{a}$, i.e., the set of points where $f(\mathbf{x}) = f(\mathbf{a})$. **Maximum Rate of Change:** The maximum value of the directional derivative $D_{\mathbf{u}}f(\mathbf{a})$ occurs when $\mathbf{u}$ is in the direction of the gradient $\nabla f(\mathbf{a})$. **Minimum Rate of Change:** The minimum value of the directional derivative $D_{\mathbf{u}}f(\mathbf{a})$ occurs when $\mathbf{u}$ is opposite to the direction of the gradient $\nabla f(\mathbf{a})$.

Higher-order partial derivatives are obtained by taking the partial derivative of a function multiple times with respect to one or more variables. Specifically, for a function $f(x_1, x_2, \ldots, x_n)$, the second-order partial derivatives are given by:

$$\frac{\partial^2 f}{\partial x_i^2}, \quad \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Mixed Partial Derivatives and Symmetry of Second-Order Derivatives

Mixed partial derivatives are derivatives where the function is differentiated with respect to more than one variable. For example, the mixed partial derivative of a function $f(x,y)$ with respect to x and then y is:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

Theorem (Symmetry of Second-Order Partial Derivatives (Clairaut's Theorem))

If the partial derivatives $\frac{\partial^2 f}{\partial x \partial y}$ *and* $\frac{\partial^2 f}{\partial y \partial x}$ *exist and are continuous in a neighbourhood of a point* (x_0, y_0)*, then:*

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

URJC

Definition (Hessian Matrix)

Suppose $f : U \subset \mathbb{R}^n \to \mathbb{R}$ has second-order continuous derivatives $(\partial^2 f/\partial x_i \partial x_j)(\mathbf{x}_0)$ for $i, j = 1, \ldots, n$ at a point $\mathbf{x}_0 \in U$. The **Hessian matrix** of f at $\mathbf{x}_0$ is defined as

$$H_f(\mathbf{x_0}) \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \frac{\partial^2 f}{\partial x_i \partial x_j} & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}_{\mathbf{x}=\mathbf{x_0}}$$

Remark

Note that if the function $f \in C^2(D)$ then Clairaut's theorem holds, and the hessian matrix must be **symmetric**.

Example (Hessian matrix)

Determine the Hessian matrix of the function $f(x,y,z) = x^2y + yz^3 + \sin(xz)$

Compute the first-order partial derivatives

$$\frac{\partial f}{\partial x} = 2xy + z\cos(xz) \qquad \frac{\partial f}{\partial y} = x^2 + z^3 \qquad \frac{\partial f}{\partial z} = 3yz^2 + x\cos(xz)$$

Compute the second-order partial derivatives

$$\frac{\partial^2 f}{\partial x^2} = 2y - z^2\sin(xz) \qquad \frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x} = 2x \qquad \frac{\partial^2 f}{\partial x\partial z} = \frac{\partial^2 f}{\partial z\partial x} = \cos(xz) - xz\sin(xz)$$

$$\frac{\partial^2 f}{\partial y^2} = 0 \qquad \frac{\partial^2 f}{\partial y\partial z} = \frac{\partial^2 f}{\partial z\partial y} = 3z^2 \qquad \frac{\partial^2 f}{\partial z^2} = 6yz - x^2\sin(xz)$$

The Hessian matrix is

$$H_f = \begin{bmatrix} 2y - z^2\sin(xz) & 2x & \cos(xz) - xz\sin(xz) \\ 2x & 0 & 3z^2 \\ \cos(xz) - xz\sin(xz) & 3z^2 & 6yz - x^2\sin(xz) \end{bmatrix}$$

This theorem implies that the order in which the partial derivatives are taken does not matter, provided the mixed second-order partial derivatives are continuous.

Example

Let $f(x,y) = x^2y + y^3$. The mixed partial derivatives are:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(2xy) = 2x,$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(2xy) = 2x.$$

Since these mixed partial derivatives are equal, they satisfy Clairaut's Theorem.

Example (Symmetry in Higher Dimensions)

Let $f(x,y,z) = x^2yz$. The mixed second-order partial derivatives are:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(2xyz) = 2xz,$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(2xyz) = 2xz.$$

Again, these derivatives are equal, showing the symmetry of mixed partials in this case.

Higher-Order Derivatives in Multivariable Calculus

Higher-order partial derivatives can extend to any number of orders. For example, the third-order partial derivative of a function $f(x,y,z)$ with respect to x, y, and z is given by:

$$\frac{\partial^3 f}{\partial x \partial y \partial z}.$$

In practice, the order of differentiation for mixed derivatives can often be interchanged, as long as the function is sufficiently smooth (i.e., has continuous partial derivatives of the required order).

Consider a function $f:\mathbb{R}^n \to \mathbb{R}^m$, such that $f(\mathbf{x}) = f(x_1,x_2,\ldots,x_n) = (f_1(\mathbf{x}),f_2(\mathbf{x}),\ldots,f_m(\mathbf{x}))$. The functions $f_i:\mathbb{R}^n \to \mathbb{R}$ are component functions. In this section f_i are differentiable functions.

Definition (Jacobian matrix)

The **Jacobian matrix** of a function $f:\mathbb{R}^n \to \mathbb{R}^m$ at $\mathbf{a} = (a_1,a_2,\ldots,a_n)$ is defined as the $m\times n$ matrix:

$$Df(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

where all partial derivatives are evaluated at **a**

Definition (Differentiation for Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$)

Let $f:\mathbb{R}^n \to \mathbb{R}^m$ be a function that maps vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$. If $\mathbf{x} = (x_1,x_2,\ldots,x_n)$ is a vector in $\mathbb{R}^n$, and $f(\mathbf{x}) = (f_1(\mathbf{x}),f_2(\mathbf{x}),\ldots,f_m(\mathbf{x}))$, then the derivative of f at a point $\mathbf{a}\in\mathbb{R}^n$ is the matrix of partial derivatives, also known as the *Jacobian matrix*, denoted by $Df(\mathbf{a})$.

Example

Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by:

$$f(x,y) = \begin{bmatrix} x^2 + y \\ xy \end{bmatrix}.$$

The Jacobian matrix of f is:

$$Df(x,y) = \begin{bmatrix} 2x & 1 \\ y & x \end{bmatrix}.$$

The notion of differentiability as approximation by linear maps is extended naturally to vector functions.

Definition (Differentiable Vector Function)

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *differentiable* at $\mathbf{a} \in \mathbb{R}^n$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$, represented by the Jacobian matrix $Df(\mathbf{a})$, such that:

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

In practice we must check for differentiability in each component.

Theorem (Differentiability Criteria)

Suppose that S is an open subset of $\mathbb{R}^n$. Then a function $\mathbf{f} : S \to \mathbb{R}^m$ is differentiable at a point $\mathbf{a} \in S$ if and only if the component functions f_j are differentiable at $\mathbf{a}$ for every $j \in 1, \ldots, m$. Moreover,

$$D\mathbf{f}(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

(where all the partial derivatives are evaluated at a.) Furthermore, if all partial derivatives $\frac{\partial f_j}{\partial x_i}$ (for $i = 1, \ldots, n$ and $j = 1, \ldots, m$) exist and are continuous in S, then $\mathbf{f}$ is differentiable in S.

Proposition

If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\mathbf{a} \in \mathbb{R}^n$, then f is continuous at $\mathbf{a}$.

Proof. Since f is differentiable at $\mathbf{a}$, we have:

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

Therefore, as $\mathbf{h} \to 0$, the term $Df(\mathbf{a})\mathbf{h}$ approaches 0, implying that:

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}),$$

which proves the continuity of f at $\mathbf{a}$. $\square$

Example

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = (x^2 + y^2, x)$. The Jacobian matrix $Df(x,y)$ is:

$$Df(x,y) = \begin{bmatrix} 2x & 2y \\ 1 & 0 \end{bmatrix}.$$

Note that the component functions $f_1(x,y)$ and $f_2(x,y) = x$ are differentiable, therefore, the function f is differentiable.

Example

Consider the function $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$\mathbf{f}(x,y,z) = \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \end{pmatrix} = \begin{pmatrix} |x| + z \\ |y-1| + xz \end{pmatrix}$$

Note that f_1 is not differentiable when $x = 0$, and f_2 is not differentiable when $y = 1$. Taking these into account, the matrix of partial derivatives at a point (x,y,z) is given by

$$\begin{pmatrix} \frac{x}{|x|} & 0 & 1 \\ z & \frac{y-1}{|y-1|} & x \end{pmatrix} \quad \text{if } x \neq 0 \text{ and } y \neq 1$$

The entries of this matrix are all continuous everywhere in

$$S = \left\{ (x,y,z) \in \mathbb{R}^3 : x \neq 0 \text{ and } y \neq 1 \right\}$$

which is an open set, so we conclude that $\mathbf{f}$ is differentiable everywhere in this set, and that the derivative $D\mathbf{f}(x,y,z)$ is given by the above matrix.

The Chain Rule for Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

The chain rule extends naturally to functions that map between spaces of different dimensions. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a differentiable function, and $g : \mathbb{R}^p \to \mathbb{R}^n$ is also differentiable, then the composition $f \circ g : \mathbb{R}^p \to \mathbb{R}^m$ is differentiable, and its derivative is given by:

$$D(f \circ g)(\mathbf{a}) = Df(g(\mathbf{a}))Dg(\mathbf{a}).$$

Theorem (Multivariable Chain Rule)

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^p \to \mathbb{R}^n$ be differentiable at points $g(\mathbf{a})$ and $\mathbf{a}$, respectively. Then the composition $f \circ g : \mathbb{R}^p \to \mathbb{R}^m$ is differentiable at $\mathbf{a}$, and:

$$D(f \circ g)(\mathbf{a}) = Df(g(\mathbf{a}))Dg(\mathbf{a}).$$

Proof. The derivative of the composition $f(g(\mathbf{a}))$ can be computed using the linear maps corresponding to Df and Dg. By the definition of differentiability, the total change in $f(g(\mathbf{a}+\mathbf{h}))$ can be approximated by:

$$f(g(\mathbf{a}+\mathbf{h})) - f(g(\mathbf{a})) \approx Df(g(\mathbf{a}))\,(Dg(\mathbf{a})\mathbf{h})\,,$$

which shows that the derivative of the composition is the product of the two Jacobian matrices. □

Example

Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by:

$$f(x,y) = \begin{bmatrix} x^2 \\ y^2 \end{bmatrix},$$

and let $g : \mathbb{R} \to \mathbb{R}^2$ be defined by:

$$g(t) = \begin{bmatrix} t \\ \sin(t) \end{bmatrix}.$$

The Jacobian matrices are:

$$Df(x,y) = \begin{bmatrix} 2x & 0 \\ 0 & 2y \end{bmatrix}, \quad Dg(t) = \begin{bmatrix} 1 \\ \cos(t) \end{bmatrix}.$$

The derivative of the composition $f \circ g$ is:

$$D(f \circ g)(t) = \begin{bmatrix} 2t & 0 \\ 0 & 2\sin(t) \end{bmatrix} \begin{bmatrix} 1 \\ \cos(t) \end{bmatrix} = \begin{bmatrix} 2t \\ 2\sin(t)\cos(t) \end{bmatrix}.$$

Properties of Differentiation for Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

- **Linearity:** The derivative of a sum of functions is the sum of their derivatives.

$$D(f+g)(\mathbf{a}) = Df(\mathbf{a}) + Dg(\mathbf{a}).$$

- **Product Rule:** If $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$, then:

$$D(fg)(\mathbf{a}) = f(\mathbf{a})Dg(\mathbf{a}) + g(\mathbf{a})Df(\mathbf{a}).$$

Chain Rule: For compositions of differentiable functions f and g, the chain rule applies as shown in the previous theorem.

Example

Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $f(x,y,z) = \begin{bmatrix} x^2+y^2 \\ y+z^2 \end{bmatrix}$, and let $g : \mathbb{R} \to \mathbb{R}^3$ be defined by $g(t) = \begin{bmatrix} t \\ \sin(t) \\ \cos(t) \end{bmatrix}$. The derivative of the composition $f \circ g$ can be calculated by applying the Jacobian matrices for both f and g.

§II. Differentiation in Several Variables

- Extrema of Functions and Optimization
  - Differentials and Taylor's Theorem
  - Extrema of Functions
  - Lagrange Multipliers: Finding Constrained Extrema

Differentials provide a linear approximation of changes in a function based on the function's derivative.

Definition

Differentials Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function at $\mathbf{x} \in \mathbb{R}^n$. The *differential* of f at $\mathbf{x}$, denoted by df, is given by:

$$df(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot d\mathbf{x},$$

where $\nabla f(\mathbf{x})$ is the gradient of f at $\mathbf{x}$, and $d\mathbf{x}$ represents a small change in $\mathbf{x}$.

Remark

The differential $df(\mathbf{x})$ represents the best linear approximation to the change in f near $\mathbf{x}$. For small changes in $\mathbf{x}$, the approximation becomes more accurate.

Taylor's Theorem

Taylor's Theorem allows us to approximate a function around a point using a polynomial expansion.

Theorem (Taylor's Theorem)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. The Taylor expansion of f around $\mathbf{x}_0 \in \mathbb{R}^n$ up to order k is:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \cdots + R_k(\mathbf{x}),$$

where $H_f(\mathbf{x}_0)$ is the Hessian matrix of second partial derivatives of f at $\mathbf{x}_0$, and $R_k(\mathbf{x})$ is the remainder term.

Example

Consider $f(x) = e^x$. The second-order Taylor expansion of $f(x)$ around $x_0 = 0$ is:

$$f(x) = 1 + x + \frac{x^2}{2} + O(x^3).$$

This polynomial provides an approximation to e^x near $x = 0$.

Theorem

Second-Order Taylor Formula Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ have continuous partial derivatives of third order. [2] *Then we may write*

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{x}_0) + \frac{1}{2} \sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) + R_2(\mathbf{x}_0, \mathbf{h}),$$

where $R_2(\mathbf{x}_0, \mathbf{h}) / \|\mathbf{h}\|^2 \to 0$ as $\mathbf{h} \to \mathbf{0}$ and the second sum is over all i 's and j 's between 1 and n (so there are n^2 terms).

Example

To find the first- and second-order Taylor approximations to $f(x,y) = \sin(xy)$ at the point $(x_0, y_0) = (1, \pi/2)$ compute

$$f(x_0,y_0) = \sin(x_0 y_0) = \sin(\pi/2) = 1$$
$$f_x(x_0,y_0) = y_0 \cos(x_0 y_0) = \frac{\pi}{2}\cos(\pi/2) = 0$$
$$f_y(x_0,y_0) = x_0 \cos(x_0 y_0) = \cos(\pi/2) = 0$$

$$f_{xx}(x_0,y_0) = -y_0^2 \sin(x_0 y_0) = -\frac{\pi^2}{4}\sin(\pi/2) = -\frac{\pi^2}{4}$$
$$f_{xy}(x_0,y_0) = \cos(x_0 y_0) - x_0 y_0 \sin(x_0 y_0) = -\frac{\pi}{2}$$
$$f_{yy}(x_0,y_0) = -x_0^2 \sin(x_0 y_0) = -\sin(\pi/2) = -1$$

Example

Thus, the linear (first-order) approximation is

$$\begin{aligned} l(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &= 1+0+0 = 1 \end{aligned}$$

and the second-order (or quadratic) approximation is

$$\begin{aligned} g(x,y) =& 1+0+0+\frac{1}{2}\left(-\frac{\pi^2}{4}\right)(x-1)^2 + \left(-\frac{\pi}{2}\right)(x-1)\left(y-\frac{\pi}{2}\right) \\ &+ \frac{1}{2}(-1)\left(y-\frac{\pi}{2}\right)^2 \\ =& 1 - \frac{\pi^2}{8}(x-1)^2 - \frac{\pi}{2}(x-1)\left(y-\frac{\pi}{2}\right) - \frac{1}{2}\left(y-\frac{\pi}{2}\right)^2 \end{aligned}$$

Remark

Higher-order terms in the Taylor expansion improve the accuracy of the approximation, particularly for larger deviations from the point of expansion. Here we only focus to approximations of at most order two.

Definition (Local Maximum)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and let $\mathbf{x}_0 \in \mathbb{R}^n$ be a point in the domain of f. We say that f has a *local maximum* at $\mathbf{x}_0$ if there exists a neighborhood U of $\mathbf{x}_0$ such that:

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{for all} \quad \mathbf{x} \in U.$$

In other words, $f(\mathbf{x}_0)$ is greater than or equal to $f(\mathbf{x})$ for all $\mathbf{x}$ near $\mathbf{x}_0$.

Definition (Local Minimum)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and let $\mathbf{x}_0 \in \mathbb{R}^n$ be a point in the domain of f. We say that f has a *local minimum* at $\mathbf{x}_0$ if there exists a neighborhood U of $\mathbf{x}_0$ such that:

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) \quad \text{for all} \quad \mathbf{x} \in U.$$

In other words, $f(\mathbf{x}_0)$ is less than or equal to $f(\mathbf{x})$ for all $\mathbf{x}$ near $\mathbf{x}_0$.

URJC

Remark

A point $\mathbf{x}_0$ where f has a local maximum or minimum is called a *critical point*. These points often occur where the gradient $\nabla f(\mathbf{x}_0)$ is zero, but not all critical points are local extrema. To find local maxima or minima, we first locate the critical points, where the gradient vanishes.

Definition (Critical Point)

A point $\mathbf{x}_0 \in \mathbb{R}^n$ is called a *critical point* of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ if $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

Remark

Critical points are candidates for local extrema but do not necessarily guarantee maxima or minima.

URJC

Definition (Hessian Matrix)

Suppose $f : U \subset \mathbb{R}^n$ has second-order continuous derivatives $(\partial^2/\partial x_i \partial x_j)(\mathbf{x}_0)$ for $i, j = 1, \ldots, n$ at a point $\mathbf{x}_0 \in U$. The **Hessian matrix** of f at $\mathbf{x}_0$ is defined as

$$\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \frac{\partial^2 f}{\partial x_i \partial x_j} & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Theorem (Second Derivative Test)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function, and let $\mathbf{x}_0$ be a critical point. Consider the Hessian matrix $H_f(\mathbf{x}_0)$:

- *If $H_f(\mathbf{x}_0)$ is positive definite, then $\mathbf{x}_0$ is a local minimum.*
- *If $H_f(\mathbf{x}_0)$ is negative definite, then $\mathbf{x}_0$ is a local maximum.*
- *If $H_f(\mathbf{x}_0)$ has both positive and negative eigenvalues, then $\mathbf{x}_0$ is a saddle point.*

  If $|H_f(\mathbf{x}_0)| = 0$ we are in the degenerate case, and nothing can be said about the functions from the Hessian matrix

Definition (Leading Minor)

Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. The *k-leading minor* of A, denoted Δ_k, is the determinant of the $k \times k$ leading principal submatrix of A. Specifically, it is the submatrix formed by the first k rows and the first k columns of A.

More formally, if $A = [a_{ij}]$ is the matrix, then the k-leading principal submatrix A_k is:

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}.$$

The k-leading minor Δ_k is then given by:

$$\Delta_k = \det(A_k).$$

Definition (Positive Definite Matrix)

A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is *positive definite* if and only if all its principal minors are positive. That is, for each k (where $1 \le k \le n$), the leading principal minors Δ_k of A satisfy:

$$\Delta_k > 0.$$

Here, Δ_k is the determinant of the $k \times k$ leading principal submatrix of A.

Definition (Negative Definite Matrix)

A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is *negative definite* if and only if the signs of the principal minors alternate, starting with a negative sign for the first principal minor. Specifically, for each k (where $1 \le k \le n$), the leading principal minors Δ_k satisfy:

$$\operatorname{sgn} \Delta_k = (-1)^{k+1}$$

where sgn stands for the function sign

$$\operatorname{sgn}(x) = \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 0. \end{cases}$$

Remark

The signs of the principal minors provide a convenient criterion for definiteness of symmetric matrices. Positive definiteness is indicated by **all positive principal minors**, while negative definiteness is indicated by **alternating signs of principal minors**, starting with a negative sign.

Example

Consider the function $f(x,y,z) = x^2 + y^2 + z^2$. The gradient is:

$$\nabla f(x,y) = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}.$$

The only critical point is $(0,0,0)$. The Hessian matrix is:

$$H_f(0,0,0) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

The determinant of the leading minors are

$$\Delta_1 = |2| = 2 > 0 \qquad \Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0 \qquad \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 8 > 0$$

Therefore, the Hessian matrix is definite positive and $(0,0,0)$ is a local minimum.

Theorem (Second-Derivative Maximum-Minimum Test for Functions of Two Variables)

Let $f(x,y)$ be of class C^2 on an open set U in $\mathbb{R}^2$. A point (x_0,y_0) is a (strict) local minimum of f provided the following three conditions hold:

(i) $\frac{\partial f}{\partial x}(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = 0$

(ii) $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) > 0$

(iii) $D = \left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0$ *at* (x_0,y_0)

(D is called the discriminant of the Hessian.) If in (ii) we have < 0 instead of > 0 and condition (iii) is unchanged, then we have a (strict) local maximum. If $D < 0$ (e.g., if $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) = 0$ or $\frac{\partial^2 f}{\partial y^2}(x_0,y_0) = 0$, but $\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) \neq 0$), then (x_0,y_0) is of saddle type (neither a maximum nor a minimum).

Example

Locate the relative maxima, minima, and saddle points of the function

$$f(x,y) = \log\left(x^2+y^2+1\right)$$

We must first locate the critical points of this function; therefore, according to Theorem 3 , we calculate

$$\nabla f(x,y) = \frac{2x}{x^2+y^2+1}\mathbf{i} + \frac{2y}{x^2+y^2+1}\mathbf{j}$$

Thus, $\nabla f(x,y) = \mathbf{0}$ if and only if $(x,y) = (0,0)$, and so the only critical point of f is $(0,0)$. Now we must determine whether this is a maximum, a minimum, or a saddle point. The second partial derivatives are

$$\frac{\partial^2 f}{\partial x^2} = \frac{2\left(x^2+y^2+1\right)-(2x)(2x)}{\left(x^2+y^2+1\right)^2}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{2\left(x^2+y^2+1\right)-(2y)(2y)}{\left(x^2+y^2+1\right)^2}$$

Example

On the other hand

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{-2x(2y)}{(x^2+y^2+1)^2}$$

Therefore,

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2 = \frac{\partial^2 f}{\partial y^2}(0,0) \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 0$$

which yields

$$D = 2 \cdot 2 = 4 > 0$$

Because $\left(\partial^2 f / \partial x^2\right)(0,0) > 0$, we conclude that $(0,0)$ is a local minimum. (Can you show this just from the fact that $\log t$ is an increasing function of $t > 0$?)

Lagrange Multipliers: Finding Constrained Extrema

URJC

Lagrange multipliers are a powerful method for finding the extrema (maximum or minimum) of a function subject to constraints. This technique is especially useful in optimization problems where direct methods are not feasible due to the constraints.

Definition (Lagrange Multipliers)

Let $f(x,y,z,\dots)$ be a function of several variables, and let the constraint be given by $g(x,y,z,\dots)=0$. The method of **Lagrange multipliers** seeks to find the extrema of f subject to the constraint by introducing a scalar λ, called the **Lagrange multiplier**, such that:

$$\nabla f = \lambda \nabla g$$

This equation implies that the gradient of the objective function f is parallel to the gradient of the constraint g. The parameter λ adjusts the magnitudes of the gradients while ensuring their directions are aligned.

Proposition (Constrained Optimization Problem)

Let $f(x,y,z)$ be a smooth function, and let the constraint $g(x,y,z)=0$ define a smooth surface. To find the extrema of f subject to the constraint, solve the system of equations:

$$\nabla f = \lambda \nabla g \quad \text{and} \quad g(x,y,z)=0$$

Proof. If f has an extremum on the constraint surface defined by $g(x,y,z) = 0$, the gradient ∇f must be perpendicular to the surface at that point. The gradient of the constraint ∇g also defines the normal direction of the surface. Hence, ∇f and ∇g must be parallel, which leads to the equation $\nabla f = \lambda \nabla g$. □

Proposition (Properties of Lagrange Multipliers)

If f has a local extremum subject to the constraint $g(x,y,z) = 0$, then there exists a scalar λ such that $\nabla f = \lambda \nabla g$.

This result provides a necessary condition for constrained optimization problems.

Remark

The method of Lagrange multipliers does not distinguish between maxima and minima. After solving for x, y, z, one must check the second derivative or use other techniques to determine the nature of the extremum.

Example

Find the extrema of the function $f(x,y) = x^2 + y^2$ subject to the constraint $g(x,y) = x + y - 1 = 0$.
Solution: 1. The gradients of f and g are:

$$\nabla f = (2x, 2y), \quad \nabla g = (1, 1)$$

2. Using the Lagrange multiplier condition $\nabla f = \lambda \nabla g$, we get:

$$(2x, 2y) = \lambda (1, 1)$$

This gives two equations:

$$2x = \lambda, \quad 2y = \lambda$$

Thus, $x = y$.

Example

3. Using the constraint $x+y=1$, substitute $y=x$:

$$x+x=1 \quad \Rightarrow \quad x=\frac{1}{2}, \quad y=\frac{1}{2}$$

4. The extrema occur at $x=\frac{1}{2}, y=\frac{1}{2}$. Substituting into $f(x,y)$:

$$f\left(\frac{1}{2},\frac{1}{2}\right)=\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2=\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$$

Thus, the extremum of $f(x,y)=x^2+y^2$ subject to the constraint $x+y=1$ is $\frac{1}{2}$.

Multiple Constraints

Proposition

For a function $f(x_1, x_2, \ldots, x_n)$ with multiple constraints $g_1(x_1, x_2, \ldots, x_n) = 0$ and $g_2(x_1, x_2, \ldots, x_n) = 0$, the method of Lagrange multipliers extends as follows:

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$$

Example

Find the extrema of $f(x,y,z) = x^2 + y^2 + z^2$ subject to two constraints:

$$g_1(x,y,z) = x + y + z - 1 = 0, \quad g_2(x,y,z) = x - y = 0$$

Solution: 1. Compute the gradients:

$$\nabla f = (2x, 2y, 2z), \quad \nabla g_1 = (1,1,1), \quad \nabla g_2 = (1,-1,0)$$

2. The Lagrange multiplier condition is:

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$$

Example

This gives the system of equations:

$$2x = \lambda_1 + \lambda_2, \quad 2y = \lambda_1 - \lambda_2, \quad 2z = \lambda_1$$

3. Use the constraints $x+y+z=1$ and $x-y=0$. From $x-y=0$, we get $x=y$. Substituting into $x+y+z=1$, we get:

$$2x+z=1$$

Substituting $z=x$, we find $3x=1$, so $x=y=z=\frac{1}{3}$. Thus, the extremum occurs at $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$, and the value of the function is:

$$f\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) = \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{1}{3}$$

The method of Lagrange multipliers provides a systematic way to find extrema of functions subject to constraints. By transforming the problem into one involving gradients and multipliers, we can solve for the extrema while satisfying the given constraints.

Theorem (Second-Derivative test for constrained optimization)

Let $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ and $g : U \subset \mathbb{R}^2 \to \mathbb{R}$ be smooth (at least C^2) functions. Let $\mathbf{v}_0 \in U, g(\mathbf{v}_0) = c$, and S be the level curve for g with value c. Assume that $\nabla g(\mathbf{v}_0) \neq \mathbf{0}$ and that there is a real number λ such that $\nabla f(\mathbf{v}_0) = \lambda \nabla g(\mathbf{v}_0)$. Form the auxiliary function $h = f - \lambda g$ and the bordered Hessian determinant

$$|\overline{H}| = \begin{vmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 h}{\partial x \partial y} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 h}{\partial x \partial y} & \frac{\partial^2 h}{\partial y^2} \end{vmatrix} \quad \text{evaluated at } \mathbf{v}_0$$

(i) If $|\bar{H}| > 0$, then $\mathbf{v}_0$ is a local maximum point for $f \mid S$.

(ii) If $|\bar{H}| < 0$, then $\mathbf{v}_0$ is a local minimum point for $f \mid S$.

(iii) If $|\bar{H}| = 0$, the test is inconclusive and $\mathbf{v}_0$ may be a minimum, a maximum, or neither.

Remark

This result is an application of the second derivative test studied before to the lagrange function. Furthemore, it can be extended for the case of more that two variables and several restrictions.

URJC

Theorem

Optimization on Compact Sets Let K be a compact subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ a continuous function on K. Then, there are $x_m, x_M \in K$ such that $f(x_m) = \min_K f$ and $f(x_m) = \min_K f$, i.e. f reach its maximum and minimum values on K.

Steps for General Optimization on Compact Sets

Let f be a differentiable function on a closed and bounded region $D = U \cup \partial U, U$ open in $\mathbb{R}^n$, with smooth boundary ∂U. To find the absolute maximum and minimum of f on D :

(i) Locate all critical points of f in U.

(ii) Use the method of Lagrange multiplier to locate all the critical points of $f \mid \partial U$.

(iii) Compute the values of f at all these critical points.

(iv) Select the largest and the smallest.

Example

Find the absolute maximum and minimum of the function $f(x,y,z) = x^2 + y^2 + z^2 - x + y$ on the set
$D = \{(x,y,z) \mid x^2 + y^2 + z^2 \leq 1\}$
As in the previous examples, we know the absolute maximum and minimum exists. Now $D = U \cup \partial U$, where

$$U = \{(x,y,z) \mid x^2 + y^2 + z^2 < 1\}$$

and

$$\partial U = \{(x,y,z) \mid x^2 + y^2 + z^2 = 1\}$$
$$\nabla f(x,y,z) = (2x-1, 2y+1, 2z)$$

Thus, $\nabla f = 0$ at $(1/2, -1/2, 0)$ which is in U, the interior of D. Let $g(x,y,z) = x^2 + y^2 + z^2$. Then ∂U is the level set $g(x,y,z) = 1$. By the method of Lagrange multipliers, the maximum and minimum must occur at a critical point of $f \mid \partial U$; that is, at a point $\mathbf{x}_0$ where $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$ for some scalar λ.

URJC

Example

Thus, $(2x-1, 2y+1, 2z) = \lambda(2x, 2y, 2z)$ or

(i) $2x-1 = 2\lambda x$

(ii) $2y+1 = 2\lambda y$

(iii) $2z = 2\lambda z$

If $\lambda = 1$, then we would have $2x-1 = 2x$ or $-1 = 0$, which is impossible. We may assume that $\lambda \neq 0$ since if $\lambda = 0$, we only get an interior point as above. Thus (iii) implies that $z = 0$ and
(iv) $x^2 + y^2 = 1$
Solving (i) and (ii) for x and y we find,

(v) $x = 1/2(1-\lambda)$

(vi) $y = -1/2(1-\lambda)$

Applying (iv) we can solve for λ, namely $\lambda = 1 \pm (1/\sqrt{2})$. Thus, from (v) and (vi) we have that $x = \pm(1/\sqrt{2})$ and $y = \pm(1/\sqrt{2})$; that is, we have four critical points on ∂U. Evaluating f at each of these points, we see that the maximum value for f on ∂U is $1 + 2/\sqrt{2} = 1+\sqrt{2}$ and the minimum value is $1-\sqrt{2}$. The value of f at $(1/2, -1/2)$ is $-1/2$. Comparing these values, noting that $-1/2 < 1-\sqrt{2}$, we see that the absolute minimum is $-1/2$, occurring at $(1/2, -1/2)$, and that absolute maximum is $1+\sqrt{2}$, occurring at $(-1/\sqrt{2}, 1/\sqrt{2})$.

§II. Integration in Several Variables

- Integration in Several Variables
  - Double Integrals.
  - Changing the Order of Integration.
  - Triple Integrals.
  - Jacobians. Transformation of Integrals.
  - Areas and Volumes. Applications of Integration

Double and triple integrals are essential tools for calculating areas, volumes, and other physical properties of regions in two and three dimensions.

Definition

Double Integrals over Rectangular Regions Let $f(x,y)$ be a continuous function on a rectangular region $R = [a,b] \times [c,d]$. The double integral of f over R is given by:

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

In rectangular regions, the limits of integration are constants, making the evaluation of the integral straightforward.

Example

Evaluate the double integral $\iint_R x^2 + y^2 \, dA$, where $R = [0,1] \times [0,2]$.
Solution: We evaluate the integral as follows:

$$\iint_R (x^2 + y^2) \, dA = \int_0^1 \int_0^2 (x^2 + y^2) \, dy \, dx$$

First, integrate with respect to y:

$$\int_0^2 (x^2 + y^2) \, dy = x^2 y + \frac{y^3}{3} \Big|_0^2 = 2x^2 + \frac{8}{3}$$

Now, integrate with respect to x:

$$\int_0^1 \left(2x^2 + \frac{8}{3} \right) dx = \frac{2x^3}{3} + \frac{8x}{3} \Big|_0^1 = \frac{2}{3} + \frac{8}{3} = \frac{10}{3}$$

Thus, the value of the double integral is $\frac{10}{3}$.

When the region R is not rectangular, the limits of integration may depend on one of the variables, and we must carefully choose the order of integration.

Definition (y-simple region)

Let $f(x,y)$ be continuous on a general region R bounded by curves. If R is defined by $a \leq x \leq b$ and $\alpha(x) \leq y \leq \beta(x)$, the double integral is:

$$\iint_R f(x,y)\,dA = \int_a^b \int_{\alpha(x)}^{\beta(x)} f(x,y)\,dy\,dx$$

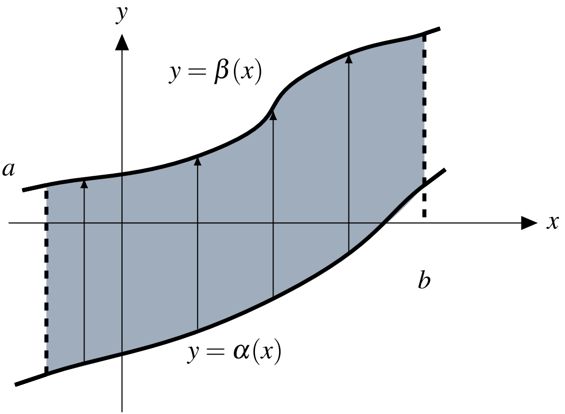


Figure: x- simple region

Example

Find the area of the region R bounded by $y = x^2$ and $y = 1$.
The region is bounded by $y = x^2$ and $y = 1$, and the limits of x are from -1 to 1. The area is given by the double integral:

$$\text{Area}(R) = \int_{-1}^{1}\int_{x^2}^{1} 1\,dy\,dx$$

First, integrate with respect to y:

$$\int_{x^2}^{1} 1\,dy = 1 - x^2$$

Now, integrate with respect to x:

$$\int_{-1}^{1}(1-x^2)\,dx = \left[x - \frac{x^3}{3}\right]_{-1}^{1} = \left(1-\frac{1}{3}\right) - \left(-1+\frac{1}{3}\right) = \frac{4}{3}$$

Thus, the area of the region is $\frac{4}{3}$.

Definition (x-simple region)

Let $f(x,y)$ be continuous on a general region R bounded by curves. If R is defined by $c \leq y \leq d$ and $\gamma(y) \leq x \leq \delta(y)$, the double integral is:

$$\iint_R f(x,y)\,dA = \int_c^d \int_{\gamma(y)}^{\delta(x)} f(x,y)\,dx\,dy$$

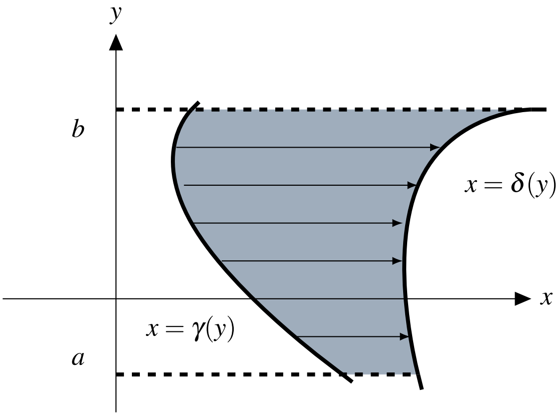


Figure: y- simple region

A region that is x-simple and y-simple simultaneously is called **simple region**.

Example

Let's consider the region R bounded by $x = y$ and $x = y^2$, with y ranging from 0 to 1. We want to integrate a function $f(x,y) = y$ over this region.

The integral can be written as $\int_R f(x,y)\,dA = \int_0^1 \left(\int_{y^2}^{y} f(x,y)\,dx \right) dy$

Set $f(x,y) = y$, the integral simplifies to $\int_0^1 \left(\int_{y^2}^{y} y\,dx \right) dy$

Evaluating the inner integral: $\int_{y^2}^{y} dx = y - y^2$

So the integral becomes $\int_0^1 y(y - y^2)\,dy = \frac{1}{12}$

Remark

1. Double integrals are useful for calculating areas, masses, and centroids in two-dimensional regions, while triple integrals extend these ideas to three-dimensional regions, allowing for volume and mass calculations.
2. There are regions that can be written as y-simple and x-simple at the same time, for example the semi-circle, in that case we may choose the simplest way
3. The evaluation of double and triple integrals can be simplified by changing the order of integration or transforming to polar, cylindrical, or spherical coordinates, depending on the symmetry of the region.

Changing the Order of Integration

In this section we present how to change limits of integrations in two variables. This result is naturally extended to more variables in an straightforward way.

Definition

Let $f(x,y)$ be a continuous function defined on a rectangular region $R = [a,b] \times [c,d] \subset \mathbb{R}^2$. A *double integral* is the integral of a function of two variables over this region, denoted by

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

Theorem (Fubini's Theorem)

Let $f(x,y)$ be a continuous function on the rectangular region $R = [a,b] \times [c,d]$. Then, the double integral of $f(x,y)$ over R can be computed as an iterated integral in either order:

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

Remark

Fubini's Theorem allows us to compute a double integral by performing two single-variable integrals in succession. The choice of the order of integration depends on the limits of the region of integration and may simplify the calculation. This theorem is extended in a natural way to more variables.

Proposition

Let R be a **simple region**, i.e.,

$$R = \{(x,y) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\} = \{(x,y) \mid c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\},$$

is a region that can be written as y and x-simple, then Fubini's Theorem applies:

$$\iint_R f(x,y)\, dA = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\, dy \right) dx = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\, dx \right) dy.$$

Remark

When the region R is non-rectangular, changing the order of integration requires careful analysis of the bounds. The limits of integration for the inner integral must be adjusted accordingly.

Example

Consider the function $f(x,y) = xy$ over the triangular region R defined by $0 \leq x \leq 1$ and $0 \leq y \leq x$. The double integral is:

$$\iint_R xy\, dA = \int_0^1 \left(\int_0^x xy\, dy \right) dx.$$

First, integrate with respect to y:

$$\int_0^x xy\, dy = x \cdot \frac{y^2}{2}\Big|_0^x = x \cdot \frac{x^2}{2} = \frac{x^3}{2}.$$

Now, integrate with respect to x:

$$\int_0^1 \frac{x^3}{2}\, dx = \frac{1}{2} \int_0^1 x^3\, dx = \frac{1}{2} \cdot \frac{x^4}{4}\Big|_0^1 = \frac{1}{8}.$$

Thus, the value of the integral is $\frac{1}{8}$.

Example

Let us reverse the order now. Note that the triangular region R can be written also as

$$R = \{(x,y) \in \mathbb{R}^2 : 0 \leq y \leq 1; y \leq x \leq 1\}$$

Therefore, according to Fubini's theorem we have

$$\iint_R xy\,dA = \int_0^1 \left(\int_0^x xy\,dy \right) dx = \int_0^1 \left(\int_x^1 xy\,dx \right) dy$$

The remaining calculation is left as exercise.

Example

Let $f(x,y) = x+y$ over the region R bounded by $0 \leq y \leq 1$ and $y \leq x \leq 1$. This region can be written as:

$$\iint_R (x+y)\,dA = \int_0^1 \left(\int_y^1 (x+y)\,dx \right) dy.$$

First, integrate with respect to x:

$$\int_y^1 (x+y)\,dx = \left(\frac{x^2}{2} + yx \right) \Big|_y^1 = \left(\frac{1}{2} + y \right) - \left(\frac{y^2}{2} + y^2 \right).$$

Simplifying this:

$$= \frac{1}{2} + y - \frac{y^2}{2} - y^2 = \frac{1}{2} + y - \frac{3y^2}{2}.$$

URJC

Example

Now, integrate with respect to y:

$$\int_0^1 \left(\frac{1}{2} + y - \frac{3y^2}{2}\right) dy = \frac{y}{2} + \frac{y^2}{2} - \frac{y^3}{2}\Big|_0^1 = \frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2}.$$

Therefore, the value of the integral is $\frac{1}{2}$.

Remark

Changing the order of integration can greatly simplify the computation in certain cases. The region of integration may have a simpler description when switching the roles of the variables.

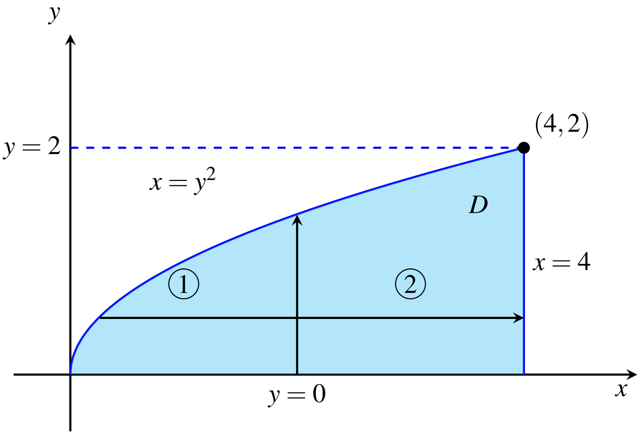
Example

Sometimes changing the order of integration can make an impossible calculation possible.
Consider the evaluation of the following iterated integral

$$\int_0^2 \int_{y^2}^4 y\cos\left(x^2\right) dxdy$$

Change the order of integration in two steps:

1. Use the limits of integration in the original iterated integral to identify the region D in $\mathbb{R}^2$
2. Assuming that the region D in Step 1 is a **simple region**, change the order of integration



Example

The limits of integration in this particular case imply that D can be described as

$$D = \left\{(x,y) \mid y^2 \leq x \leq 4, 0 \leq y \leq 2\right\}$$

Note that $x = y^2$ corresponds to $y = \sqrt{x}$ over this region. Use this information to change the order of integration

$$\int_0^2 \int_{y^2}^4 y\cos\left(x^2\right) dxdy = \int_0^4 \int_0^{\sqrt{x}} y\cos\left(x^2\right) dydx$$

It is now possible to complete the calculation

$$\int_0^2 \int_{y^2}^4 y\cos\left(x^2\right) dxdy = \int_0^4 \int_0^{\sqrt{x}} y\cos\left(x^2\right) dydx = \int_0^4 \left(\frac{y^2}{2}\cos\left(x^2\right)\Big|_{y=0}^{y=\sqrt{x}} \right) dx = \int_0^4 \frac{x}{2}\cos\left(x^2\right) dx$$

$$u = x^2 \Rightarrow du = 2xdx$$

$$= \frac{1}{4}\int_0^{16} \cos u du = \frac{1}{4}\sin u\Big|_0^{16} = \frac{1}{4}\sin 16$$

Definition (Triple Integral)

A triple integral is an integral of a function of three variables over a three-dimensional region. It is denoted as

$$\iiint_D f(x,y,z)\,dV,$$

where D is the region in $\mathbb{R}^3$ over which we are integrating, and dV represents the volume element in Cartesian coordinates.

Definition (Rectangular Box)

A **rectangular box** R is a subset of $\mathbb{R}^3$ such that

$$R = \{(x,y,z) \in \mathbb{R}^2 : a \leq x \leq b; c \leq y \leq d; e \leq z \leq f\}$$

Theorem

If $f(x,y,z)$ is continuous on a region D and D is a rectangular box defined by $a \leq x \leq b$, $c \leq y \leq d$, and $e \leq z \leq f$, then

$$\iiint_D f(x,y,z)\,dV = \int_a^b \int_c^d \int_e^f f(x,y,z)\,dz\,dy\,dx.$$

Proposition

The order of integration in a triple integral can be changed as long as the limits of integration are appropriately adjusted. For example, if D is a region bounded by x from a to b, y from $g_1(x)$ to $g_2(x)$, and z from $h_1(x,y)$ to $h_2(x,y)$, then

$$\iiint_D f(x,y,z)\,dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)\,dz\,dy\,dx.$$

Example

Consider the function $f(x,y,z) = x^2 + y^2 + z^2$ over the rectangular box with limits $0 \leq x \leq 1$, $0 \leq y \leq 2$, and $0 \leq z \leq 3$. The triple integral is

$$\iiint_D (x^2 + y^2 + z^2)\, dV.$$

Evaluating this, we have:

$$\int_0^1 \int_0^2 \int_0^3 (x^2 + y^2 + z^2)\, dz\, dy\, dx.$$

First integrate with respect to z:

$$\int_0^1 \int_0^2 \left[x^2 z + y^2 z + \frac{z^3}{3}\right]_0^3 dy\, dx = \int_0^1 \int_0^2 \left(3x^2 + 3y^2 + 9\right) dy\, dx.$$

Next, integrate with respect to y:

$$\int_0^1 \left[3x^2 y + 3\frac{y^3}{3} + 9y\right]_0^2 dx = \int_0^1 \left(6x^2 + 8 + 18\right) dx = \int_0^1 \left(6x^2 + 26\right) dx.$$

Finally, integrate with respect to x:

$$\left[2x^3 + 26x\right]_0^1 = 2 + 26 = 28.$$

Definition (Cylindrical region)

A region D is called **cylindrical** if it can be described in cylindrical coordinates (r,θ,z) as

$$D=\{(r,\theta,z)\mid 0\leq r\leq R,\,0\leq\theta\leq 2\pi,\,h_1(r)\leq z\leq h_2(r)\}.$$

Definition (Spherical region)

A region D is called **spherical** if it can be described in spherical coordinates (ρ,θ,ϕ) as

$$D=\{(\rho,\theta,\phi)\mid 0\leq\rho\leq\rho_0,\,0\leq\theta\leq 2\pi,\,0\leq\phi\leq\pi\}.$$

Remark

To solve triple integrals over regions of this type we must introduce a change of coordinates, either cylindrical or spherical, and like in the case of one variable, we must introduce a "change of variables theorem", in this case we use Jacobian matrices to reflect the fact that we are moving trough different multidimensional spaces.

When transforming integrals from one set of variables to another, the Jacobian determinant plays a crucial role. It helps adjust the integral measure to account for the change of variables. The Jacobian matrix generalizes the concept of the derivative to higher dimensions and allows for the transformation of integrals.

Definition (Jacobian Matrix)

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ be two sets of variables. Suppose each x_i is a function of the variables $u_1, u_2, \ldots, u_n$, that is:

$$x_i = x_i(u_1, u_2, \ldots, u_n) \quad \text{for} \quad i = 1, 2, \ldots, n.$$

The *Jacobian matrix* of the transformation $\mathbf{x} = \mathbf{x}(\mathbf{u})$ is the matrix of first-order partial derivatives:

$$D\mathbf{x} = \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{bmatrix}.$$

Definition (Jacobian Determinant)

The *Jacobian determinant* of a transformation $\mathbf{x}$, denoted as $\frac{\partial(x_1,x_2,\dots,x_n)}{\partial(u_1,u_2,\dots,u_n)}$ or $\det(D\mathbf{x})$, is the determinant of the Jacobian matrix:

$$\det(D\mathbf{x}) = \frac{\partial(x_1,x_2,\dots,x_n)}{\partial(u_1,u_2,\dots,u_n)} = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{vmatrix}.$$

This determinant gives the factor by which volumes (or areas in two dimensions) are scaled under the transformation.

Proposition (Multiplicative Property)

If $\mathbf{x} = \mathbf{x}(\mathbf{u})$ and $\mathbf{u} = \mathbf{u}(\mathbf{v})$, then the Jacobian of the composition of transformations $\mathbf{x} = \mathbf{x}(\mathbf{u}(\mathbf{v}))$ is the product of the Jacobians:

$$\frac{\partial(x_1,\dots,x_n)}{\partial(v_1,\dots,v_n)} = \frac{\partial(x_1,\dots,x_n)}{\partial(u_1,\dots,u_n)} \cdot \frac{\partial(u_1,\dots,u_n)}{\partial(v_1,\dots,v_n)}.$$

Proposition (Inverse Transformations)

If the transformation $\mathbf{x} = \mathbf{x}(\mathbf{u})$ is invertible, then the Jacobian of the inverse transformation $\mathbf{u} = \mathbf{u}(\mathbf{x})$ is the reciprocal of the Jacobian of the original transformation:

$$\frac{\partial(u_1,\dots,u_n)}{\partial(x_1,\dots,x_n)} = \left(\frac{\partial(x_1,\dots,x_n)}{\partial(u_1,\dots,u_n)}\right)^{-1}.$$

Example (Polar Coordinates)

Consider the transformation $P : \mathbb{R}_+ \times [0, 2\pi] \to \mathbb{R}^2$ from polar coordinates (r, θ) to Cartesian coordinates (x, y), where:

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

The Jacobian matrix is:

$$DP = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

The Jacobian determinant is:

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

Thus, when transforming an integral from Cartesian to polar coordinates, we include the factor r in the integral:

$$\int_R f(x, y)\, dx\, dy = \int_{R'} f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta.$$

Example (Cylindrical Coordinates)

Consider the transformation $C : \mathbb{R}_+ \times [0, 2\pi] \times \mathbb{R} \to \mathbb{R}^3$ from cylindrical coordinates $(r, \theta), z$ to Cartesian coordinates (x, y, z), where:

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

The Jacobian matrix is:

$$DC = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & 0 \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & 0 \\ 0 & 0 & \frac{\partial z}{\partial z} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The Jacobian determinant is:

$$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = r.$$

Thus, when transforming an integral from Cartesian to polar coordinates, we include the factor r in the integral:

$$\int_R f(x, y)\, dx\, dy = \int_{R'} f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta.$$

Example (Spherical Coordinates)

Let $T:\mathbb{R}^3\to\mathbb{R}^3$ the spherical change of coordinates given as $T(r,\phi,\theta)=(x,y,z)$, where

$$x=\rho\sin\phi\cos\theta,\quad y=\rho\sin\phi\sin\theta,\quad z=\rho\cos\phi.$$

The Jacobian determinant is given as

$$\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}=\begin{vmatrix}\sin\phi\cos\theta & -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta\\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta\\ \cos\phi & 0 & -\rho\sin\phi\end{vmatrix}.$$

Expanding along the last row, we get

$$\begin{aligned}\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}&=\cos\phi\begin{vmatrix}-\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta\\ \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta\end{vmatrix}-\rho\sin\phi\begin{vmatrix}\sin\phi\cos\theta & -\rho\sin\phi\sin\theta\\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta\end{vmatrix}\\ &=-\rho^2\cos^2\phi\sin\phi\sin^2\theta-\rho^2\cos^2\phi\sin\phi\cos^2\theta-\rho^2\sin^3\phi\cos^2\theta-\rho^2\sin^3\phi\sin^2\theta\\ &=-\rho^2\cos^2\phi\sin\phi-\rho^2\sin^3\phi=-\rho^2\sin\phi.\end{aligned}$$

Theorem (Change of Variables Theorem)

*Let f be a continuous function on a region R in $\mathbb{R}^n$, and suppose the **one-to-one**transformation $\mathbf{x} = \mathbf{x}(\mathbf{u})$ maps a region R' in $\mathbb{R}^n$ **onto** R. The integral of f over R can be transformed into an integral over R' using the Jacobian determinant. Specifically, the change of variables formula is:*

$$\int_R f(x_1, x_2, \ldots, x_n)\, d\mathbf{x} = \int_{R'} f(x_1(\mathbf{u}), x_2(\mathbf{u}), \ldots, x_n(\mathbf{u})) \left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} \right| d\mathbf{u}.$$

*Here, $|\det(J)|$ accounts for the scaling effect of the transformation on the **differential element** $d\mathbf{x}$.*

Remark

1. **One-to-one**: $\mathbf{x} = \mathbf{x}(\mathbf{u} : R' \to R$ is one-to-one if it holds that $\mathbf{x}(u_1) \neq \mathbf{x}(u_2)$ for every $u_1, u_1 \in R'$. This is the notion of injectivity.
2. **Onto**: $\mathbf{x} = \mathbf{x}(\mathbf{u} : R' \to R$ is onto if for all $u \in R$, there is a u' such that $\mathbf{x}(u') = u$. This corresponds to sobrejectivity (or suprajectivity).
3. Every linear transformation T is one-to-one and onto if and only $\det(DT) \neq 0$.
4. The absolute value of the Jacobian determinant $|\det(J)|$ is crucial because the transformation can flip orientation, which would result in a negative determinant. The absolute value ensures that the volume or area scaling is positive.

URJC

Theorem (Change of Variables for Polar, Spherical and Cylindrical Coordinates)

Let W be a region in Cartesian coordinates and W^* its corresponding transformation into polar (cylindrical, spherical) coordinates. Then it holds

$$\text{(polar)} \quad \iiint_W f(x,y)dxdy = \iiint_{W^*} f(r\cos\theta, \sin\theta)rdrd\theta.$$

$$\text{(cyl.)} \quad \iiint_W f(x,y,z)dxdydz = \iiint_{W^*} f(r\cos\theta, \sin\theta, z)rdrd\theta dz.$$

$$\text{(sph.)} \quad \iiint_W f(x,y,z)dxdydz = \iiint_{W^*} f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\rho^2\sin\phi d\rho d\theta d\phi.$$

Example (Polar coordinates)

Evaluate $\iint_D \log\left(x^2+y^2\right) dxdy$, where D is the region in the first quadrant lying between the arcs of the circles $x^2+y^2=a^2$ and $x^2+y^2=b^2$, where $0<a<b$.

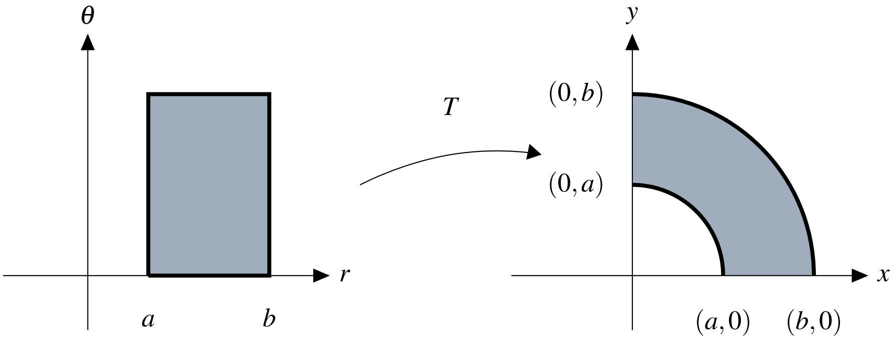


Figure: Representation of the polar region of the example

Example

These circles have the simple equations $r = a$ and $r = b$ in polar coordinates. Moreover, $r^2 = x^2 + y^2$ appears in the integrand. Thus, a change to polar coordinates will simplify both the integrand and the region of integration.

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

sends the rectangle D^* given by $a \le r \le b, 0 \le \theta \le \pi/2$ onto the region D. This transformation is one-to-one on D^* and so, by formula (7), we have

$$\iint_D \log\left(x^2+y^2\right) dxdy = \int_a^b \int_0^{\pi/2} r\log r^2 d\theta dr = \frac{\pi}{2}\int_a^b r\log r^2 dr = \frac{\pi}{2}\int_a^b 2r\log r dr$$

Applying integration by parts, or using the formula

$$\int x\log x dx = \frac{x^2}{2}\log x - \frac{x^2}{4}$$

we obtain the result

$$\frac{\pi}{2}\int_a^b 2r\log r dr = \frac{\pi}{2}\left[b^2\log b - a^2\log a - \frac{1}{2}\left(b^2-a^2\right)\right]$$

Remark

The Jacobian determinant r reflects the fact that small areas in polar coordinates grow by a factor of r as we move away from the origin. This scaling is captured by the Jacobian and is essential for correct integration in new coordinates.

Definition (Area)

The **area** of a region R in the plane can be computed using a double integral. If R is a region in $\mathbb{R}^2$, the area is given by:

$$\text{Area}(R) = \iint_R 1\, dA$$

where dA represents the differential area element.

This means that to find the area of a region R, we simply integrate 1 over R. This technique is useful when the region is bounded by curves, making it difficult to apply standard geometric formulas.

Definition (Volume)

The **volume** of a region D in $\mathbb{R}^3$ can be computed using a triple integral. The volume is given by:

$$\text{Volume}(D) = \iiint_D 1\, dV$$

where dV is the differential volume element.

Triple integrals extend the idea of double integrals to three dimensions and allow for the calculation of volumes of more complex regions in space.

Example

Find the volume of the solid region D bounded by the planes $x=0$, $y=0$, $z=0$, and the plane $x+y+z=1$.
The region D is a tetrahedron with vertices at $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. The volume can be computed as:

$$\text{Volume}(D) = \iiint_D 1\,dV = \int_0^1 \int_0^{1-x} \int_0^{1-x-y} dz\,dy\,dx$$

Evaluate the innermost integral:

$$\int_0^{1-x-y} dz = 1-x-y$$

Substitute and evaluate the second integral:

$$\int_0^{1-x} (1-x-y)\,dy = (1-x)y - \frac{y^2}{2}\bigg|_0^{1-x} = \frac{(1-x)^2}{2}$$

Finally, evaluate the outermost integral:

$$\int_0^1 \frac{(1-x)^2}{2}\,dx = \frac{1}{2}\int_0^1 (1-2x+x^2)\,dx = \frac{1}{2}\left[x - x^2 + \frac{x^3}{3}\right]_0^1 = \frac{1}{6}$$

Thus, the volume of the tetrahedron is $\frac{1}{6}$.

Definition (Mass)

The mass M of a continuous body occupying a region $D \subset \mathbb{R}^3$ with density function $\rho(x)$ is given by:

$$M = \iiint_D \rho(x)\, dV,$$

where $\rho(x)$ represents the density at the point $x \in D$, and dV is the differential volume element.

Definition (Center of Mass)

The center of mass $\mathbf{C} = (C_1, C_2, C_3)$ of a continuous body with density function $\rho(x)$ is the point at which the entire mass can be considered to be concentrated. It is given by:

$$C_i = \frac{1}{M} \iiint_D x_i \rho(x)\, dV \quad \text{for} \quad i = 1, 2, 3,$$

where x_i represents the i-th coordinate of the point x, and M is the total mass of the body.

Definition (Moment of Inertia about an Axis)

The moment of inertia I of a continuous body with density function $\rho(x)$ about an axis (say, the z-axis) is given by:

$$I_z = \iiint_D (x^2 + y^2)\rho(x,y,z)\,dV,$$

where (x,y,z) are the coordinates of a point in D, and dV is the differential volume element.

Definition (Moment of Inertia about a Point)

The moment of inertia I_p of a continuous body about a point p is given by:

$$I_p = \iiint_D r^2 \rho(x)\,dV,$$

where r is the distance from the point p to the point $x \in D$, and dV is the differential volume element.

Definition (Average Values)

The **average value** of a function of one variable on the interval $[a,b]$ is defined by

$$[f]_{\mathrm{av}} = \frac{\int_a^b f(x)dx}{b-a}$$

Likewise, for functions of two variables, the ratio of the integral to the area of D,

$$[f]_{\mathrm{av}} = \frac{\iint_D f(x,y)dxdy}{\iint_D dxdy}$$

is called the average value of f over D. Similarly, the average value of a function f on a region W in 3 -space is defined by

$$[f]_{\mathrm{av}} = \frac{\iiint_W f(x,y,z)dxdydz}{\iiint_W dxdydz}$$

§II. Vector Calculus

- Intr. Differential Geometry and Vector Calculus
  - Parametrized Paths. Arclength
  - Differential Geometry. Operators

URJC

A parametrized curve represents a path in space (plane) described by a vector-valued function of one variable.

Definition (Path in $\mathbb{R}^n$)

A **path** in $\mathbb{R}^n$ is a continuous function $\mathbf{x} : I \subseteq \mathbb{R} \to \mathbb{R}^n$. If $I = [a,b]$ for some numbers $a < b$, then the points $\mathbf{x}(a)$ and $\mathbf{x}(b)$ are called the endpoints of the path $\mathbf{x}$.

Remark

I can be on any of the forms $[a,b]$, (a,b), $[a,b)$, $(a,b]$, $[a,\infty)$, (a,∞), $(-\infty,b]$, $(-\infty,b)$, or $(-\infty,\infty) = \mathbb{R}$

Example

Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{R}^3$ with $\mathbf{a} \neq \mathbf{0}$. Let $\mathbf{x} : (-\infty,\infty) \to \mathbb{R}^3$ be the function given by

$$\mathbf{x}(t) = \mathbf{b} + t\mathbf{a}$$

Then, this function $\mathbf{x}$ defines the path along the straight line parallel to $\mathbf{a}$, and passing through the endpoint of the position vector of $\mathbf{b}$

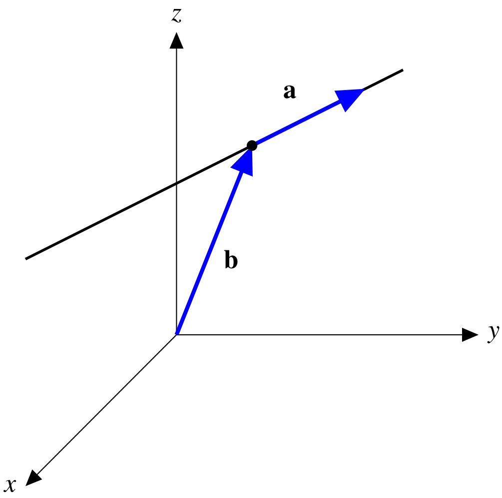


Figure: Parametric line in the space

URJC

Example

The parametrized curve of a circle of radius R in the plane $\mathbb{R}^2$ can be described by:

$$\mathbf{r}(t) = \begin{bmatrix} R\cos(t) \\ R\sin(t) \end{bmatrix}, \quad t \in [0, 2\pi].$$

This represents a circle centred at the origin with radius R.

Definition (Velocity, Speed, Acceleration and Tangent)

Let $\mathbf{x} : I \subseteq \mathbb{R} \to \mathbb{R}^n$ be a differentiable path. Then, the **velocity** $\mathbf{v}(t) = \mathbf{x}'(t)$ exists, and we define the **speed** of $\mathbf{x}$ to be the magnitude of velocity

$$\text{Speed} = \|\mathbf{v}(t)\|$$

If $\mathbf{v}$ is itself differentiable, then we call $\mathbf{v}'(t) = \mathbf{x}''(t)$ the **acceleration** of $\mathbf{x}$ and denote it by $\mathbf{a}(t)$.
The **tangent line** to $\mathbf{x}$ at $\mathbf{x}_0 = \mathbf{x}(t_0) \in \mathbf{x}(t)$ is the line

$$\mathbf{l}(s) = \mathbf{x}_0 + s\mathbf{v}_0 \qquad \text{or} \qquad \mathbf{l}(t) = \mathbf{x}_0 + (t - t_0)\mathbf{v}_0$$

Example (Circular Helix)

Consider the parametrized curve of a helix in $\mathbb{R}^3$:

$$\mathbf{r}(t) = \begin{bmatrix} R\cos(t) \\ R\sin(t) \\ ct \end{bmatrix}, \quad t \in [0, 2\pi].$$

The velocity and acceleration vectors are:

$$\mathbf{v}(t) = \mathbf{r}'(t) = (-R\sin(t), R\cos(t), c),$$
$$\mathbf{a}(t) = \mathbf{r}''(t) = (-R\cos(t), -R\sin(t), c).$$

The speed of the helix is:

$$|\mathbf{r}'(t)| = \sqrt{(-R\sin(t))^2 + (R\cos(t))^2 + c^2} = \sqrt{R^2 + c^2}.$$

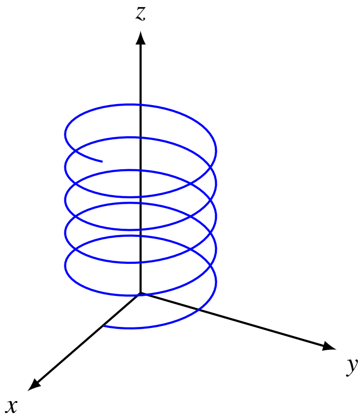


Figure: Circular Helix

Example

Let $\mathbf{x}(t) = (3t+2, t^2-7, t-t^2)$. We find parametric equations for the line tangent to $\mathbf{x}$ at $\mathbf{x}(1) = (5,-6,0)$ For this path

$$\mathbf{v}(t) = \mathbf{x}'(t) = 3\mathbf{i} + 2t\mathbf{j} + (1-2t)\mathbf{k}$$

So that

$$\mathbf{v}_0 = \mathbf{v}(t_0) = \mathbf{v}(1) = 3\mathbf{i} + 2\mathbf{j} - \mathbf{k}$$

Thus,

$$\mathbf{l}(t) = \mathbf{x}_0 + (t-t_0)\mathbf{v}_0 = (5\mathbf{i} - 6\mathbf{j}) + (t-1)(3\mathbf{i} + 2\mathbf{j} - k)$$

Arclength of a Parametrized Curve

The arclength of a curve is the total distance traveled along the curve between two points. It can be computed using the velocity vector of the parametrized curve.

Definition (Arclenght)

Let $\mathbf{r} : [a,b] \to \mathbb{R}^n$ be a smooth parametrized path, $\mathbf{r}(t) = (x_1(t), \ldots, x_n(t))$. The **arclength** of the curve from $t = a$ to $t = b$ is given by the integral:

$$s = \int_a^b |\mathbf{r}'(t)|\, dt = \int_a^b \sqrt{\left(\frac{dx_1}{dt}\right)^2 + \cdots + \left(\frac{dx_n}{dt}\right)^2}\, dt.$$

Example (Arclength of a Circular Arc)

Consider the parametrized curve of a quarter-circle in $\mathbb{R}^2$ with radius R:

$$\mathbf{r}(t) = \begin{bmatrix} R\cos(t) \\ R\sin(t) \end{bmatrix}, \quad t \in \left[0, \frac{\pi}{2}\right].$$

The velocity vector is:

$$\mathbf{r}'(t) = \begin{bmatrix} -R\sin(t) \\ R\cos(t) \end{bmatrix}.$$

The magnitude of the velocity vector is:

$$|\mathbf{r}'(t)| = \sqrt{(-R\sin(t))^2 + (R\cos(t))^2} = R.$$

The arclength of the curve from $t = 0$ to $t = \frac{\pi}{2}$ is:

$$s = \int_0^{\frac{\pi}{2}} R\,dt = R\,[t]_0^{\frac{\pi}{2}} = R\frac{\pi}{2}.$$

Thus, the length of the quarter-circle is $\frac{\pi R}{2}$.

URJC

Example (Arclength of a Helix)

Consider the helix $\mathbf{r}(t) = \begin{bmatrix} R\cos(t) \\ R\sin(t) \\ ct \end{bmatrix}$, where $t \in [0, 2\pi]$. The magnitude of the velocity vector is:

$$|\mathbf{r}'(t)| = \sqrt{(-R\sin(t))^2 + (R\cos(t))^2 + c^2} = \sqrt{R^2 + c^2}.$$

The arclength of the helix from $t = 0$ to $t = 2\pi$ is:

$$s = \int_0^{2\pi} \sqrt{R^2 + c^2}\, dt = \sqrt{R^2 + c^2} \int_0^{2\pi} dt = 2\pi\sqrt{R^2 + c^2}.$$

Thus, the length of the helix is $2\pi\sqrt{R^2 + c^2}$.

Definition (Arclenght parametrization)

A curve in $\mathbb{R}^n$, given by a parametric equation $\mathbf{r}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$, can be re-parametrized using its arc length, s, as the parameter. The arc length of a curve between two points is defined as:

$$s(t) = \int_{t_0}^{t} \|\mathbf{r}'(u)\| \, du,$$

where:

- t_0 is the starting parameter value,
- $\mathbf{r}'(u)$ is the derivative of the position vector $\mathbf{r}(u)$,
- $\|\mathbf{r}'(u)\|$ is the magnitude of the velocity vector at u.

If t is the original parameter, the arc length $s(t)$ measures the total distance traveled along the curve from t_0 to t.

Proposition (Properties of Arc-Length Parametrization)

1. **Unit Speed**: When a curve is parametrized by its arc length s, the magnitude of its velocity vector is always 1:
$$\|\mathbf{r}'(s)\| = 1.$$
2. **Reparametrization**: Any regular curve (one where $\|\mathbf{r}'(t)\| > 0$) can be reparametrized by its arc length. This involves inverting the relationship $s(t)$ to express t as a function of s, and substituting this back into $\mathbf{r}(t)$.

Steps to Find the Arc-Length Parametrization

1. **Compute the Arc Length:** Calculate the arc length $s(t)$ from a given starting point t_0:
$$s(t) = \int_{t_0}^{t} \|\mathbf{r}'(u)\|\, du.$$
2. **Invert the Relationship:** Solve $s(t)$ for t as a function of s: $t = t(s)$.
3. **Reparametrize the Curve:**
Substitute $t(s)$ into the original curve $\mathbf{r}(t)$ to obtain the arc-length parametrization $\mathbf{r}(s)$.

Example

Consider the parametric curve of a helix:

$$\mathbf{r}(t) = (\cos(t), \sin(t), bt),$$

where $b > 0$ is a constant. To find the arc-length parametrization:

1. Compute $\|\mathbf{r}'(t)\|$:

$$\mathbf{r}'(t) = (-\sin(t), \cos(t), b), \quad \|\mathbf{r}'(t)\| = \sqrt{(-\sin(t))^2 + (\cos(t))^2 + b^2} = \sqrt{1+b^2}.$$

2. Find $s(t)$:

$$s(t) = \int_0^t \sqrt{1+b^2}\, du = \sqrt{1+b^2}\, t.$$

3. Solve for $t(s)$ $\qquad t = \frac{s}{\sqrt{1+b^2}}$.
4. Reparametrize: Substitute $t = \frac{s}{\sqrt{1+b^2}}$ into $\mathbf{r}(t)$:

$$\mathbf{r}(s) = \left(\cos\left(\frac{s}{\sqrt{1+b^2}}\right), \sin\left(\frac{s}{\sqrt{1+b^2}}\right), \frac{bs}{\sqrt{1+b^2}}\right).$$

This gives the arc-length parametrization of the helix.

Definition (Reparametrization)

Let $\mathbf{x} : [a,b] \to \mathbb{R}^n$ be a piecewise C^1 path, and consider another C^1 path $\mathbf{y} : [c,d] \to \mathbb{R}^n$. We say that $\mathbf{y}$ is a **reparametrization** of $\mathbf{x}$ if there is a one-one and onto function $u : [c,d] \to [a,b]$ of class C^1. With inverse $u^{-1} : [a,b] \to [c,d]$ that is also of class C^1, such that $\mathbf{y}(t) = \mathbf{x}(u(t))$, that is, $\mathbf{y} = \mathbf{x} \circ u$.

1. If $\mathbf{x}(a) = \mathbf{y}(c)$ and $\mathbf{x}(b) = \mathbf{y}(d)$ we say that that reparametrization is orientation- preserving, $\mathbf{x}$ and $\mathbf{y}$ trace the curve in the same direction.
2. If $\mathbf{x}(a) = \mathbf{y}(d)$ and $\mathbf{x}(b) = \mathbf{y}(c)$ we say that that reparametrization is orientation- reversing, $\mathbf{x}$ and $\mathbf{y}$ trace the curve in opposite direction.

Remark

Thus, any reparametrization of a path must have the same underlying image curve as the original path.

Example

Consider the path

$$\mathbf{x}(t) = (1+2t, 2-t, 3+5t), \quad 0 \leq t \leq 1$$

It traces the line segment from the point $(1,2,3)$ to the point $(3,1,8)$
Also does the path

$$\mathbf{y}(t) = (1+2t^2, 2-t^2, 3+5t^2), \quad 0 \leq t \leq 1$$

We have that **y** is a reparametrization of **x** via the change of variable

$$u(t) = t^2$$

Note that this parametrization preserves the orientation.

Example (Opposite path)

$\mathbf{x} : [a,b] \to \mathbb{R}^n$ be a piecewise C^1 path. We can define $\mathbf{x}_{opp} : [a,b] \to \mathbb{R}^n$ as

$$\mathbf{x}_{opp}(t) = \mathbf{x}(a+b-t)$$

Note that $\mathbf{x}_{opp}$ is a reparametrization of C via $u(t) = a+b-t$, that reverses the orientation.

Proposition (The speed depends on the parametrization)

A reparametrization u changes the speed by a factor $|u'(t)|$

Proof. From the chain rule

$$\|\mathbf{y}'(t)\| = |u'(t)| \|\mathbf{x}'(t)\|$$

Since u is one-one, it follows that either

- $u'(t) \geq 0$ for all $t \in [a,b]$ (orientation is preserved) or
- $u'(t) \leq 0$ for all $t \in [a,b]$ (orientation is reversed)

□

In differential geometry, curvature and torsion describe the bending and twisting of a curve in space.

Definition (Curvature)

Let $\mathbf{r}(t)$ be a smooth parametrized curve in $\mathbb{R}^3$. The **curvature** $\kappa(t)$ of the curve at a point is a measure of how quickly the curve changes direction and is defined as:

$$\kappa(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3}.$$

Definition (Torsion)

The **torsion** $\tau(t)$ of a curve at a point measures how much the curve twists out of the plane of curvature. It is defined as:

$$\tau(t) = \frac{(\mathbf{r}'(t) \times \mathbf{r}''(t)) \cdot \mathbf{r}'''(t)}{|\mathbf{r}'(t) \times \mathbf{r}''(t)|^2}.$$

Example

Consider the helix parametrized by $\mathbf{r}(t) = \begin{bmatrix} R\cos(t) \\ R\sin(t) \\ ct \end{bmatrix}$. The curvature $\kappa(t)$ and torsion $\tau(t)$ are:

$$\kappa(t) = \frac{R}{R^2 + c^2}, \quad \tau(t) = \frac{c}{R^2 + c^2}.$$

Thus, the curvature is constant and describes how the helix bends, while the torsion describes how the helix twists.

A vector field assigns a vector to each point in a region of space. Vector fields are used to model various physical phenomena, such as fluid flow, electromagnetic fields, and gravitational fields.

Definition (Vector Field)

A **vector field** on $\mathbb{R}^n$ is a function $\mathbf{F} : A \subset \mathbb{R}^n \to \mathbb{R}^n$ that assigns to each point $\mathbf{x} \in A \subset \mathbb{R}^n$ a vector $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$. For $n = 2, 3$ we call it vector field in the plane, space, respectively.

Example

An example of a vector field in $\mathbb{R}^3$ is the gravitational field of a point mass:

$$\mathbf{F}(x,y,z) = \frac{-GM}{(x^2+y^2+z^2)^{3/2}} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

where G is the gravitational constant and M is the mass.

Example

Consider the vector field

$$F(x,y) = (u(x,y), v(x,y)) = \left(\frac{1}{1+(x+y)^2}, \frac{x+y}{1+(x+y)^2} \right)$$

See representation in figure 9

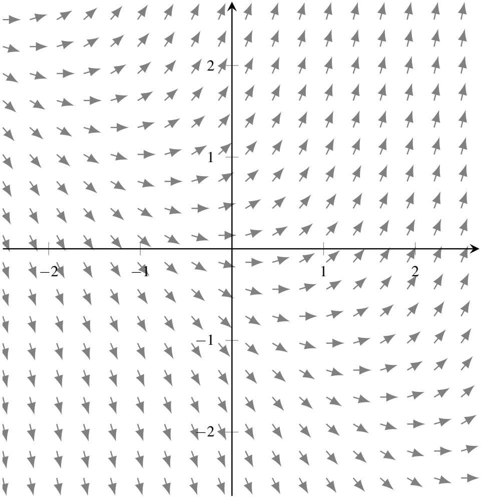


Figure: Normalized gradient vector field

Example (Radial vector Field)

A simple vector field in $\mathbb{R}^2$ is the radial vector field:

$$\mathbf{F}(x,y) = \begin{bmatrix} x \\ y \end{bmatrix}.$$

At each point, the vector points outward from the origin, with its magnitude proportional to the distance from the origin. See figure 10

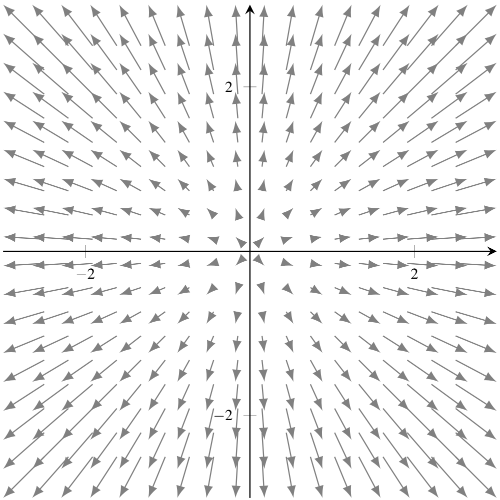


Figure: Radial Vector Field

Example (Outward vector field)

Consider the vector field

$$F(x,y) = (u(x,y), v(x,y)) = (-y, x)$$

Note in figure 11 how the directions are tangent to concentric circles.

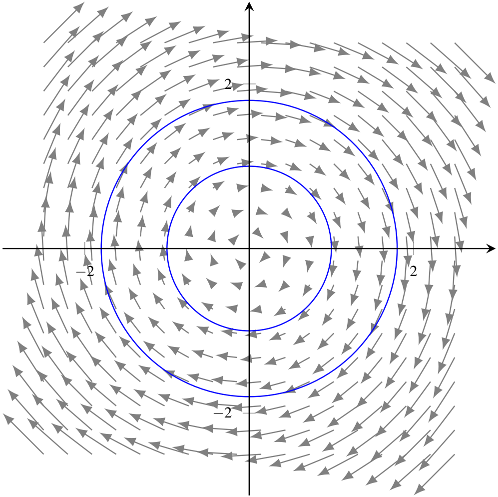


Figure: Outward vector field

URJC

In vector calculus, the gradient, divergence, and curl are operations that provide information about scalar and vector fields. These operations are expressed using the **del operator** (also called the **nabla operator**), denoted ∇.

Definition (Gradient of a Scalar Field)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar field. The **gradient** of f, denoted ∇f, is the vector of partial derivatives:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}.$$

Example

Let $f(x,y,z) = x^2 + y^2 + z^2$. Then the gradient of f is:

$$\nabla f(x,y,z) = (2x, 2y, 2z)^t.$$

Definition (Divergence of a Vector Field)

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field. The **divergence** of $\mathbf{F}$, denoted $\nabla \cdot \mathbf{F}$, is the scalar field defined by the dot product of ∇ with $\mathbf{F}$:

$$\nabla \cdot \mathbf{F} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \cdots + \frac{\partial F_n}{\partial x_n}.$$

Example

For the vector field $\mathbf{F}(x,y,z) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, the divergence is:

$$\nabla \cdot \mathbf{F} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3.$$

Definition (Curl of a Vector Field)

Let $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field. The **curl** of $\mathbf{F} = (M, N, P)$, denoted $\nabla \times \mathbf{F}$, is the vector field defined by:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ M & N & P \end{vmatrix} = (P_y - N_z)\mathbf{i} + (M_z - P_x)\mathbf{j} + (N_x - M_y)\mathbf{k}$$

where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the unit vectors in the x-, y-, and z-directions, respectively.

Example

Consider the vector field $\mathbf{F}(x, y, z) = \begin{bmatrix} -y \\ x \\ 0 \end{bmatrix}$. The curl of $\mathbf{F}$ is:

$$\nabla \times \mathbf{F} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Remark

The **del operator** ∇, when applied to scalar and vector fields, produces the following:

- The **gradient** of a scalar field f: ∇f, a vector field.
- The **divergence** of a vector field $\mathbf{F}$: $\nabla \cdot \mathbf{F}$, a scalar field.
- The **curl** of a vector field $\mathbf{F}$: $\nabla \times \mathbf{F}$, a vector field in $\mathbb{R}^3$.

§II. Vector Calculus

- Integral Theorems of Vector Calculus
  - Line Integrals
  - Conservative Fields
  - Green's Theorem
  - Parametric Surfaces. Surface Integrals
  - Stokes' and Gauss' Theorems

Definition (Scalar Line Integral)

Let $f : \mathbb{R}^n \to \mathbb{R}$ whose domain contains the C^1 path $\mathbf{x}$. The **Scalar Line Integral** of f along $\mathbf{x}$ is denoted as $\int_{\mathbf{x}} f\, ds$ and defined as

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

Remark

The scalar line integral measures the "weighted length" of the curve with respect to the scalar field f. It is often interpreted as the total quantity of f accumulated along the curve C.

Proposition (The line integral is independent from the parametrization)

Let $\mathbf{x} : [a,b] \to \mathbb{R}^n$ be a piecewise C^1 path and $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a continuous function whose domain X contains the image of $\mathbf{x}$. If $\mathbf{y} : [c,d] \to \mathbb{R}^n$ is any reparametrization of $\mathbf{x}$, then

$$\int_{\mathbf{y}} f\, ds = \int_{\mathbf{x}} f\, ds$$

Example

Compute the scalar line integral of $f(x,y) = x^2 + y^2$ along the curve C parametrized by $\mathbf{r}(t) = (\cos t, \sin t)$ for $t \in [0, 2\pi]$.

Solution: The magnitude of the derivative of the parametrization is:

$$\left|\mathbf{r}'(t)\right| = |(-\sin t, \cos t)| = 1.$$

Thus, the scalar line integral becomes:

$$\int_C f(x,y)\,ds = \int_0^{2\pi} (\cos^2 t + \sin^2 t)\,dt = \int_0^{2\pi} 1\,dt = 2\pi.$$

The total value of the scalar line integral is 2π, which corresponds to the circumference of the unit circle.

Definition (Vector Line Integral)

Let $\mathbf{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a vector field, and let $C \subset D$ be a smooth curve parametrized by $\mathbf{r}(t)$ for $t \in [a,b]$. The *vector line integral* of $\mathbf{F}$ along C is given by:

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

where $\mathbf{F}(\mathbf{r}(t))$ is the evaluation of the vector field at each point on the curve, and $\mathbf{r}'(t)$ is the tangent vector to the curve.

Remark

The vector line integral measures the "work" done by the vector field $\mathbf{F}$ along the curve C, which can be thought of as the total force applied in the direction of motion along the curve.

Proposition (Scalar Line Integrals and Arc Length)

If $f(x,y,z) = 1$, the scalar line integral reduces to the arc length of the curve C:

$$\int_C 1\, ds = \int_C ds = \text{Arc Length of } C.$$

Proposition (Vector Line Integrals and Conservative Fields)

If the vector field $\mathbf{F}$ is conservative, meaning there exists a potential function ϕ such that $\mathbf{F} = \nabla\phi$, the vector line integral over a curve C depends only on the endpoints of C:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)).$$

This is a special case of the Fundamental Theorem for Line Integrals.

Theorem (Fundamental Theorem of Line Integrals)

Let $\mathbf{F} = \nabla\phi$ *be a conservative vector field, and let* C *be a smooth curve from* $\mathbf{r}(a)$ *to* $\mathbf{r}(b)$*. Then the vector line integral of* $\mathbf{F}$ *along* C *is given by:*

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)).$$

Remark

The Fundamental Theorem of Line Integrals simplifies the computation of vector line integrals in conservative fields by relating the integral to the potential function.

Example

Let $\mathbf{F}(x,y) = (-y,x)$ be a vector field and compute the vector line integral of $\mathbf{F}$ along the unit circle parametrized by $\mathbf{r}(t) = (\cos t, \sin t)$ for $t \in [0, 2\pi]$.
Let us compute $\int_C F$. The vector field evaluated along the curve is $\mathbf{F}(\cos t, \sin t) = (-\sin t, \cos t)$, and the tangent vector is $\mathbf{r}'(t) = (-\sin t, \cos t)$. The dot product is:

$$\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = (-\sin t, \cos t) \cdot (-\sin t, \cos t) = \sin^2 t + \cos^2 t = 1.$$

Thus, the vector line integral is:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} 1\, dt = 2\pi.$$

The total value of the vector line integral is 2π, which corresponds to the work done by the vector field around the unit circle.

Example

Let $\mathbf{F}$ be the radial vector field on $\mathbb{R}^3$ given by

$$\mathbf{F} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

Let $\mathbf{x} : [0,1] \to \mathbb{R}^3$ be the path

$$\mathbf{x}(t) = (t, 3t^2, 2t^3)$$

Then

$$\begin{aligned}
\mathbf{x}'(t) &= (1, 6t, 6t^2) \\
\int_{\mathbf{x}} \mathbf{F} \cdot d\mathbf{s} &= \int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) dt \\
&= \int_0^1 \left(t\mathbf{i} + 3t^2\mathbf{j} + 2t^3\mathbf{k}\right) \cdot \left(\mathbf{i} + 6t\mathbf{j} + 6t^2\mathbf{k}\right) dt \\
&= \int_0^1 \left(t + 18t^3 + 12t^5\right) dt = \left(\frac{1}{2}t^2 + \frac{9}{2}t^4 + 2t^6\right)\Big|_0^1 = 7
\end{aligned}$$

Proposition

Let $\mathbf{x} : [a,b] \to \mathbb{R}^n$ be a piecewise C^1 path. Let $\mathbf{F} : X \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a continuous vector field whose domain X contains the image of $\mathbf{x}$. If $\mathbf{y} : [c,d] \to \mathbb{R}^n$ is any reparametrization of $\mathbf{x}$, then

$$\int_{\mathbf{y}} \mathbf{F} \cdot d\mathbf{s} = \int_{\mathbf{x}} \mathbf{F} \cdot d\mathbf{s} \quad \text{or} \quad \int_{\mathbf{y}} \mathbf{F} \cdot d\mathbf{s} = -\int_{\mathbf{x}} \mathbf{F} \cdot d\mathbf{s}$$

Remark

1. This proposition states that vector line integrals are independent of reparametrization up to a sign.
2. This sign depends only on whether the reparametrization preserves or reverses orientation.

Let us introduce the definition of conservative vector field.

Definition (Conservative Vector Field)

A vector field $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is called *conservative* if there exists a scalar potential function f such that:

$$\mathbf{F} = \nabla f.$$

In other words, the vector field is the gradient of some potential function f.

Theorem (Criteria for Conservative Vector Fields)

A vector field $\mathbf{F} = (M, N, P)$ *is conservative in a simply connected region if and only if:*

$$Curl F = \nabla \times F = 0$$

If $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F = (M, N)$, *this is equivalent to*

$$N_y - M_x = \frac{\partial N}{\partial y} - \frac{\partial M}{\partial x} = 0$$

Example

Consider the vector field $\mathbf{F}(x,y) = (2xy, x^2)$. Note that $\nabla f = F$, where

$$f(x,y) = x^2 y$$

thus, this is a gradient field. This means that the vector field is conservative.
Now let us use the curl condition

$$\frac{\partial P}{\partial y} = \frac{\partial}{\partial y}(2xy) = 2x, \quad \frac{\partial Q}{\partial x} = \frac{\partial}{\partial x}(x^2) = 2x.$$

Since $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, the vector field is conservative.
Let us now apply a method to find the potential function $f(x,y)$. Note that

$$\nabla f = (2xy, x^2).$$

Integrating $2xy$ with respect to x, we get:

$$f(x,y) = x^2 y + g(y),$$

where $g(y)$ is an arbitrary function of y. Taking the partial derivative with respect to y and setting it equal to x^2, we find that $g(y)$ must be constant. Therefore, the potential function is:

$$f(x,y) = x^2 y.$$

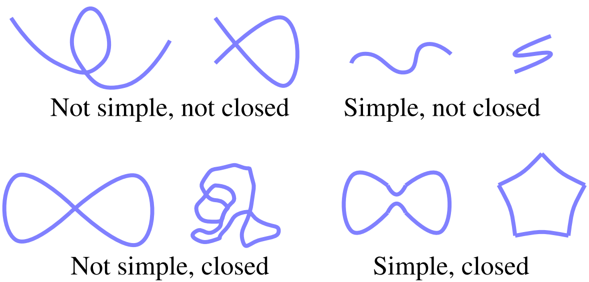
Theorem (Fundamental Theorem of Line Integrals)

Let $\mathbf{F} = \nabla\phi$ *be a conservative vector field, and let* C *be a smooth curve from* $\mathbf{r}(a)$ *to* $\mathbf{r}(b)$*. Then the vector line integral of* $\mathbf{F}$ *along* C *is given by:*

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \phi(\mathbf{r}(b)) - \phi(\mathbf{r}(a)).$$

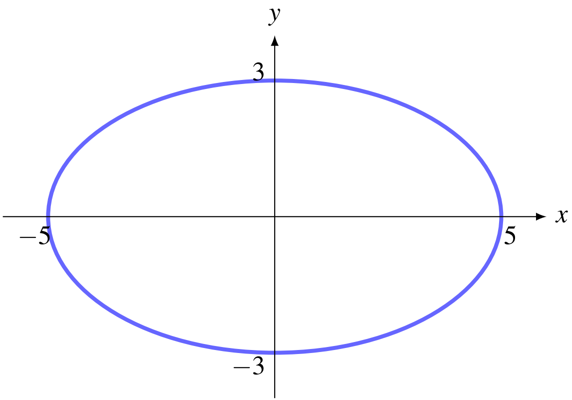
Remark

The Fundamental Theorem of Line Integrals simplifies the computation of vector line integrals in conservative fields by relating the integral to the potential function. Note that the value of the integral is independent from the parametrization of C, only depends on the endpoints., this is called **path independence**.



Definition (Simple and Closed Curves)

We say that a path $\mathbf{x} : [a,b] \to \mathbb{R}^n$ is **closed** if $\mathbf{x}(a) = \mathbf{x}(b)$.
We say that a path $\mathbf{x} : [a,b] \to \mathbb{R}^n$ is **simple** if it has not self-intersections, that is $\mathbf{x}$ is one-to-one for $t \in (a,b)$



Example

Consider the ellipse

$$\frac{x^2}{25} + \frac{y^2}{9} = 1$$

Example

It is a **simple**, **closed curve** that may be parametrized by either

$$\mathbf{x}(t) = (5\cos t, 3\sin t), \quad \mathbf{x} : [0, 2\pi] \to \mathbb{R}^2$$

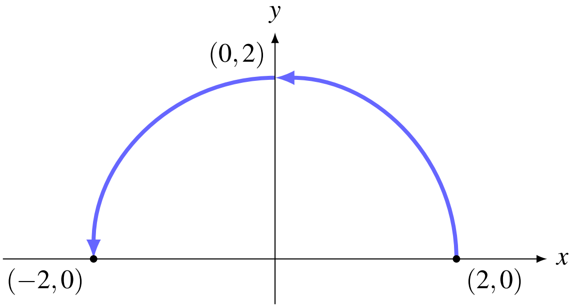
or

$$\mathbf{y}(t) = (5\cos 2(\pi - t), 3\sin 2(\pi - t)), \quad \mathbf{y} : [0, \pi] \to \mathbb{R}^2$$

However, if we consider the parametrization

$$\mathbf{z}(t) = (5\cos t, 3\sin t), \quad \mathbf{z} : [0, 6\pi] \to \mathbb{R}^2$$

is not a **simple** parametrization, since it traces the ellipse three times as t increases from 0 to 6π. $\mathbf{z}$ is not one-one.



Example

Let C be the upper semicircle of radius 2, centered at $(0,0)$ and oriented counterclockwise from $(2,0)$ to $(-2,0)$. We calculate

$$I = \int_C (x^2 - y^2 + 1)\,ds$$

Example

If we consider the parametrization

$$\mathbf{x}(t) = (2\cos t, 2\sin t), \quad 0 \leq t \leq \pi$$

we obtain $I = 2\pi$. On the other hand, if we consider

$$\mathbf{y}(t) = (-2\cos 2t, -2\sin 2t), \quad -\frac{\pi}{2} \leq t \leq 0$$

Then $I = -2\pi$
Note how the change of orientation occurs in the parametrization and in the domain.

Example

Consider the force $\mathbf{F} = (x, -y, x+y+z)$. Calculate the work done by $\mathbf{F}$ on a particle that move from $(0,0,0)$ to $(2,12,0)$ along the parabola $y = 3x^3$, $z = 0$.
Parametrize the parabola by

$$x = t, y = 3t^2, z = 0 \text{ for } 0 \leq t \leq 2$$

Then,

$$\text{Work} = \int_C \mathbf{F} \cdot d\mathbf{s} = \int_{\mathbf{x}} \mathbf{F} \cdot d\mathbf{s} = \int_0^2 \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) dt = -70$$

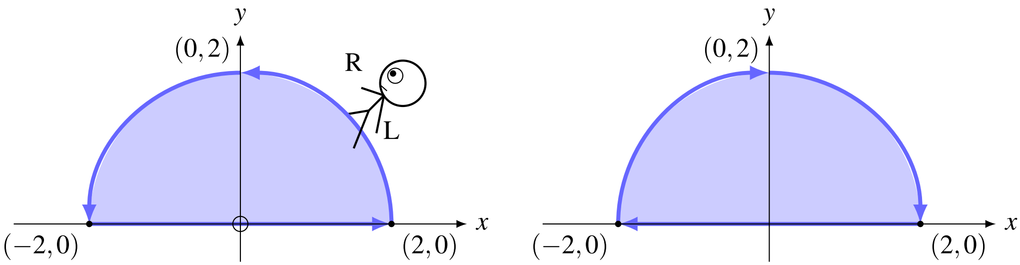
If we change the direction of the movement, the sign of the work would change.

Definition (Positive orientation)

Let ∂D be a closed curve surrounding a set D. We say that the parametrization $\mathbf{x}$ of ∂D is positively oriented if the interior of the set is always at left of the direction.

Example

The left-side semicircle is oriented positively (see figure), the right-side semicircle is oriented negatively. Note that if we move trough the boundary of the semicircle, the interior of the set is always at left (in case of positive orientation).



Theorem (Green's Theorem)

Let C be a positively oriented, simple, closed curve in the plane, and let R be the region enclosed by C. If $M(x,y)$ and $N(x,y)$ have continuous partial derivatives on an open region containing R, then Green's Theorem states that:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int\int_R \left(\frac{\partial M}{\partial x} - \frac{\partial N}{\partial y} \right) dA,$$

where $\mathbf{F} = M(x,y)\mathbf{i} + N(x,y)\mathbf{j}$ is a vector field and $d\mathbf{r}$ is the differential vector along C.

Remark

Green's Theorem converts the circulation (line integral) of a vector field around a closed curve into a double integral over the region it encloses. This result is especially useful in converting otherwise difficult line integrals into simpler area integrals. The symbol $\oint_C$ indicates that the line integral is taken over one or more **closed** curves

Proof. The proof of Green's Theorem involves breaking the region R into smaller rectangles, applying the Fundamental Theorem of Calculus to each, and summing the contributions to relate the line integral to the double integral. □

Example

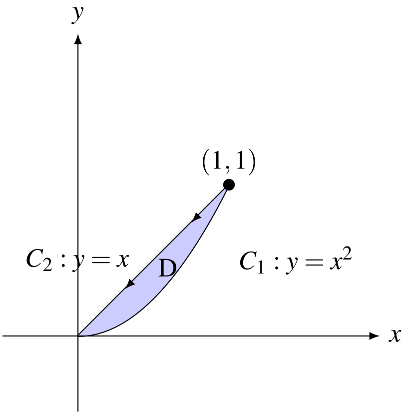
Consider the vector field $\mathbf{F}(x,y) = (-y,x)$ and the curve C as the unit circle $x^2+y^2=1$.
The line integral is:

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \oint_C (-y\,dx + x\,dy).$$

Using Green's Theorem, the corresponding double integral is:

$$\int\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA = \int\int_R (1+1)\,dA = 2\times \text{Area}(R).$$

Since R is the unit circle, its area is π, so the integral becomes 2π, which is the value of the original line integral.



Example

Determine

$$\oint_{\partial D} \mathbf{F} \cdot d\mathbf{s} = \oint_{\partial D} xy\, dx + y^2\, dy$$

Let $\mathbf{F} = xy\mathbf{i} + y^2\mathbf{j}$ and let D be the first quadrant region bounded by the line $y = x$ and the parabola $y = x^2$, see figure.

Example

We need to parametrize the two C^1 pieces of ∂D separately

$$C_1 : \begin{cases} x = t \\ y = t^2 \end{cases} ,0 \le t \le 1 \text{ and } C_2 : \begin{cases} x = 1-t \\ y = 1-t \end{cases} ,0 \le t \le 1$$

Note the orientations of C_1 and C_2

Then

$$\oint_{\partial D} xy\,dx + y^2\,dy = \oint_{C_1} xy\,dx + y^2\,dy + \oint_{C_2} xy\,dx + y^2\,dy$$

$$= \int_0^1 \left(t \cdot t^2 + t^4 \cdot 2t\right) dt + \int_0^1 \left((1-t)^2 + (1-t)^2\right)(-dt)$$

$$= \int_0^1 \left(t^3 + 2t^5\right) dt + \int_0^1 2(1-t)^2(-dt)$$

$$= \left(\frac{1}{4}t^4 + \frac{2}{6}t^6\right)\Bigg|_0^1 + \left(\frac{2}{3}(1-t)^3\right)\Bigg|_0^1 = \frac{1}{4} + \frac{2}{6} - \frac{2}{3} = -\frac{1}{12}$$

Example

On the other hand

$$\int\int_D \left(\frac{\partial}{\partial x}\left(y^2\right) - \frac{\partial}{\partial y}(xy)\right) dxdy = \int_0^1 \int_{x^2}^{x} -x \, dydx$$
$$= \int_0^1 -x\left(x-x^2\right)\, dx = \int_0^1 \left(x^3-x^2\right)\, dx = \left(\frac{1}{4}x^4 - \frac{1}{3}x^3\right)\Bigg|_0^1$$
$$= \frac{1}{4} - \frac{1}{3} = -\frac{1}{12}$$

Green's Theorem holds, the line integral and the double integral agree

Surface integrals allow us to generalize the concept of integration to curved surfaces in three-dimensional space. To compute surface integrals, we often use parametrized surfaces, which provide a way to describe surfaces using a set of parameters.

Definition (Parametrized Surface)

A *parametrized surface* in $\mathbb{R}^3$ is a surface defined by a vector-valued function $\mathbf{r}(u,v)$ of two parameters u and v, such that:

$$\mathbf{r}(u,v) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix}, \quad (u,v) \in D \subset \mathbb{R}^2.$$

The function $\mathbf{r}(u,v)$ gives a mapping from the parameter domain D onto the surface in $\mathbb{R}^3$.

Example

The parametric equations for a sphere of radius R centered at the origin are given by:

$$\mathbf{r}(u,v) = \begin{bmatrix} R\sin u\cos v \\ R\sin u\sin v \\ R\cos u \end{bmatrix},$$

where $0 \leq u \leq \pi$ and $0 \leq v < 2\pi$ are the parameters.

Remark

Parametrizing a surface allows us to compute surface integrals and evaluate geometric quantities such as area. The differential surface element dS can be computed from the parametrization.

Definition (Scalar Surface Integral)

Let S be a surface in $\mathbb{R}^3$ parametrized by $\mathbf{r}(u,v)$, and let $f(x,y,z)$ be a scalar function defined on the surface. The *scalar surface integral* of f over S is given by:

$$\int_S f(x,y,z)\,dS = \int_D f(\mathbf{r}(u,v))\,|\mathbf{r}_u \times \mathbf{r}_v|\,du\,dv,$$

where $\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u}$ and $\mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}$ are the partial derivatives of the parametrization.

Remark

The magnitude of the cross product $\mathbf{r}_u \times \mathbf{r}_v$ gives the area of the parallelogram formed by the tangent vectors at each point on the surface. Thus, $|\mathbf{r}_u \times \mathbf{r}_v|\,du\,dv$ is the differential element of surface area.

Example

Consider the unit sphere S^2 parametrized by spherical coordinates as:

$$\mathbf{r}(u,v) = \begin{bmatrix} \sin u \cos v \\ \sin u \sin v \\ \cos u \end{bmatrix},$$

where $u \in [0,\pi]$ and $v \in [0,2\pi]$. The differential element of surface area is:

$$dS = |\mathbf{r}_u \times \mathbf{r}_v|\, du\, dv = \sin u\, du\, dv.$$

The surface area of the sphere is then:

$$\int_{S^2} dS = \int_0^{\pi} \int_0^{2\pi} \sin u\, dv\, du = 4\pi.$$

Definition (Vector Surface Integral)

Let $\mathbf{F} = \begin{bmatrix} F_1(x,y,z) \\ F_2(x,y,z) \\ F_3(x,y,z) \end{bmatrix}$ be a vector field defined on a surface S. The *vector surface integral*, or the flux of $\mathbf{F}$ across S, is given by:

$$\int_S \mathbf{F} \cdot d\mathbf{S} = \int_D \mathbf{F}(\mathbf{r}(u,v)) \cdot (\mathbf{r}_u \times \mathbf{r}_v)\, du\, dv.$$

Remark

The vector surface integral computes the total flux of the vector field $\mathbf{F}$ through the surface S. The cross product $\mathbf{r}_u \times \mathbf{r}_v$ represents a normal vector to the surface, ensuring that we integrate in the direction perpendicular to the surface.

Example

Consider the vector field $\mathbf{F}(x,y,z) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and the unit sphere S^2 as described above. The flux through the surface is:

$$\int_{S^2} \mathbf{F} \cdot d\mathbf{S} = \int_0^{\pi} \int_0^{2\pi} \mathbf{r}(u,v) \cdot (\sin u \mathbf{k})\, dv\, du,$$

where $\mathbf{k}$ is the unit normal vector to the surface. Evaluating this integral gives zero, indicating that the flux of $\mathbf{F}$ through the closed surface of the sphere is zero, consistent with the divergence theorem for this field.

Proposition (Additivity)

The surface integral is additive. If a surface S is divided into two disjoint parts S_1 and S_2, then:

$$\int_S f(x,y,z)\,dS = \int_{S_1} f(x,y,z)\,dS + \int_{S_2} f(x,y,z)\,dS.$$

Proposition (Orientation Dependence)

The vector surface integral depends on the orientation of the surface. Reversing the orientation of the surface changes the sign of the integral:

$$\int_S \mathbf{F}\cdot d\mathbf{S} = -\int_S \mathbf{F}\cdot d\mathbf{S}_{\text{reversed}}.$$

Remark

In vector surface integrals, the orientation of the surface (given by the direction of the normal vector) is crucial. Changing the direction of the normal vector will reverse the sign of the flux through the surface.

In vector calculus, integral theorems play a central role in relating the behavior of a vector field in a region to its behavior on the boundary of that region. Two important integral theorems are Stokes' Theorem and Gauss's (Divergence) Theorem.

Theorem (Stokes' Theorem)

Stokes' Theorem relates the surface integral of the curl of a vector field over a surface S to the line integral of the vector field $\mathbf{F}$ around the boundary C of the surface:

$$\int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint_C \mathbf{F} \cdot d\mathbf{r}.$$

Here:

- $\mathbf{F}$ *is a continuously differentiable vector field.*
- S *is a smooth, oriented surface with boundary* C.
- $d\mathbf{S}$ *is the oriented surface element.*
- $d\mathbf{r}$ *is the differential element along the curve* C.

Remark

Stokes' Theorem provides a direct connection between a surface integral and a line integral. It can be seen as a generalization of the fundamental theorem of calculus to higher dimensions.

Example

Consider a vector field $\mathbf{F}(x,y,z) = (-y,x,0)$ and let S be the disk in the xy-plane with radius R and center at the origin, oriented upward. The boundary C of S is the circle $x^2+y^2=R^2$ in the plane. According to Stokes' Theorem:

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_S (\nabla\times\mathbf{F})\cdot d\mathbf{S}.$$

First, compute the curl of $\mathbf{F}$:

$$\nabla\times\mathbf{F} = (0,0,2).$$

The surface integral becomes:

$$\int_S (\nabla\times\mathbf{F})\cdot d\mathbf{S} = \int_S 2\,dS = 2\pi R^2.$$

Now compute the line integral around C. Parameterize the boundary curve C as $\mathbf{r}(t) = (R\cos t, R\sin t, 0)$ with $t\in[0,2\pi]$. Then:

$$\mathbf{F}\cdot d\mathbf{r} = (-R\sin t, R\cos t, 0)\cdot(-R\sin t, R\cos t, 0)\,dt = R^2\,dt.$$

The line integral is:

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} R^2\,dt = 2\pi R^2.$$

Thus, Stokes' Theorem holds, as both the surface and line integrals yield the same result.

Theorem (Gauss's Theorem (Divergence Theorem))

Gauss's Theorem, also known as the Divergence Theorem, relates the flux of a vector field $\mathbf{F}$ *through a closed surface* S *to the volume integral of the divergence of* $\mathbf{F}$ *over the region* V *enclosed by* S*:*

$$\int_V (\nabla \cdot \mathbf{F})\, dV = \oint_S \mathbf{F} \cdot d\mathbf{S}.$$

Here:

- $\mathbf{F}$ *is a continuously differentiable vector field.*
- V *is a volume in* $\mathbb{R}^3$ *enclosed by the surface* S*.*
- $\nabla \cdot \mathbf{F}$ *is the divergence of* $\mathbf{F}$*.*
- $d\mathbf{S}$ *is the outward-oriented surface element.*

Remark

Gauss's Theorem connects the flux through a closed surface to the behaviour of the vector field inside the surface. It is a generalization of the divergence form of the fundamental theorem of calculus to three dimensions.

Example

Consider the vector field $\mathbf{F}(x,y,z) = (x,y,z)$ and let V be the unit sphere $x^2+y^2+z^2 \leq 1$. The surface S is the boundary of V, the unit sphere itself.
First, compute the divergence of $\mathbf{F}$:

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}x + \frac{\partial}{\partial y}y + \frac{\partial}{\partial z}z = 3.$$

The volume integral is:

$$\int_V (\nabla \cdot \mathbf{F})\, dV = 3\int_V dV = 3 \times \frac{4}{3}\pi(1^3) = 4\pi.$$

Now compute the surface integral:

$$\oint_S \mathbf{F} \cdot d\mathbf{S}.$$

For the unit sphere, the outward normal vector at any point is simply $\hat{n} = (x,y,z)$, and $dS = 1\, dA$, where dA is the differential surface area element. Thus:

$$\mathbf{F} \cdot d\mathbf{S} = (x,y,z) \cdot (x,y,z)\, dA = (x^2+y^2+z^2)\, dA = 1\, dA.$$

The surface integral is:

$$\oint_S dA = 4\pi.$$

Hence, Gauss's Theorem holds, as both the volume and surface integrals yield the same result.

Mathematical Methods for Bioengineering

Biomedic Engineering Degree

Javier Quintero

Slides – Material docente en abierto de la Universidad Rey Juan Carlos
Avaible at Institutional Open Archive Universidad Rey Juan Carlos (BURJC Digital)
https://burjcdigital.urjc.es/home

2024 – 2025 v1.02