Advanced Calculus

Semester 2, 2023

Dr Pascal Buenzli

Tarang Janawalkar

Contents

1 Euclidean Space

The Euclidean space $\mathbb{R}^n$ is an n-dimensional vector space of real numbers. This space is closed under addition and scalar multiplication.

1.1 Operations

1.1.1 Addition

The sum of two vectors $\mathbf{x}$ and $\mathbf{y}$ is defined element-wise

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

In a coordinate system, the vectors $\mathbf{x}$ and $\mathbf{y}$ are added tip-to-tail.

1.1.2 Scalar Multiplication

The scalar multiplication of a vector $\mathbf{x}$ by a scalar $\lambda \in \mathbb{R}$ is defined element-wise

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

In a coordinate system, λ scales the vector $\mathbf{x}$ along the same line.

1.1.3 Norm

The norm (length) of a vector $\mathbf{x}$ is defined as

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

The norm of a vector $\mathbf{x}$ is the distance from the origin to the tip of the vector. This allows us to define the unit vector $\hat{\mathbf{x}}$ as

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$$

which is a vector of length 1 in the same direction as $\mathbf{x}$.

1.1.4 Scalar Product

The scalar product (dot product) of two vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

The scalar product allows us to define the angle θ between two vectors $\mathbf{x}$ and $\mathbf{y}$ as

$$\cos(\theta) = \hat{\mathbf{x}} \cdot \hat{\mathbf{y}}$$

where we use the unit vectors of $\mathbf{x}$ and $\mathbf{y}$, as the angle between two vectors is invariant under scaling. Additionally, we can determine the projection of the vector $\mathbf{x}$ onto the vector $\mathbf{y}$ using trigonometry

$$\text{proj}_{\mathbf{y}}(\mathbf{x}) = (\|\mathbf{x}\| \cos(\theta))\,\hat{\mathbf{y}} = (\|\mathbf{x}\|\,(\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}))\,\hat{\mathbf{y}} = (\mathbf{x} \cdot \hat{\mathbf{y}})\,\hat{\mathbf{y}}$$

where $\mathbf{x} \cdot \hat{\mathbf{y}}$ is the norm of the projection vector.

1.2 Additional Properties

1.2.1 Triangle Inequality

$$\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$$

1.2.2 Inverse Triangle Inequality

$$\|\mathbf{x} - \mathbf{y}\| \geqslant |\|\mathbf{x}\| - \|\mathbf{y}\||$$

1.2.3 Cauchy-Schwarz Inequality

$$|\mathbf{x} \cdot \mathbf{y}| \leqslant \|\mathbf{x}\|\|\mathbf{y}\|$$

1.3 Multivariable Functions

A multivariable function f maps a vector $\mathbf{x} \in \mathbb{R}^n$ to a real number $f(\mathbf{x}) \in \mathbb{R}$. This function can be expressed in **explicit form** as

$$z = f(x,\, y)$$

or in **implicit form** as

$$F(x,\, y,\, z) = z - f(x,\, y)$$

These equations define a surface in $\mathbb{R}^3$.

1.3.1 Level Curves

The level curves of a function $f(x,\, y)$ are the curves in $\mathbb{R}^2$ where

$$f(x,\, y) = c$$

where c is the height of the curve. Implicitly, this is equivalent to

$$F(x,\, y,\, z) = 0.$$

Level curves represent paths of equal height on the surface defined by $z = f(x,\, y)$.

1.4 Special Regions

1.4.1 Balls

In an Euclidean space, an open ball of radius $r > 0$ centred at a point $\mathbf{p} \in \mathbb{R}^n$ is denoted $B_r(\mathbf{p})$, and is defined as

$$B_r(\mathbf{p}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{p}\| < r\}.$$

This region includes all points less than a distance r from the vector $\mathbf{p}$, where the distance is typically defined by the L_2-norm:

$$\|\mathbf{x} - \mathbf{p}\|_2 = \left(\sum_{i=1}^{n} (x_i - p_i)^2 \right)^{1/2}.$$

1.5 Mathematical Representation of Curves

1.5.1 Explicit Form

A curve in $\mathbb{R}^2$ can be represented in explicit form as

$$y = f(x)$$

but this is not possible in $\mathbb{R}^3$ as a 3D curve requires two equations. For a 2D explicit curve:

- x is an independent variable such that we have 1 degree of freedom.

1.5.2 Implicit Form

A curve in $\mathbb{R}^2$ can be represented in implicit form as

$$F(x,\ y) = 0.$$

In 3D, we must impose an additional equation that intersects a surface.

$$\begin{cases} F(x,\ y,\ z) = 0 \\ G(x,\ y,\ z) = 0 \end{cases}$$

In both cases, we have 1 degree of freedom as the degrees of freedom is the difference between the number of variables and the number of equations.

1.5.3 Parametric Form

In parametric form, curves are parametrised in terms of a parameter t. In 2D, this is represented as

$$\mathbf{r}(t) = \langle x(t),\ y(t) \rangle$$

and similarly in 3D,

$$\mathbf{r}(t) = \langle x(t),\ y(t),\ z(t) \rangle$$

1.6 Converting Between Representations

1.6.1 Explicit to Implicit

The equation $y = f(x)$ can always be converted to implicit form by rewriting it as

$$F(x,\ y) = y - f(x) = 0.$$

1.6.2 Implicit to Explicit

The equation $F(x,\ y) = 0$ can be converted to explicit form if we can solve for y (or x):

1.6.3 Parametric to Explicit/Implicit

The equation $\mathbf{r}(t) = \langle x(t),\ y(t)\rangle$ can be written in explicit or implicit form, if the parameter t can be eliminated from the simultaneous equations.

1.6.4 Explicit to Parametric

The equation $y = f(x)$ can always be converted to parametric form by choosing the parameter $t = x$, so that

$$\mathbf{r}(t) = \langle t,\ f(t)\rangle.$$

1.6.5 Implicit to Parametric

The equation $F(x,\ y) = 0$ can be converted to parametric form if we can find $x = p(t)$ and $y = q(t)$, such that $F(p(t),\ q(t)) = 0$, and

$$\mathbf{r}(t) = \langle p(t),\ q(t)\rangle$$

for all t.

1.7 Paramaterisation

To parametrise a curve, consider the following strategies:

- For a closed curve, consider the polar parametrisation in terms of the angle θ:

$$\mathbf{r}(\theta) = \langle R(\theta)\cos(\theta),\ R(\theta)\sin(\theta)\rangle.$$

- For a curve that is the intersection of two surfaces, consider one of the following mappings:

$$x \mapsto \begin{bmatrix} x \\ y(x) \\ z(x) \end{bmatrix} \qquad y \mapsto \begin{bmatrix} x(y) \\ y \\ z(y) \end{bmatrix} \qquad z \mapsto \begin{bmatrix} x(z) \\ y(z) \\ z \end{bmatrix}$$

- Otherwise, consider a vector construction.

1.7.1 Line Segments

To parametrise a line segment from point A to B, define the parameter $t \in [0, 1]$. Then, consider the vectors $\mathbf{a} = \overline{OA}$ and $\mathbf{b} = \overline{OB}$. By scaling the vector from A to B by t, we can parametrise the line segment as

$$\mathbf{r}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) = \mathbf{a}(1 - t) + \mathbf{b}t.$$

1.7.2 Circles

To parametrise a circle of radius R centred at the $\langle x_0,\, y_0 \rangle$, first parametrise the curve in terms of the angle θ, then shift the curve by $\langle x_0,\, y_0 \rangle$.

$$\mathbf{r}(\theta) = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} R\cos(\theta) \\ R\sin(\theta) \end{bmatrix} = \begin{bmatrix} x_0 + R\cos(\theta) \\ y_0 + R\sin(\theta) \end{bmatrix}$$

1.7.3 Velocity Vectors

The velocity vector of a parametrised curve $\mathbf{r}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ is defined as

$$\mathbf{v}(t) = \mathbf{r}'(t) = \lim_{\Delta t \to 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$$

where $\mathbf{v}(t)$ is a tangent vector to the curve at the point $\mathbf{r}(t)$, for all t.

1.7.4 Tangent Vectors

Following from the definition of the velocity vector, the tangent vectors of a parametrised curve are unit vectors in the direction of the velocity vector.

$$\hat{\boldsymbol{\tau}}(t) = \pm \frac{\mathbf{v}(t)}{\|\mathbf{v}(t)\|} = \pm \mathbf{v}\hat{(t)}$$

For a curve given in explicit form $y = f(x)$, the tangent vectors are given by

$$\hat{\boldsymbol{\tau}}(x) = \pm \frac{1}{\sqrt{1 + (f'(x))^2}} \begin{bmatrix} 1 \\ f'(x) \end{bmatrix}$$

2 Multivariable Calculus

2.1 Multivariable Functions

Definition 2.1 (Multivariable Function)**.** A multivariable function f maps several independent variables to a real number:

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

where

$$\langle x_1,\, x_2,\, \ldots,\, x_n \rangle \mapsto z = f(x_1,\, x_2,\, \ldots,\, x_n)$$
$$\mathbf{x} \mapsto z = f(\mathbf{x})$$

Definition 2.2 (Domain)**.** The domain of f is the subset of $\mathbb{R}^n$ for which f is defined. It corresponds to the set of all possible inputs to f.

$$\mathscr{D}(f) = E$$

Definition 2.3 (Range)**.** The range of f is the image of the domain E under f:

$$\mathscr{R}(f) = \{f(\mathbf{x}) \in \mathbb{R} : \mathbf{x} \in E\}$$

Definition 2.4 (Graph)**.** The graph of f is the defined as the set

$$G = \{\langle \mathbf{x},\ f(\mathbf{x})\rangle : \mathbf{x} \in E\} \subset \mathbb{R}^{n+1}$$

2.2 Curves of Intersection

Definition 2.5 (Curves of Intersection)**.** The curves of intersection of f with the plane $x_i = c$ are defined as

$$\{\langle x_1,\ x_2,\ \ldots,\ x_{i-1},\ c,\ x_{i+1},\ \ldots,\ x_n\rangle : \langle x_1,\ x_2,\ \ldots,\ x_n\rangle \in E\}$$

In 3D, curves of intersection of $z = f(x,\ y)$ with the planes perpendicular to the x, y, or z-axes allow us to represent the function in 3D.

$$\begin{array}{ccc} \perp x\ (x = c) & \perp y\ (y = c) & \perp z\ (z = c) \\ z = f(c,\ y) = g(y) & z = f(x,\ c) = g(x) & c = f(x,\ y) \implies y = g(x) \end{array}$$

Definition 2.6 (Level Set)**.** The level sets of f are the set of all points in the domain of f that map to a given value c.

$$\{\mathbf{x} \in E : f(\mathbf{x}) = c\}$$

In 2D, level sets are called **level curves**.

Definition 2.7 (Contour Map)**.** The projection of all level curves onto the xy-plane is called the **contour map** of f. T he lines of a contour map are called **contours**.

2.3 Derivatives

Definition 2.8 (Continuity)**.** A function f is continuous at a point $\mathbf{x}_0$ if

$$\forall \varepsilon > 0,\ \exists \delta > 0 : \|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon$$

Equivalently, f is continuous at $\mathbf{x}_0$ when

$$\lim_{n\to\infty} f(\mathbf{x}_n) = f(\mathbf{x}_0)$$

for all sequences $\{\mathbf{x}_n\}$ where $\mathbf{x}_n \to \mathbf{x}_0$.
A function f is continuous on a set E if it is continuous at every point in E.

Definition 2.9 (Partial Derivatives)**.** Partial derivatives represent the rate of change of a function with respect to one of its variables, holding all other variables constant.

$$\frac{\partial f}{\partial x_i} = f_{x_i} = \lim_{h\to 0} \frac{f(x_1,\ x_2,\ \ldots,\ x_i + h,\ \ldots,\ x_n) - f(x_1,\ x_2,\ \ldots,\ x_i,\ \ldots,\ x_n)}{h}$$

Definition 2.10 (Higher-order Partial Derivatives)**.** Higher-order partial derivatives are defined as

$$\frac{\partial^n f}{\partial x_i^n} = \frac{\partial}{\partial x_i}\left(\frac{\partial^{n-1} f}{\partial x_i^{n-1}}\right)$$

Definition 2.11 (Mixed Partial Derivatives)**.** Mixed partial derivatives are given the following notation

$$\frac{\partial^2 f}{\partial x_i \, \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = f_{x_i x_j}$$

Theorem 2.3.1 (Schwarz's Theorem (or Clairaut's Theorem))**.** *For a function $f : E \subset \mathbb{R}^2 \to \mathbb{R}$ with $\mathbf{x}_a \in E$, if f_{xy} and f_{yx} are continuous on $\mathbf{x}_0$, then*

$$f_{xy} = f_{yx}$$

Remark 1. For an arbitrary n and $k \leqslant n$, if all partials of order $\leqslant k$ are continuous in the neighbourhood of $\mathbf{x}_0$, then mixed partials of order k are equal for any permutation of indices:

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

2.4 Chain Rule

For a function with multiple arguments, each argument of f that has an implicit dependence on the variable of differentiation must be differentiated using the chain rule, where all contributions are summed.

$$\frac{\mathrm{d}}{\mathrm{d}t} f\left(x\left(t\right),\, y\left(t\right)\right) = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

$$\frac{\partial}{\partial u} f\left(x\left(u,\, v\right),\, y\left(u,\, v\right)\right) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}$$

Definition 2.12 (Total derivative)**.** The total derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial f}{\partial x_1}\frac{\mathrm{d}x_1}{\mathrm{d}t} + \frac{\partial f}{\partial x_2}\frac{\mathrm{d}x_2}{\mathrm{d}t} + \cdots + \frac{\partial f}{\partial x_n}\frac{\mathrm{d}x_n}{\mathrm{d}t}$$

Proof. Consider the case where $n = 2$. The total derivative can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} f\left(x\left(t\right),\, y\left(t\right)\right) = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[f\left(x\left(t + \Delta t\right),\, y\left(t + \Delta t\right)\right) - f\left(x\left(t\right),\, y\left(t\right)\right)\right]$$

Using the substitutions $x\left(t + \Delta t\right) = x\left(t\right) + \Delta x$ and $y\left(t + \Delta t\right) = y\left(t\right) + \Delta y$,

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}t} &= \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[f\left(x + \Delta x,\, y + \Delta y\right) - f\left(x,\, y\right)\right] \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[f\left(x + \Delta x,\, y + \Delta y\right) - f\left(x,\, y + \Delta y\right) + f\left(x,\, y + \Delta y\right) - f\left(x,\, y\right)\right] \\
&= \lim_{\Delta t \to 0}\left[\frac{\Delta x}{\Delta t}\frac{f\left(x + \Delta x,\, y + \Delta y\right) - f\left(x,\, y + \Delta y\right)}{\Delta x} + \frac{\Delta y}{\Delta t}\frac{f\left(x,\, y + \Delta y\right) - f\left(x,\, y\right)}{\Delta y}\right] \\
&= \frac{\mathrm{d}x}{\mathrm{d}t}\frac{\partial f}{\partial x} + \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\partial f}{\partial y}
\end{aligned}$$

□

Definition 2.13 (Gradient)**.** The gradient is an operator that collects all partial derivatives of a function into a vector.

$$\boldsymbol{\nabla} = \begin{bmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \vdots \\ \partial_{x_n} \end{bmatrix} \implies \boldsymbol{\nabla} f = \begin{bmatrix} \partial_{x_1} f \\ \partial_{x_2} f \\ \vdots \\ \partial_{x_n} f \end{bmatrix}$$

The gradient allows us to define the chain rule using the dot product:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \boldsymbol{\nabla} f \cdot \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$$

2.5 Directional Derivative

The directional derivative of a function $f : E \subset \mathbb{R}^n \to \mathbb{R}$ at a point $\mathbf{x}_0 \in E$ in the direction of a unit vector $\hat{\mathbf{u}} \in \mathbb{R}^n$ is the slope of f in the direction of $\mathbf{u}$:

$$D_{\mathbf{u}} f(\mathbf{x}_0) = \partial_{\mathbf{u}} f(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\hat{\mathbf{u}}) - f(\mathbf{x}_0)}{h}$$

Proposition 2.5.1. *The directional derivative can be computed using the gradient of f*

$$D_{\mathbf{u}} f(\mathbf{x}_0) = \boldsymbol{\nabla} f(\mathbf{x}_0) \cdot \hat{\mathbf{u}}.$$

Proof. Consider the path in the output space $g(s) = f(\mathbf{x}_0 + s\hat{\mathbf{u}}) \in \mathbb{R}$ for $s \in \mathbb{R}$. The derivative of g w.r.t. s is given by:

$$\frac{\mathrm{d}g}{\mathrm{d}s} = \lim_{h \to 0} \frac{g(s+h) - g(s)}{h}$$

where if we evaluate g at $s = 0$, we get

$$\left.\frac{\mathrm{d}g}{\mathrm{d}s}\right|_{s=0} = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\hat{\mathbf{u}}) - f(\mathbf{x}_0)}{h} = D_{\mathbf{u}} f(\mathbf{x}_0).$$

Using the derivative of the parametrised vector $\mathbf{x} = \mathbf{x}_0 + s\hat{\mathbf{u}} \in \mathbb{R}^n$:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = \hat{\mathbf{u}}$$

we can evaluate the derivative of g w.r.t. s using the total derivative:

$$\begin{aligned} \left.\frac{\mathrm{d}g}{\mathrm{d}s}\right|_{s=0} &= \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_{s=0} \left.\frac{\mathrm{d}x_i}{\mathrm{d}s}\right|_{s=0} \\ &= \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_{s=0} \hat{u}_i \\ &= \boldsymbol{\nabla} f(\mathbf{x})_{s=0} \cdot \hat{\mathbf{u}} \\ &= \boldsymbol{\nabla} f(\mathbf{x}_0) \cdot \hat{\mathbf{u}} \end{aligned}$$

where $\hat{u}_i$ is the i-th component of $\hat{\mathbf{u}}$. Therefore $D_{\mathbf{u}} f(\mathbf{x}_0) = \boldsymbol{\nabla} f(\mathbf{x}_0) \cdot \hat{\mathbf{u}}$. □

Remark 1. Partial derivatives are directional derivatives in the direction of the canonical basis vectors $\hat{\mathbf{e}}_i$.

$$\frac{\partial f}{\partial x_i} = D_{\hat{\mathbf{e}}_i} f$$

We can therefore say that the directional derivative is a generalisation of the partial derivative for any direction $\hat{\mathbf{u}}$.

Proposition 2.5.2. *The gradient of a function $\boldsymbol{\nabla} f$ is orthogonal to the level curves of f.*

Proof. Consider the path $\mathbf{x}(s)$ on a contour of f. As f is constant on the contour,

$$\frac{\partial f}{\partial s} = 0.$$

Using the chain rule,

$$\frac{\partial f}{\partial s} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial s} = \boldsymbol{\nabla} f \cdot \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = 0$$

therefore as the dot product is zero, $\boldsymbol{\nabla} f$ is orthogonal to the path $\mathbf{x}(s)$. □

Proposition 2.5.3. *The directional derivative is maximised when $\hat{\mathbf{u}}$ is parallel to $\boldsymbol{\nabla} f$.*

Proof. Using the angle definition of the dot product, the directional derivative is given by

$$D_{\mathbf{u}} f = \boldsymbol{\nabla} f \cdot \hat{\mathbf{u}} = \|\boldsymbol{\nabla} f\| \|\hat{\mathbf{u}}\| \cos\theta = \|\boldsymbol{\nabla} f\| \cos\theta$$

This expression is maximised when $\cos\theta = 1$, or when $\mathbf{u}$ is parallel to $\boldsymbol{\nabla} f$. □

Remark 2. The maximum slope of f at $\mathbf{x}_0$ is given by the magnitude of the gradient at $\mathbf{x}_0$:

$$\max_{\hat{\mathbf{u}}} D_{\mathbf{u}} f(\mathbf{x}_0) = \|\boldsymbol{\nabla} f(\mathbf{x}_0)\|$$

2.6 Normal Vectors to Curves

2.6.1 Parametric Curves

For a parametric curve $\mathbf{r}(t)$, we find a normal vector $\hat{\mathbf{n}}$ such that

$$\hat{\mathbf{n}} \cdot \mathbf{r}'(t) = 0$$

where $\mathbf{r}'(t)$ is a tangent vector to the curve.

2.6.2 Implicit Curves

For an implicit curve $F(x,\, y) = 0$, the normal vector is given by

$$\hat{\mathbf{n}} = \pm \frac{\boldsymbol{\nabla} F}{\|\boldsymbol{\nabla} F\|}$$

2.6.3 Explicit Curves

For an explicit curve $y = f(x)$, we must convert the curve to implicit or parametric form.

2.7 Normal Vectors to Surfaces

2.7.1 Parametric Surfaces

2.7.2 Implicit Surfaces

For an implicit surface $F(x,\ y,\ z) = 0$, the normal vector is given by

$$\hat{\mathbf{n}} = \pm\frac{\boldsymbol{\nabla} F}{\|\boldsymbol{\nabla} F\|}$$

2.7.3 Explicit Surfaces

For an explicit surface $z = f(x,\ y)$, we can either convert the surface to implicit form, or consider the tangents of two curves $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$ on the surface. Then we can find the normal vector by taking the cross product of the tangent vectors:

$$\hat{\mathbf{n}} = \pm\frac{\mathbf{r}_1'(t) \times \mathbf{r}_2'(t)}{\|\mathbf{r}_1'(t) \times \mathbf{r}_2'(t)\|}$$

2.8 Tangent Vectors to Curves

2.8.1 Parametric Curves

For a parametric curve $\mathbf{r}(t)$, the tangent vector is given by

$$\hat{\boldsymbol{\tau}} = \pm\frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$$

2.8.2 Implicit Curves in 2D

For an implicit curve $F(x,\ y) = 0$, the tangent vector can be found by first determining the normal vector $\hat{\mathbf{n}}$ which is proportional to the gradient of F:

$$\hat{\mathbf{n}} = \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \propto \boldsymbol{\nabla} f$$

such that the tangent vector is given by

$$\hat{\boldsymbol{\tau}} = \pm\begin{bmatrix} -n_2 \\ n_1 \end{bmatrix}$$

2.8.3 Implicit Curves in 3D

Given the intersectin of two implicit curves $F(x,\ y,\ z) = 0$ and $G(x,\ y,\ z) = 0$, the tangent vector along the intersection is given by

$$\hat{\boldsymbol{\tau}} = \pm\frac{\boldsymbol{\nabla} F \times \boldsymbol{\nabla} G}{\|\boldsymbol{\nabla} F \times \boldsymbol{\nabla} G\|}$$

2.8.4 Explicit Curves in 2D

For an explicit curve $y = f(x)$, we can either:

- convert the curve to parametric form: $\mathbf{r}(t) = \langle t,\ f(t) \rangle$
- convert the curve to implicit form: $F(x,\ y) = y - f(x) = 0$

2.9 Differentiability

2.10 Single Variable Functions

$f(x)$ is differentiable at a if $f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$ exists. This is equivalent to saying that the tangent line at $\langle a,\ f(a) \rangle$ is well-defined:

$$f(x) = \underbrace{f(a) + f'(a)\, f(x - a)}_{\text{tangent line}} + R(x - a) \qquad \text{with } \lim_{x \to a} \frac{R(x - a)}{x - a} = 0$$

2.11 Multivariable Functions

$f(\mathbf{x})$ is differentiable at $\mathbf{a} \in \Omega \subset \mathbb{R}^n$ if there exists a linear map $f'(\mathbf{a}) : \begin{cases} \mathbb{R}^n \to \mathbb{R} \\ \xi \mapsto f'(\mathbf{a}) \cdot \xi \end{cases}$ such that

$$f(\mathbf{x}) = \underbrace{f(\mathbf{a}) + f'(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})}_{\text{tangent plane}} + R(\mathbf{x} - \mathbf{a}) \qquad \text{with } \lim_{\mathbf{x} \to \mathbf{a}} \frac{R(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|} = 0$$

The linear map $f'(\mathbf{a})$ is then the derivative of f at $\mathbf{a}$.

Theorem 2.11.1 (Derivative Equivalence with Gradient)**.** *If f is differentiable, then*

$$f'(\mathbf{a}) = \boldsymbol{\nabla} f(\mathbf{a})$$

Additionally, if $\boldsymbol{\nabla} f(\mathbf{a})$ exists $\forall \mathbf{a} \in \Omega$, and all partial derivatives are continuous (i.e., $\frac{\partial f}{\partial x_i} : \Omega \subset \mathbb{R}^n \to R$ are continuous), then f is differentiable everywhere in Ω.

The first result tells us that the derivative is unique, and that f is differentiable at $\mathbf{a}$ if it has a tangent plane at $\mathbf{a}$. The second result gives us a sufficient condition for differentiability.

2.12 Taylor Series Expansion

The Taylor series expansion of a function f at a point $\mathbf{a}$ is given by

$$f(\mathbf{x}) = f(\mathbf{a}) + \sum_{i=0}^{n} \frac{\partial f(\mathbf{a})}{\partial x_i} (x_i - a_i) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(\mathbf{a})}{\partial x_i\, \partial x_j} (x_i - a_i)\left(x_j - a_j\right) + \cdots$$
$$+ \frac{1}{k!} \sum_{i_1, \ldots, i_k}^{n} \frac{\partial^3 f(\mathbf{a})}{\partial x_{i_1}\, \partial \cdots \partial x_{i_k}} \left(x_{i_1} - a_{i_1}\right) \cdots \left(x_{i_k} - a_{i_k}\right) + \cdots$$

This allows us to compute the tangent plane, paraboloid, and so on, of a function f at $\mathbf{a}$, by increasing the order k of the Taylor series expansion.

3 Double Integrals

Integrals represent continuous sums of infinitesimal quantities, and allow us to measure extensive and average properties of objects, such as lengths, surface areas, volumes, masses, centres of mass, and so on.

3.1 Riemann Sums

The signed area under the function $f(x)$ is represented by

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\,\Delta x$$

Similarly, the signed volume under the surface $f(x,\ y)$ is represented by

$$\iint_R f(x,\ y)\,\mathrm{d}A = \lim_{\Delta A_i \to 0} \sum_{i=1}^{n} f(x_i,\ y_i)\,\Delta A_i$$

where R is the region of integration.

3.2 Lebesgue Integrals

Lebesgue integrals are a top down approach to multiple integrals that allow us to integrate a wider class of nonnegative functions. This includes improper integrals where f is discontinuous or singular, or when the domain of integration is unbounded.

Theorem 3.2.1 (Fubini's Theorem)**.** *For a nonnegative function $f : \mathbb{R}^2 \to [0, \infty)$, the following equality holds,*

$$0 \leqslant \iint_{\mathbb{R}^2} f(x,\ y)\,\mathrm{d}A = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,\ y)\,\mathrm{d}x\right] \mathrm{d}y = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,\ y)\,\mathrm{d}y\right] \mathrm{d}x \leqslant \infty$$

The above theorem holds for particular classes of functions. The following techniques can be used to integrate arbitrary functions.

3.2.1 Finite Domains

For double integrals over domains $R \subset \mathbb{R}^2$, we can define a new function $\tilde{f}$ such that

$$\iint_R f(x,\ y)\,\mathrm{d}A \equiv \iint_{\mathbb{R}^2} \tilde{f}(x,\ y)\,\mathrm{d}A, \qquad \text{where } \tilde{f}(x,\ y) = \begin{cases} f(x,\ y) & (x,\ y) \in R \\ 0 & \text{otherwise} \end{cases}$$

3.2.2 Nonpositive Functions

For nonpositive functions $f : \mathbb{R}^2 \to (-\infty, \infty)$, we can express f as $f = f^+ - f^-$, where f^+ and f^- are the positive and negative parts of f respectively:

$$f^+(x,\, y) = \max\{0,\, f(x,\, y)\} = \begin{cases} f(x,\, y) & f(x,\, y) \geqslant 0 \\ 0 & f(x,\, y) < 0 \end{cases}$$

$$f^-(x,\, y) = -\min\{0,\, f(x,\, y)\} = \begin{cases} -f(x,\, y) & f(x,\, y) < 0 \\ 0 & f(x,\, y) \geqslant 0 \end{cases}$$

Therefore, the double integral is defined

$$\iint_{\mathbb{R}^2} f(x,\, y)\, \mathrm{d}A = \iint_{\mathbb{R}^2} f^+(x,\, y)\, \mathrm{d}A - \iint_{\mathbb{R}^2} f^-(x,\, y)\, \mathrm{d}A$$

Definition 3.1 (Integrable Function)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is **integrable** if

$$\int_{\mathbb{R}^n} |f(\mathbf{x})|\, \mathrm{d}\mathbf{x} < \infty$$

Theorem 3.2.2 (Fubini's Theorem for Integrable Functions)**.** *If $f : \mathbb{R}^2 \to \mathbb{R}$ is an integrable function*

$$\iint_{\mathbb{R}^2} f(x,\, y)\, \mathrm{d}A = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,\, y)\, \mathrm{d}x \right] \mathrm{d}y = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,\, y)\, \mathrm{d}y \right] \mathrm{d}x$$

3.3 Measure of a Region

Definition 3.2 (Indicator Function)**.** The indicator function of R is defined:

$$1_R(x,\, y) = \begin{cases} 1 & (x,\, y) \in R \\ 0 & \text{otherwise} \end{cases}$$

Definition 3.3 (Measure)**.** The measure of a region R is defined as

$$\mu(R) = \iint_{\mathbb{R}^2} 1_R\, \mathrm{d}A$$

In two dimensions, the measure is the area of the region.

3.3.1 Rectangular Regions

For a rectangular region $R = [a,\, b] \times [c,\, d]$, the indicator function is given by

$$\begin{aligned} 1_R(x,\, y) &= \begin{cases} 1 & (x,\, y) \in D \\ 0 & \text{otherwise} \end{cases} \\ &= 1_{[a,\, b]}(x)\, 1_{[c,\, d]}(y) \end{aligned}$$

so that the area of R is given by

$$\mu(R) = \iint_{\mathbb{R}^2} 1_R \,\mathrm{d}A = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{[a,\,b]}(x)\, 1_{[c,\,d]}(y)\,\mathrm{d}y\,\mathrm{d}x = \int_{-\infty}^{\infty} 1_{[a,\,b]}(x) \int_{-\infty}^{\infty} 1_{[c,\,d]}(y)\,\mathrm{d}y\,\mathrm{d}x$$
$$= \int_a^b 1 \left[\int_c^d 1\,\mathrm{d}y\right]\mathrm{d}x = \int_a^b (d-c)\,\mathrm{d}x = (b-a)(d-c)$$

3.3.2 Integrating Functions over Regions

When integrating over a nonuniform density f, we can multiply f with the indicator function over R:

$$\iint_R f(x,\,y)\,\mathrm{d}A = \iint_{\mathbb{R}^2} f(x,\,y)\,1_R\,\mathrm{d}A$$

3.4 Simple Domains

3.4.1 Type I — y-Simple

Let Y be defined as a region bounded by $x_1 \leqslant x \leqslant x_2$ and $y_1(x) \leqslant y \leqslant y_2(x)$. Y is called y-simple as it can be split into subdomains R_i with continuous lines **parallel to the y-axis**.
The double integral over Y can then be decomposed into the sum of integrals over each subdomain Y_i:

$$\iint_Y f(x,\,y)\,\mathrm{d}A = \sum_{i=1}^{n} \iint_{Y_i} f(x,\,y)\,\mathrm{d}A$$

which is equivalent to integrating between the y-bounds of each subdomain Y_i within the x-bounds of Y:

$$\iint_Y f(x,\,y)\,\mathrm{d}A = \int_{x_1}^{x_2} \left[\int_{y_1(x)}^{y_2(x)} f(x,\,y)\,\mathrm{d}y\right]\mathrm{d}x$$

3.4.2 Type II — x-Simple

Let X be defined as a region bounded by $x_1(y) \leqslant x \leqslant x_2(y)$ and $y_1 \leqslant y \leqslant y_2$. X is called x-simple as it can be split into subdomains R_i with continuous lines **parallel to the x-axis**.
The double integral over X can then be decomposed into the sum of integrals over each subdomain X_i:

$$\iint_X f(x,\,y)\,\mathrm{d}A = \sum_{i=1}^{n} \iint_{X_i} f(x,\,y)\,\mathrm{d}A$$

which is equivalent to integrating between the x-bounds of each subdomain X_i within the y-bounds of X:

$$\iint_X f(x,\,y)\,\mathrm{d}A = \int_{y_1}^{y_2} \left[\int_{x_1(y)}^{x_2(y)} f(x,\,y)\,\mathrm{d}x\right]\mathrm{d}y$$

3.5 Transformation of Coordinates

In single variable calculus, we can use the transformation $x = g(u)$ to change the variable of integration from x to u, using the chain rule, $\mathrm{d}x = g'(u)\,\mathrm{d}u$. For multiple integrals, the same can be accomplished using the Jacobian matrix.

Definition 3.4 (Jacobian Matrix)**.** Consider the transformation $\mathbf{x} = T(\mathbf{u})$, where $T : \mathbb{R}^m \to \mathbb{R}^n$ is once differentiable. The Jacobian matrix of T is defined as an $m \times n$ matrix, denoted $\mathbf{J}$, whose $(i,\, j)$-th entry is given by

$$\mathbf{J}_{i,\,j} = \frac{\partial x_i}{\partial u_j}.$$

Explicitly, the Jacobian matrix is given by

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial x_1}{\partial u_1} & \cdots & \dfrac{\partial x_1}{\partial u_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x_m}{\partial u_1} & \cdots & \dfrac{\partial x_m}{\partial u_n} \end{bmatrix}.$$

It may also be notated as

$$\mathbf{J} = \frac{\partial\,(x_1,\, \ldots,\, x_m)}{\partial\,(u_1,\, \ldots,\, u_n)}$$

Definition 3.5 (Jacobian)**.** When $m = n$, we can determine the ratio of the area of the original region R to the area of the transformed region $T(R)$ using the determinant of the Jacobian matrix:

$$|\mathbf{J}| = \det(\mathbf{J}) = \left|\frac{\partial\,(x_1,\, \ldots,\, x_n)}{\partial\,(u_1,\, \ldots,\, u_n)}\right|.$$

In the case of two variables, $(x,\, y) = T(u,\, v)$, and the Jacobian matrix is defined as

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix}.$$

The Jacobian is then given by

$$|\mathbf{J}| = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

3.5.1 General Transformation

Using this matrix determinant allows us to change the variables of integration from $(x,\, y)$ to $(u,\, v)$ using the following formula:

$$\iint_R f(x,\, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{T(R)} f(T(u,\, v))\,|\det(\mathbf{J})|\,\mathrm{d}u\,\mathrm{d}v$$

where $f(T(u,\, v))$ is multiplied by the absolute value of the Jacobian.

3.5.2 Polar Coordinate Transformation

Definition 3.6 (Polar Coordinates)**.** A polar coordinate system is defined by the transformation

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned}$$

where $r > 0$ is the radius of the circle and $\theta \in [0, 2\pi)$ is the angle of rotation measured anticlockwise from the positive x-axis. The Jacobian is given by

$$\begin{aligned} \mathbf{J} &= \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} \\ \det(\mathbf{J}) &= r\cos^2\theta + r\sin^2\theta \\ &= r \end{aligned}$$

The infinitesimal area element $\mathrm{d}A$ is therefore

$$\mathrm{d}A = |\det(\mathbf{J})|\,\mathrm{d}r\,\mathrm{d}\theta = r\,\mathrm{d}r\,\mathrm{d}\theta.$$

Using polar coordinates, the double integral over a region R is defined

$$\iint_R f(x,\, y)\,\mathrm{d}A = \int_0^{\infty}\int_0^{2\pi} f(r\cos\theta,\, r\sin\theta)\, r\,\mathrm{d}\theta\,\mathrm{d}r$$

3.6 Strategies for Evaluating Double Integrals

1. Decompose the region of integration into simple domains, and if this is not possible, divide the region into subregions.
2. Change the variables of integration to simplify the integral.
3. Swap the order of integration to simplify the integral.

4 Multiple Integrals

All results and properties for double integrals can be extended to triple integrals.

Theorem 4.0.1 (Fubini's Theorem in n-dimensions)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^+$, or $f : \mathbb{R}^n \to \mathbb{R}$ is integrable, then all $n!$ permutations of integrals are equal:*

$$\int_{\mathbb{R}^n} f(x_1,\, \ldots,\, x_n)\,\mathrm{d}x_1 \cdots \mathrm{d}x_n = \int_{-\infty}^{\infty}\left[\cdots\left[\int_{-\infty}^{\infty} f(x_1,\, \ldots,\, x_n)\,\mathrm{d}x_1\right]\cdots\right]\mathrm{d}x_n$$

As with double integrals, we can evaluate integrals of nonpositive functions by using the positive and negative parts of the function.

4.1 Vector-Valued Functions

Definition 4.1 (Vector-Valued Function)**.** A vector-valued function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is a function that maps a vector $\mathbf{x} \in \mathbb{R}^n$ to a vector $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$.

A vector-valued function is integrable if either:

- the norm of the function is integrable

$$\int_{\mathbb{R}^n} \|\mathbf{f}(\mathbf{x})\| \, \mathrm{d}\mathbf{x} < \infty$$

- or if all components of the function are integrable

$$\int_{\mathbb{R}^n} |f_i(\mathbf{x})| \, \mathrm{d}\mathbf{x} < \infty \qquad \forall i \in \{1, \ldots, m\}$$

The integral of a vector-valued function is defined as

$$\int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \left\langle \int_{\mathbb{R}^n} f_1(\mathbf{x}) \, \mathrm{d}\mathbf{x}, \ldots, \int_{\mathbb{R}^n} f_m(\mathbf{x}) \, \mathrm{d}\mathbf{x} \right\rangle$$

4.2 Change of Variables

4.2.1 Cylindrical Coordinate Transformation

Definition 4.2 (Cylindrical Coordinates)**.** A cylindrical coordinate system is defined by the transformation

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}$$

where $r > 0$ is the radius of the cylinder, $\theta \in [0, 2\pi)$ is the angle of rotation measured anticlockwise from the positive x-axis, and $z \in \mathbb{R}$ is the height of the cylinder. The Jacobian is given by

$$\begin{aligned} \mathbf{J} &= \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} & \dfrac{\partial x}{\partial z} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} & \dfrac{\partial y}{\partial z} \\ \dfrac{\partial z}{\partial r} & \dfrac{\partial z}{\partial \theta} & \dfrac{\partial z}{\partial z} \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \det(\mathbf{J}) &= r\cos^2\theta + r\sin^2\theta \\ &= r \end{aligned}$$

The infinitesimal area element $\mathrm{d}V$ is therefore

$$\mathrm{d}V = |\det(\mathbf{J})| \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}z = r \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}z.$$

Using cylindrical coordinates, the triple integral over a region R is defined

$$\iiint_R f(x,\ y,\ z)\,\mathrm{d}V = \int_{-\infty}^{\infty}\int_0^{\infty}\int_0^{2\pi} f(r\cos\theta,\ r\sin\theta,\ z)\,r\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}z$$

4.2.2 Spherical Coordinate Transformation

Definition 4.3 (Spherical Coordinates)**.** A spherical coordinate system is defined by the transformation

$$\begin{aligned} x &= \rho\sin\phi\cos\theta \\ y &= \rho\sin\phi\sin\theta \\ z &= \rho\cos\phi \end{aligned}$$

where $\rho > 0$ is the radius of the sphere, $\phi \in [0,\pi)$ is the polar angle measured down from the positive z-axis, and $\theta \in [0,2\pi)$ is the azimuthal angle measured anticlockwise from the positive x-axis. The Jacobian is given by

$$\begin{aligned} \mathbf{J} &= \begin{bmatrix} \dfrac{\partial x}{\partial\rho} & \dfrac{\partial x}{\partial\phi} & \dfrac{\partial x}{\partial\theta} \\ \dfrac{\partial y}{\partial\rho} & \dfrac{\partial y}{\partial\phi} & \dfrac{\partial y}{\partial\theta} \\ \dfrac{\partial z}{\partial\rho} & \dfrac{\partial z}{\partial\phi} & \dfrac{\partial z}{\partial\theta} \end{bmatrix} \\ &= \begin{bmatrix} \sin\phi\cos\theta & \rho\cos\phi\cos\theta & -\rho\sin\phi\sin\theta \\ \sin\phi\sin\theta & \rho\cos\phi\sin\theta & \rho\sin\phi\cos\theta \\ \cos\phi & -\rho\sin\phi & 0 \end{bmatrix} \\ \det(\mathbf{J}) &= \rho^2\sin^3\phi\cos^2\theta + \rho^2\sin\phi\cos^2\phi\cos^2\theta - \rho\sin\phi\sin\theta\left(-\rho\sin^2\phi\sin\theta - \rho\cos^2\phi\sin\theta\right) \\ &= \rho^2\sin\phi\cos^2\theta\left(\sin^2\phi+\cos^2\phi\right) + \rho^2\sin\phi\sin^2\theta\left(\sin^2\phi+\cos^2\phi\right) \\ &= \rho^2\sin\phi\left(\sin^2\theta+\cos^2\theta\right) \\ &= \rho^2\sin\phi \end{aligned}$$

The infinitesimal area element $\mathrm{d}V$ is therefore

$$\mathrm{d}V = |\det(\mathbf{J})|\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta = \rho^2\sin\phi\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta.$$

Using spherical coordinates, the triple integral over a region R is defined

$$\iiint_R f(x,\ y,\ z)\,\mathrm{d}V = \int_0^{\infty}\int_0^{2\pi}\int_0^{\pi} f(\rho\sin\phi\cos\theta,\ \rho\sin\phi\sin\theta,\ \rho\cos\phi)\,\rho^2\sin\phi\,\mathrm{d}\phi\,\mathrm{d}\theta\,\mathrm{d}\rho$$

4.3 Interpretations of Integrals

Measure

The measure of a region $R \in \mathbb{R}^n$ is given by:

$$\mu(R) = \int_R \mathrm{d}\mathbf{x}$$

In two dimensions, the measure is the area of the region:

$$\mu(R) = \iint_R \mathrm{d}A$$

In three dimensions, the measure is the volume of the region:

$$\mu(R) = \iiint_R \mathrm{d}V$$

Mass

The mass of a region $R \in \mathbb{R}^n$ with density function ρ is given by:

$$M = \int_R \rho(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

Centroid

The average position of a region $R \in \mathbb{R}^n$ with uniform density is given by:

$$\langle \mathbf{r} \rangle = \frac{1}{\mu(R)} \int_R \mathbf{x}\,\mathrm{d}\mathbf{x}$$

This point is the geometric centre of the region where the region would balance if it were made of a uniform material.

Centre of Mass

The average position of a region $R \in \mathbb{R}^n$ with density function ρ is given by:

$$\langle \mathbf{r} \rangle_\rho = \frac{1}{M} \int_R \rho(\mathbf{x})\,\mathbf{x}\,\mathrm{d}\mathbf{x}$$

This point is the centre of mass of the region where the region would balance if it were made of a material with density ρ.

Average Value

The average value of a function $f : \mathbb{R}^n \to \mathbb{R}$ over a region $R \in \mathbb{R}^n$ is given by:

$$\langle f(\mathbf{r}) \rangle = \frac{1}{\mu(R)} \int_R f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

Expected Value

The average value of a function $f : \mathbb{R}^n \to \mathbb{R}$ over a region $R \in \mathbb{R}^n$ with density function p is given by:

$$\langle f(\mathbf{r}) \rangle_p = \int_R p(\mathbf{x})\, f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

where p is a probability density function satisfying

$$\int_R p(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$$

This result is the expected value of f over R if the region is randomly sampled according to the probability density function p.

4.4 Properties of Integrals

Linearity of Integrals

If f and g are integrable functions, and $c \in \mathbb{R}$, then

$$\int_{\mathbb{R}^n} (f(\mathbf{x}) + g(\mathbf{x}))\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_{\mathbb{R}^n} g(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

and

$$\int_{\mathbb{R}^n} cf(\mathbf{x})\,\mathrm{d}\mathbf{x} = c\int_{\mathbb{R}^n} f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

Positivity of Integrals

If $f \geqslant 0$ for all $\mathbf{x} \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \geqslant 0$$

Monotonicity of Integrals

If f and g are integrable functions, and $f(\mathbf{x}) \leqslant g(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \leqslant \int_{\mathbb{R}^n} g(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

Triangle Inequality

If f is an integrable function, then

$$\left\|\int_{\mathbb{R}^n} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\right\| \leqslant \int_{\mathbb{R}^n} \|f(\mathbf{x})\|\,\mathrm{d}\mathbf{x}$$

Change of Variables

When f is positive or integrable, the bijective transformation $\mathbf{T} : R \subset \mathbb{R}^n \to R' \subset \mathbb{R}^n$ with continuous first derivative $\mathbf{T}'$ allows us to change the variables of integration from $\mathbf{x}$ to $\mathbf{u}$:

$$\int_R f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{R'} f(\mathbf{T}(\mathbf{u}))\,|\det(\mathbf{J})|\,\mathrm{d}\mathbf{u}$$

where $\mathbf{J}$ is the Jacobian matrix of $\mathbf{T}$.

Measure Zero Integrals

If $\mu(R) = \int_R \mathrm{d}\mathbf{x} = 0$, then

$$\int_R f(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0$$

Almost Equal Functions

If f and g are integrable functions, and $f(\mathbf{x}) = g(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ except for a set of measure zero, then

$$\int_{\mathbb{R}^n} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} g(\mathbf{x})\,\mathrm{d}\mathbf{x}$$