1 Basics of Ordinary Differential Equations (Grundlagen: Gewöhnliche Differentialgleichungen)

1.1 Solving a Differential Equation

Case 1: Homogeneous DE

we have: $i(t) = 0$ and $k'(t) = \delta k(t)$

Claim: solution is $k(t) = \gamma \cdot e^{\delta t}$
Set of solutions for the homogeneous DE: $\{\gamma \cdot e^{\delta t} : \gamma \in R\}$

Case 2: Inhomogeneous DE we have a general $i(t)$

1. It is sufficient to determine (particular) solution of the Inhomogeneous DE
   $K_n(t) \rightarrow$ solution for the homogeneous DE $\Rightarrow K_n'(t) = \delta \cdot K_n(t)$
   $K_p(t) \rightarrow$ solution for the Inhomogeneous DE $\Rightarrow K_p'(t) = \delta \cdot K_p(t) + i(t)$
   combining both Equations, we have:

$$\begin{aligned} K'(t) &= K_p'(t) + K_n'(t) \\ &= \delta \cdot K_p(t) + i(t) + \delta \cdot K_n(t) \\ &= \delta \cdot (K_p(t) + K_n(t)) + i(t) \\ &= \delta \cdot K(t) + i(t) \end{aligned}$$

2. Determine one solution for Inhomogeneous DE through variation of constant
   $k(t) = \gamma \cdot e^{\delta t}$
   $k(t) = c(t) \cdot e^{\delta t}$
   substituting this into the DE

$$k'(t) = c'(t) \cdot e^{\lambda t} + c(t) \cdot \delta e^{\delta t} \tag{1}$$

$$k'(t) = k(t) \cdot \lambda + i(t) \tag{2}$$

   from (1) and (2) we have:

$$c'(t) \cdot e^{-\lambda \cdot t} = i(t)$$

   Integrating both sides

$$c(t) = \int i(t) \cdot e^{-\lambda t} dt$$

Example: Solving an Ordinary Differential Equation

$$f'(x) = 1 - x + f(x),\ f(0) = 2$$

1. Homogeneous DE:

$$f'(x) = f(x) \Rightarrow f(x) = \gamma \cdot e^x$$

   check:

$$f'(x) = f(x) = \gamma \cdot e^x - \gamma \cdot e^x = 0$$

2. Variation of constant:

$$f(x) = c(x) \cdot e^x$$
$$f'(x) = c'(x) \cdot e^x + c(x) \cdot e^x$$
$$f'(x) - f(x) = 1 - x$$

Replacing $f'(x)$ and $f(x)$ with the function of $c(x)$ and $c'(x)$

$$c'(x) \cdot e^x + c(x) \cdot e^x - c(x) \cdot e^x = 1 - x$$
$$c'(x) = e^{-x} \cdot (1 - x)$$
$$c(x) = \int e^{-x} \cdot (1 - x) dx$$
$$c(x) = \int e^{-x} dx - \int e^{-x} \cdot x dx$$
$$c(x) = e^{-x} \cdot x$$

Particular solution:

$$f(x) = c(x) \cdot e^x \Rightarrow f(x) = e^{-x} \cdot x \cdot e^x$$
$$f(x) = x \Rightarrow f'(x) = 1$$

3. Set of solutions:

$$f(x) = x + \gamma \cdot e^x$$

4. Initial Value:

$$f(0) = 2$$
$$f(0) = 0 + \gamma = 2 \Rightarrow \gamma = 2$$
$$\rightarrow f(x) = x + 2 \cdot e^x$$

solves the initial value problem

Note: You can also solve the Differential Equation with the help of an integrating factor
Refer to this wikipedia article integrating factor and this text

***

1.2 System of Differential Equations

A homogeneous system of Linear Differential Equations with constant coefficients is a system of the form:

$$\begin{pmatrix} y_1'(t) \\ \vdots \\ \vdots \\ y_n'(t) \end{pmatrix} = A \cdot \begin{pmatrix} y_1(t) \\ \vdots \\ \vdots \\ y_n(t) \end{pmatrix}, \quad \begin{pmatrix} y_1(0) \\ \vdots \\ \vdots \\ y_n(0) \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \vdots \\ \gamma_2 \end{pmatrix}$$

$$\text{where A} \in R \text{ and } \begin{pmatrix} \gamma_1 \\ \vdots \\ \vdots \\ \gamma_2 \end{pmatrix} \in R^n$$

The key idea is that if A is a diagonal matrix, then each row is a Differential Equation that is independent of the others, solve each row separately **decoupled system**

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$$

We need a basis and a corresponding basis transformation S $\in R^{n\times n}$ such that $S^{-1}AS = D$ is diagonal
Then we substitute, y = SZ $\Rightarrow Z = S^{-1}y$
The DE system,

$$y' = A \cdot y$$

becomes

$$S \cdot z = A \cdot S \cdot Z$$

$$Z' = S^{-1} \cdot A \cdot S \cdot Z = D \cdot Z$$

Solve the new decoupled system and transform solution back through $y = SZ$
Example:

$$\gamma'(t) = \frac{5}{2} \cdot \gamma(t) - \phi(t)$$

$$\phi'(t) = \frac{-1}{4} \cdot \gamma(t) + \frac{5}{2}\phi(t)$$

matrix form

$$\begin{pmatrix} \gamma' \\ \phi' \end{pmatrix} = \begin{pmatrix} \frac{5}{2} & -1 \\ \frac{-1}{4} & \frac{5}{2} \end{pmatrix} \begin{pmatrix} \gamma \\ \phi \end{pmatrix}$$

now, determine the eigenvalues and eigenvectors for the matrix

$$p_A(\lambda) = \det(A - \lambda I_2) = \det \begin{pmatrix} \frac{5}{2} - \lambda & -1 \\ \frac{-1}{4} & \frac{5}{2} - \lambda \end{pmatrix} = \left(\frac{5}{2} - \lambda\right)^2 - \frac{1}{4} = \frac{25}{4} + \lambda^2 - 5\lambda - \frac{1}{4} = 6 + \lambda^2 - 5\lambda = (\lambda - 3)(\lambda - 2)$$

The roots of the equation are $\lambda_1 = 3$ and $\lambda_2 = 2$
For the respective eigenspaces, we get

$$eig_A(3) = \ker(A - 3I_2) = \ker \begin{pmatrix} \frac{-1}{2} & -1 \\ \frac{-1}{4} & \frac{-1}{2} \end{pmatrix} = \ker \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & 0 \end{pmatrix} = span\left\{\begin{pmatrix} -2 \\ 1 \end{pmatrix}\right\}$$

$$eig_A(2) = \ker(A - 2I_2) = \ker \begin{pmatrix} \frac{1}{2} & -1 \\ \frac{-1}{4} & \frac{1}{2} \end{pmatrix} = \ker \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & 0 \end{pmatrix} = span\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$$

$$S = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}$$

$$S^{-1} = \frac{-1}{4} \begin{pmatrix} 1 & -2 \\ -1 & -2 \end{pmatrix}$$

$$S^{-1}AS = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} = D$$

Solve the system:

$$Z' = DZ = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

$$Z_1(t) = \gamma_1 e^{3t}$$

$$Z_2(t) = \gamma_2 e^{2t}$$

$$\begin{pmatrix} \gamma \\ \phi \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \gamma_1 e^{3t} \\ \gamma_2 e^{2t} \end{pmatrix}$$

$$\gamma(t) = -2\gamma_1 e^{3t} + 2\gamma_2 e^{2t}$$
$$\phi(t) = \gamma_1 e^{3t} \gamma_2 e^{3t}$$

we know the initial values: $\gamma(0) = 60$ and $\phi(0) = 60$

$$-2\gamma_1 + 2\gamma_2 = 60$$
$$\gamma_1 + \gamma_2 = 60$$

solving, we get $\gamma_1 = 15$ and $\gamma_2 = 45$

$$\gamma(t) = -30e^{3t} + 90e^{2t}$$
$$\phi(t) = 15e^{3t} + 45e^{2t}$$

Basics of Complex Numbers

A complex number is represented as $Z = a + ib$ where a and b are real numbers and

$$Re(Z) = a$$
$$Im(Z) = b$$
$$i = \sqrt{-1}$$
$$i^2 = -1$$

A complex number can be represented on an argand plane with the X-axis as the real part of the complex number and the Y-part as the imaginary part of the complex number

Algebra of Complex numbers

Two complex numbers $Z_1 = a_1 + ib_1$ and $Z_2 = a_2 + ib_2$ can be added and subtracted by seperately adding/subtracting their real and imaginary parts.
Multiplication of two complex numbers,

$$Z_1 \cdot Z_2 = a_1a_2 + ia_1b_2 + ib_1a_2 + i^2b_1b_2 = a_1a_2 + i(a_1b_2 + b_1a_2) - b_1b_2$$

Division of two complex numbers,
Mulitiplying both numerator and denominator by the **conjugate** of the denominator inorder to obtain a real denominator and a complex numerator. For example,

$$\frac{4+3i}{1+2i} = \frac{(4+3i)(1-2i)}{(1+2i)(1-2i)} = \frac{4-8i+3i-6i^2}{1-4i^2} = \frac{12+5i}{5} = \frac{12}{5} + i$$

Complex Conjugate and absolute value

For a complex number $Z = a + ib$, the complex conjugate is defined as $\overline{Z} = a - ib$

$$Z\overline{Z} = (a+ib)(a-ib) = a^2 - iab + iab + b^2 = a^2 + b^2 \tag{3}$$

As per the definition of the absolute value of a complex number

$$|Z| = \sqrt{a^2 + b^2} \tag{4}$$

By (1) and (2), we have

$$Z\overline{Z} = |Z|^2$$

Polar form and Euler's form

Representing the complex number in the form of a $|Z|$ and argument ϕ

$$Z = |Z|(\cos\phi + i\sin\phi)$$

is the Polar representation of a complex number

$$e^{i\phi} = \cos\phi + i\sin\phi$$

Therefore,

$$Z = |Z|e^{i\phi}$$

is the Euler form of the complex number A lot of cool stuff related to complex number on this playlist. Not required for economic aspects though.

2 Eigenvectors and Eigenalues (Eigenwertprobleme)

$S \in R^{n\times n}$
$A \in R^{n\times n}$
$S^{-1}AS = D$ diagonal matrix, tranformation matrix
Is there a basis transformation?
$\lambda \in R$ is called an eigenvalue of A if there exists a vector $v \neq \phi$ such that

$$A \cdot v = \lambda \cdot v$$

Any such vector v in this case would be called an eigen vector of A for the eigen value λ **Gaussian Reformulation:** Determination of the eigenvalues and eigenvectors.
λ is an eigenvalue of A if there $v \neq \phi$ with $A \cdot v = \lambda \cdot v$

$$\Leftrightarrow A \cdot v - \lambda \cdot v = 0$$

$$\Leftrightarrow (A - \lambda E_n) \cdot v = 0$$

$$\Leftrightarrow v \in \ker(A \cdot \lambda E_n)$$

$$\Leftrightarrow (A - \lambda E_n) \text{ is not invertible}$$

$$\Leftrightarrow \det(A - \lambda E_n) = 0$$

The **characteristic polynomial** $p_A(\lambda)$ is defined as the determinant $\det(A - \lambda E_n)$

$$p_A(\lambda) = \det(A - \lambda E_n)$$

and the kernel $\ker(A - \lambda E_n)$ is called the **eigenspace of A for the eigenvalue** λ and written as

$$eig_A(\lambda) = \ker(A - \lambda E_n)$$

The idea to solve each and every problem here is quite simple,

1. First of all, determine the **characteristic polynomial** [Diagonalized matrix S is the matrix with the eigenspaces of the respective eigenvalues]
2. Find the roots of the characteristic polynomial, these are the eigenvalues of A
3. For each eigenvalue of A, determine the eigenspace by solving the linear equation system
   $(A - \lambda E_n) \cdot v = 0$

2.1 Diagonalization using complex numbers

characterizing which matrices can be transformed into diagonal form, ie possess an eigenvector basis.
A matrix $A \in R^{n\times n}$ is called **diagonalizable** if there is a basis of R^n that consists of eigenvectors of A
NOTE: For each eigenvalue $\rightarrow$ there is at least one eigenvector ie. $\dim(eig_A(\lambda)) \geq 1$ for each eigenvalue $\lambda \Rightarrow$ if A possesses n distinct eigenvalue, then there are at least n eigenvectors

$$p_A(\lambda) = (\lambda - \lambda*)^k \cdot g(\lambda) \text{ and } g(\lambda*) \neq 0$$

Then k is called the **algebraic multiplicity** of $\lambda*$ and $\dim(eig_A(\lambda*))$ is called the geometric mutliplicity of $\lambda*$

1. Algebraic mulitplicity $\geq$ Geometric mulitplicity
2. For pairwise distinct eigenvalues of A with corresponding eigenvectors- the set $\{v_1, \ldots, v_r\}$ is linearly independent
3. For a matrix to be diagonalizable, the algebraic and the geometric mulitplicity have to be equal for every eigenvalue of A

Proving (2) is pretty simple as you can simply use the principle of induction after solving it for 2 eigenvalues and their corresponding eigenvectors
If roots are complex, we can use the Euler's formula

$$e^{i\phi} = \cos\phi + i\sin\phi$$

what if the matrix cannot be diagonalized? **JORDAN NORMAL FORM** as generalizations and refer for Inhomogeneous Differential Equation system
`https://people.math.harvard.edu/~knill/teaching/math19b_2011/handouts/lecture29.pdf` for a recap

2.2 Predator-Prey Model

squirrels-hawks $\rightarrow$ The growth of hawk population depends on availability of squirrels and vice-versa. [Fewer hawks $\rightarrow$ higher the growth rate of squirrels]
Model this through DE system:

$$h'(t) = s(t) - 12 \text{ and } h(0) = 6$$
$$s'(t) = -h(t) \text{ and } s(0) = 20$$

1. get rid of constants: define a new system

$$y_1(t) = h(t) - 10 = -s'(t) \text{ and } y_1(0) = h(0) - 10 = -4$$
$$y_2(t) = s(t) - 12 = h'(t) \text{ and } y_2(0) = s(0) - 12 = 8$$
$$\Rightarrow y_1'(t) = h'(t) = s(t) - 12 = y_2(t)$$
$$\Rightarrow y_2'(t) = -h(t) + 10$$

   new system without constants:

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} \mid \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = \begin{pmatrix} -4 \\ 8 \end{pmatrix}$$

2. Diagonalize the matrix A [find eigenvalues and the corresponding eigenvectors then form the S matrix]

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

   not showing the calculations but the eigenvectors are $+i$, $-i$ and eigenvectors are $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ so the required S matrix is $\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$

$$S^{-1}AS = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

3. consider the decoupled system after substitution $y = SZ$

$$Z'(t) = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} Z(t)$$

$\Rightarrow$

$$Z_1'(t) = i \cdot Z_1(t) \Rightarrow Z_1(t) = \gamma_1 \cdot e^{it}$$
$$Z_2'(t) = -i \cdot Z_2(t) \Rightarrow Z_2(t) = \gamma_2 \cdot e^{-it}$$

4. determine y = SZ:

$$y(t) = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \gamma_1 e^{it} \\ \gamma_2 e^{-it} \end{pmatrix}$$

$\Rightarrow$

$$y_1(t) = \gamma_1 e^{it} + \gamma_2 e^{-it}$$
$$y_2(t) = \gamma_2 i e^{it} - \gamma_2 i e^{-it}$$

5. Determine γ_1, γ_2 through initial values:

$$y_1(0) = -4 = \gamma_1 \cdot 1 + \gamma_2 \cdot 1$$
$$y_2(0) = 8 = \gamma_1 \cdot i \cdot 1 - \gamma_2 \cdot i \cdot 1$$
$$\rightarrow \left(\begin{array}{cc|c} 1 & 1 & -4 \\ i & -i & 8 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & 1 & -4 \\ 0 & -2i & 8+4i \end{array}\right)$$
$$(-2i)\gamma_2 = 8 + 4i$$
$$\Leftrightarrow \gamma_2 = 4i - 2 = -2 + 4i$$
$$\Rightarrow \gamma_1 = -4 - \gamma_2 = -2 - 4i$$

Solutions:

$$y_1(t) = (-2-4i)e^{it} + (-2+4i)e^{-it}$$
$$y_2(t) = (4-2i)e^{it} + (4+2i)e^{-it}$$

6. Use Euler's Formula: replace e^{it} and e^{-it} with $cis\theta$ terms

$$y_1(t) = -4\cos t + 8\sin t \mid h(t) = -4\cos t + 8\sin t + 10$$
$$y_2(t) = 8\cos t + 4\sin t \mid s(t) = 8\cos t + 4\sin t + 12$$

3 Multivariable Calculus

3.1 Fundamentals

Definition: Let $f : X \to Y$ be a function with $X \subseteq \mathbb{R}^n$ (domain) and $Y \subseteq \mathbb{R}^n$ (codomain) for $m, n \in \mathbb{N}$
The graph of f is the set

$$graph(f) = \left\{ \begin{pmatrix} x \\ f(x) \end{pmatrix} : x \in X \right\}$$

A function of two variables assigns a real number to each point in a subset of $\mathbb{R}^2$. For example,

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

defines a function over $\mathbb{R}^2$. The **domain** of a multivariable function is the set of all input values for which the function is defined. The **range** is the set of all output values. Example: For $f(x, y) = \sqrt{1 - x^2 - y^2}$, the domain is the unit disk:

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

Definition: Let $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$ be a function and let $\gamma \in \mathbb{R}$
The set

$$L_f(\gamma) := \{x \in X : f(x) = \gamma\}$$

is called the level set of f for level γ

3.2 Basic Topology

3.2.1 Open Balls and δ-Neighborhoods

Let $x^* \in \mathbb{R}^n$, $\delta > 0$. Then the δ-neighborhood of x^* is the set

$$N_\delta(x^*) := \{x \in \mathbb{R}^n : ||x - x^*|| < \delta\}$$

for some $\delta > 0$. N_δ is called **open ball of radius δ around x^* and sometimes denoted by $B_\delta(x^*)$)**

3.2.2 Interior Points

A point $x \in A \subseteq \mathbb{R}^n$ is called an **interior point** of A if there exists a $\delta > 0$ such that:

$$N_\delta(x*) \subseteq X$$

The set of all interior points of X is called the **interior** of X, denoted $\text{int}(X)$.

3.2.3 Boundary Points

A point $x \in \mathbb{R}^n$ is a **boundary point** of a set A if every δ-neighborhood of x intersects both X and its complement (not in X):

$$\forall \delta > 0,\ N_\delta(x^*) \cap X \neq \emptyset \quad \text{and} \quad N_\delta(x^*) \cap (\mathbb{R}^n \setminus X) \neq \emptyset$$

The set of all boundary points is called the **boundary** of X, denoted $bd(X)$ or ∂X.
The set $cl(X) := X \cup bd(X)$ is called **closure** of X (sometimes written $\bar{X}$)

3.2.4 Closure of a Set

The **closure** of a set X, denoted $\overline{X}$, is the union of X and its accumulation points. Equivalently,

$$\overline{X} = X \cup \partial X$$

3.2.5 Open and Closed Sets

1. A set A is **open** if all its points are interior points, i.e., $A = \text{int}(A)$.
2. A set is **closed** if it contains all its boundary points.

3.2.6 Accumulation (Limit) Points

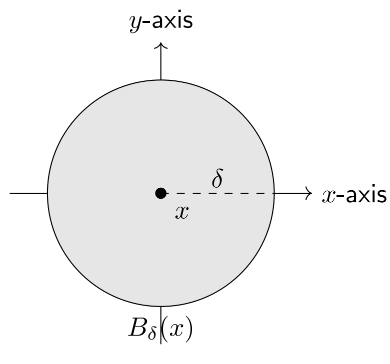
A point x is an **accumulation point** (or limit point) of a set X if every neighborhood of x contains at least one point of X different from x:

$$\forall \delta > 0, \quad (B_\delta(x) \setminus \{x\}) \cap X \neq \emptyset$$

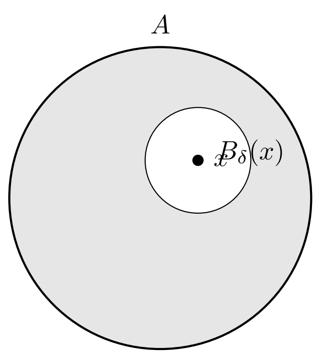
The set X is called

1. **open** if $X = int(X)$
2. **closed** if $\mathbb{R} \setminus X$ is open
3. **bounded** if there is some $\delta > 0$ such that $X \subseteq N_\delta(0)$
4. **compact** if it is closed and bounded.

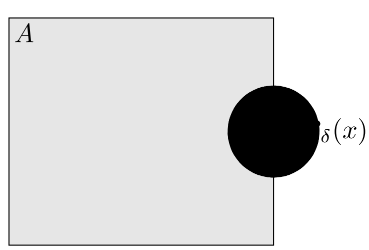
1. δ-Neighborhood (Open Ball)



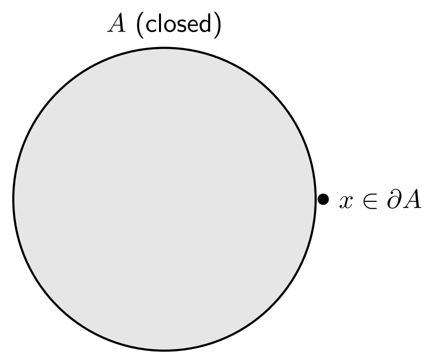
2. Interior Point



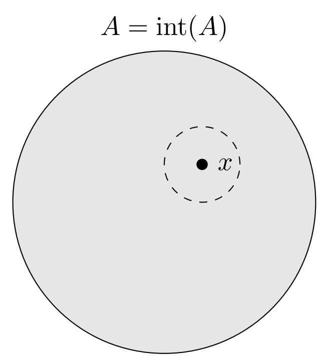
3. Boundary Point



4. Closed Set



5. Open Set



3.3 Sequences and Limits

A function in $f : \mathbb{N}_0 \to \mathbb{R}^n$ is called a **sequence** in $\mathbb{R}^n$ with $a_k := f(k)$ such a sequence is usually denoted as $(a_k)_{k\in\mathbb{N}_0}$ or just $(a_k)_{\mathbb{N}_0}$
A sequence $(a_k)_{\mathbb{N}}$ is called **convergent** if there exists a vector $a \in \mathbb{R}^n$ such that

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } ||a_k - a|| \to 0$$

as $k \to \infty$. The vector a is called the **limit of** $(a_k)_{\mathbb{N}}$ and is denoted by

$$a = \lim a_k$$

$(a_k)_{\mathbb{N}}$ is convergent with limit a if and only for $i = 1, \ldots, n$ the following holds:

1. the sequence $(\alpha_i k)_{k\in\mathbb{N}}$ is convergent
2. $\lim_k \alpha_{ik} = \alpha_i$

Definitions

1. A sequence $\{a_k\} \subset \mathbb{R}^n$ is **bounded** if the set $\{a_k : k \in \mathbb{N}_0\}$ is bounded
2. Let $(a_k)_{\mathbb{N}_0}$ be a sequence and $f : \mathbb{N}_0 \to \mathbb{N}_0$ be strictly increasing, i.e. for $p > q \implies f(p) > f(q)$.
   Then the sequence $(a'_k)_{\mathbb{N}_0}$ defined as $a'_k = a_{f(k)}$ is called a **subsequence** of $(a_k)_{\mathbb{N}_0}$
3. The limit of a convergent subsequence of some sequence $(a_k)_{\mathbb{N}_0}$ is called the **accumulation point** of $(a_k)_{\mathbb{N}_0}$

3.3.1 Bolzano-Weierstrass Theorem

Theorem: Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence. (i.e. at least one accumulation point)
in particular, if X is a compact set and $(a_k)_{\mathbb{N}_0}$ is a subsequence in X, then $(a_k)_{\mathbb{N}}$ has a convergent subsequence with limit in X.
Idea of Proof: Use the fact that $\mathbb{R}^n$ is a product of complete metric spaces, and apply the one-dimensional Bolzano-Weierstrass theorem component-wise to extract a convergent subsequence.

3.4 Continuity

Definition:
Let $X \subseteq \mathbb{R}^m$, $f : X \to \mathbb{R}^m$ and $x^* \in X$.

(3) $\alpha \in \mathbb{R}^m$ is called the limit of f for $x \to x^*$ if the following holds:
For every sequence $(a_k)_{k\in\mathbb{N}}$ in X that converges to x^*, the sequence $(f(a_k))_{k\in\mathbb{N}}$ converges to α.

$$\lim_{x\to x^*} f(x) = \alpha$$

(2) f is called continuous in x^* if

$$\lim_{x\to x^*} f(x) = f(x^*).$$

(3) f is called continuous on X if f is continuous at every point $x^* \in X$.

3.4.1 Theorem:

Let $X \subseteq \mathbb{R}^m$, $f : X \to \mathbb{R}^m$ and $g : X \to \mathbb{R}$ be continuous on X, and let $\alpha \in \mathbb{R}$. Then the following are true:

(1) $f + g$, αf, and $f \cdot g$ are continuous on X.

(2) $\frac{f}{g}$ is continuous on $X \setminus \{x \in X : g(x) = 0\}$.
$\frac{f}{\alpha}$ is continuous for $\alpha \neq 0$.

(3) Let $h : \mathbb{R}^k \to X$ be continuous, then $f \circ h$ is continuous.

(4) If

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix},$$

then f is continuous on X if and only if all $f_i : X \to \mathbb{R}$ are continuous on X.

Example:

$$f(x_1, x_2) := \begin{pmatrix} x_1 \cos(x_2) \\ e^{x_1+x_2} - 1 \end{pmatrix}, \quad f : \mathbb{R}^2 \to \mathbb{R}^2.$$

Then f is continuous:

(1) f is continuous if and only if both coordinate functions are continuous:

$$f_1(x_1, x_2) = x_1 \cos(x_2), \quad f_2(x_1, x_2) = e^{x_1+x_2} - 1.$$

(2) f_1 is a product of continuous functions, hence continuous.

(3) f_2 is the sum of continuous functions. The term $e^{x_1+x_2}$ is continuous as it is a composition of continuous functions. $\Rightarrow f_2$ is continuous.

Definition:

A multivariate polynomial is a function $p : \mathbb{R}^n \to \mathbb{R}$ of the form:

$$p(x) = \sum_I a_I x^I,$$

where I is a multi-index and $a_I \in \mathbb{R}$.

Example:

$$p(x_1, x_2, x_3) = 2x_1^2 + 4x_2^3 + 8x_3^4.$$

Multivariate polynomials are continuous on $\mathbb{R}^n$.

(1) Sum of simple functions (monomials).

(2) Each monomial is a product of constants and univariate monomials.

(3) Products and sums of continuous functions are continuous.

3.4.2 Theorem:

Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}^m$. Then f is continuous if and only if:

(i) For any open set $Y \subseteq \mathbb{R}^m$, the preimage $f^{-1}(Y)$ is open in X.

(ii) For every closed set $Y \subseteq \mathbb{R}^m$, the preimage $f^{-1}(Y)$ is closed in X.

3.4.3 Theorem:

Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}^m$ continuous. If X is compact, then $f(X)$ is also compact (closed and bounded).

This implies: For compact $X \subseteq \mathbb{R}^n$ with $X \neq \emptyset$ and continuous $f : X \to \mathbb{R}$, the function f attains its minimum and maximum on X.

An Example and a Warning

Let $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$, be a multivariate function.

$$f(x_1, x_2) = x_1^2 + x_2^2, \quad x_1 \geq x_2.$$

We can define "directional sections" to obtain one-dimensional traces from f:

- Choose a direction $d \in \mathbb{R}^n \setminus \{0\}$.
- Consider the function f_d obtained by restricting f to the line in the direction of d:

  Line: $\{0 + \lambda d : \lambda \in \mathbb{R}\}$
  $f_d : \mathbb{R} \to \mathbb{R}$
  $f_d(\lambda) = f(0 + \lambda d)$.

If f is continuous, then all directional sections are also continuous. However, the converse is not generally true. Even if all f_d for all possible directions are continuous, the function f can still be discontinuous.

Example

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x_1, x_2) := \begin{cases} \frac{x_1^2 \cdot x_2}{x_1^4 + x_2^2}, & x \neq (0,0) \\ 0, & x = (0,0) \end{cases}$$

Is f continuous at $(0,0)$?
Let $d \in \mathbb{R}^2 \setminus \{0\}$ be an arbitrary direction. We consider $f_d : \mathbb{R} \to \mathbb{R}$,

$$f_d(\lambda) = f(0 + \lambda d) = f(\lambda d_1, \lambda d_2).$$

$$f_d(\lambda) = \begin{cases} \frac{\lambda^3 d_1^2 d_2}{\lambda^4 d_1^4 + \lambda^2 d_2^2} = \frac{\lambda d_1^2 d_2}{\lambda^2 d_1^4 + d_2^2}, & \lambda \neq 0 \\ 0, & \lambda = 0 \end{cases}$$

- For $d_2 = 0$, $f_d(\lambda) = 0$ for all λ, so f_d is continuous.
- For $d_2 \neq 0$, $\lim_{\lambda \to 0} f_d(\lambda) = 0 = f_d(0)$, so f_d is continuous.

Thus, all directional sections f_d are continuous. However, consider the sequence $a_k = \left(\frac{1}{k}, \frac{1}{k^2}\right)$:

$$f(a_k) = \frac{\left(\frac{1}{k}\right)^2 \cdot \left(\frac{1}{k^2}\right)}{\left(\frac{1}{k}\right)^4 + \left(\frac{1}{k^2}\right)^2} = \frac{\frac{1}{k^4}}{\frac{1}{k^4} + \frac{1}{k^4}} = \frac{1}{2}.$$

But $f(0,0) = 0$, so $\lim_{k \to \infty} f(a_k) = \frac{1}{2} \neq f(0,0)$. Therefore, f is not continuous at $(0,0)$.

3.5 Partial Derivatives

Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$, and consider "directional cross-sections". For a direction $d \in \mathbb{R}^n \setminus \{0\}$ and a point $x^* \in \text{int}(X)$, define the function:

$$f_d : \mathbb{R} \to \mathbb{R}, \quad f_d(\lambda) := f(x^* + \lambda d).$$

This represents f along the line through x^* in direction d.

In particular, we can choose coordinate axes as directions, leading to partial functions:

$$f_i(\lambda) := f(x^* + \lambda e_i),$$

where e_i is the i-th standard basis vector.

Example:

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

$$f(x_1, x_2) = 7 - x_1^2 - x_2^2 + 3x_1x_2, \quad x^* = (0, 0).$$

(7) For directions e_1 and e_2:

$$f_1(\lambda) = f(\lambda, 0) = 7 - \lambda^2,$$
$$f_2(\lambda) = f(0, \lambda) = 7 - \lambda^2.$$

Along these axes, f attains a maximum at $x^* = (0, 0)$.

(8) For direction $d = (1, 1)$:

$$f_d(\lambda) = f(\lambda, \lambda) = 7 - \lambda^2 - \lambda^2 + 3\lambda^2 = 7 + \lambda^2.$$

Here, f attains a minimum at $x^* = (0, 0)$ along this line.

This shows that one-dimensional cross-sections may not capture all characteristics of a multivariate function.

Definition: Partial Differentiability

Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$, and $x^* \in \text{int}(X)$.

(1) The function f is *partially differentiable* at x^* with respect to x_i if the function:

$$g(\lambda) = f(x^* + \lambda e_i) = f(x_1^*, \ldots, x_{i-1}^*, x_i^* + \lambda, x_{i+1}^*, \ldots, x_n^*),$$

is differentiable at $\lambda = 0$. The partial derivative is then:

$$g'(0) = \frac{\partial f}{\partial x_i}(x^*) = \partial_i f(x^*).$$

(2) The function f is *partially differentiable* at x^* if all partial derivatives $\frac{\partial f}{\partial x_i}(x^*)$ exist. The *gradient* of f at x^* is:

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

(3) Higher-order partial derivatives are defined recursively:

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

(4) The *Hessian matrix* of f at x^* is the matrix of second-order partial derivatives:

$$H_f(x^*) = D^2 f(x^*) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) \right)_{i,j=1}^n.$$

3.6 Directional Derivatives

Partial derivatives describe the behavior of a function along lines parallel to the coordinate axes. But what about general directions?

Definition

Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$, and $x^* \in \text{int}(X)$.

For a direction $v \in \mathbb{R}^n \setminus \{0\}$, choose $\delta > 0$ such that the neighborhood $N_\delta(x^*) \subseteq X$. Define the function:

$$g : (-\delta, \delta) \to \mathbb{R}, \quad g(\lambda) = f(x^* + \lambda v).$$

If g is differentiable at $\lambda = 0$, then $g'(0)$ is called the *directional derivative* of f at x^* in direction v, denoted by:

$$g'(0) = \partial_v f(x^*) = D_v f(x^*).$$

Example

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

$$f(x_1, x_2) = \begin{cases} \frac{x_1^2 x_2}{x_1^4 + x_2^2}, & (x_1, x_2) \neq (0,0) \\ 0, & (x_1, x_2) = (0,0) \end{cases}$$

This function is not continuous at $(0,0)$, but it is continuous along every line through $(0,0)$.

Directional Derivatives at $(0,0)$

For a direction $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, define:

$$g(\lambda) = f(0 + \lambda v) = f(\lambda v_1, \lambda v_2) = \begin{cases} \frac{\lambda^3 v_1^2 v_2}{\lambda^4 v_1^4 + \lambda^2 v_2^2} = \frac{\lambda v_1^2 v_2}{\lambda^2 v_1^4 + v_2^2}, & \lambda \neq 0 \\ 0, & \lambda = 0 \end{cases}$$

The difference quotient is:

$$\frac{g(\lambda) - g(0)}{\lambda - 0} = \begin{cases} \frac{v_1^2 v_2}{\lambda^2 v_1^4 + v_2^2}, & \lambda \neq 0 \\ 0, & \lambda = 0 \end{cases}$$

Taking the limit as $\lambda \to 0$:

$$\lim_{\lambda \to 0} \frac{g(\lambda) - g(0)}{\lambda} = \begin{cases} 0, & v_2 = 0 \\ \frac{v_1^2}{v_2}, & v_2 \neq 0 \end{cases}$$

Conclusion

All directional derivatives of f at $(0,0)$ exist. However, this does not imply continuity at $(0,0)$, as shown by the earlier example.

This demonstrates that even if a function has directional derivatives in all directions, it may still fail to be continuous at a point.

Economic Application of Multivariable Functions

Keynes' Consumption Formula

$$e = c_0 + \beta \cdot I + \varepsilon_i$$

Minimizing $\sum \varepsilon_i^2$ (Least Squares Regression)

Cobb-Douglas Production Function (Economics I)

$$Y = \alpha \cdot L^b \cdot K^a$$

our objective is to minimize $f : \mathbb{R}^3 \to \mathbb{R}$

4 Total Differentiability

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}$, and $x^* \in X$.

The function f is called *(totally) differentiable* at x^* if there exists a vector $g \in \mathbb{R}^n$ and a function $r : N_\delta(0) \to \mathbb{R}$ (for some δ-neighborhood of 0) such that:

$$f(x) = f(x^*) + g^T(x - x^*) + r(x - x^*)$$

with the property:

$$\lim_{x \to x^*} \frac{r(x - x^*)}{\|x - x^*\|} = 0.$$

Interpretation

- The vector g provides a linear approximation to f at x^*.
- Locally (around x^*), the approximation is "good", meaning the error term r tends to 0 faster than the distance between x and x^* as $x \to x^*$.
- This implies the error is small compared to $\|x - x^*\|$.

Notation

$$\begin{aligned} g^T = Df(x^*) \quad &\text{is called the derivative of } f \text{ at } x^* \\ &= f'(x^*) \end{aligned}$$

Total Differentiability

A function is called *(totally) differentiable* if it is differentiable at every point in its domain.

This definition extends the one-dimensional case to higher dimensions. For $f : \mathbb{R} \to \mathbb{R}$ differentiable at $x^* \in \mathbb{R}$ in the usual sense, if f is totally differentiable, there is $g \in \mathbb{R}$ and a function $\gamma : (-\delta, \delta) \to \mathbb{R}$

$$\begin{gathered} f(x) = f(x^*) + g(x - x^*) + r(x - x^*) \\ \implies \gamma(x - x^*) = f(x) - f(x^*) - g(x - x^*) \\ \lim_{x \to x^*} \frac{\gamma(x - x^*)}{|x - x^*|} = 0 \end{gathered}$$

. therefore, after a little bit math we have $g = f'(x^*) \implies Df(x^*) = f'(x^*)$

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}$, and $x^* \in X$.

(1) If f is differentiable at x^*, then f is also continuous at x^*.

(2) If f is differentiable at x^*, then f is also partially differentiable at x^* with

$$\nabla f(x^*) = [Df(x^*)]^T = f'(x^*)$$

where $\nabla f(x^*)$ is the gradient of f at x^*.

(3) If f is partially differentiable on some δ-neighborhood $N_\delta(x^*)$ of x^* and if all partial derivatives

$$\frac{\partial f}{\partial x_i} : N_\delta(x^*) \to \mathbb{R}$$

are continuous at x^*, then f is totally differentiable at x^* and

$$f'(x^*) = \nabla f(x^*) = [Df(x^*)]^T.$$

(4) If all second partial derivatives of f exist and are continuous on some δ-neighborhood $N_\delta(x^*)$, then

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x^*) \quad \text{for all } i, j.$$

This shows the symmetry of mixed partial derivatives under continuity conditions.

4.1 Differentiation of Vector-Valued Functions

The concept of total differentiability can be extended to vector-valued functions by considering the differential as a locally good linear approximation.

Definition

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}^m$, and $x^* \in X$.

The function f is called *totally differentiable* at x^* if there exists a matrix $J_f \in \mathbb{R}^{m\times n}$ (called the Jacobian matrix) and a function $r : N_\delta(0) \to \mathbb{R}^m$ defined on some δ-neighborhood of 0 such that:

$$f(x) = f(x^*) + J_f(x - x^*) + r(x - x^*)$$

with the property:

$$\lim_{x\to x^*} \frac{r(x - x^*)}{\|x - x^*\|} = 0.$$

The matrix J_f is called the *Jacobian matrix* of f at x^*, denoted by:

$$J_f(x^*) \quad \text{or} \quad Df(x^*), \quad \text{sometimes } f'(x^*).$$

4.1.1 Theorem

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}^m$ with component functions:

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

Then f is differentiable at $x^* \in X$ if and only if each component function f_i is differentiable at x^*.
In this case, the Jacobian matrix of f at x^* is:

$$J_f(x^*) = \begin{pmatrix} \nabla f_1(x^*) \\ \vdots \\ \nabla f_m(x^*) \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{R}^{m\times n}$$

Examples

(a) Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by:

$$f(x_1, x_2) = \begin{pmatrix} e^{x_2} \\ x_1 e^{x_2} \end{pmatrix}$$

The Jacobian matrix is:

$$J_f(x) = \begin{pmatrix} 0 & e^{x_2} \\ e^{x_2} & x_1 e^{x_2} \end{pmatrix}$$

(b) Consider $f : [0, \infty) \to \mathbb{R}^2$ defined by:

$$f(t) = \begin{pmatrix} t\cos t \\ t \sin t \end{pmatrix}$$

The Jacobian matrix is:

$$J_f(t) = \begin{pmatrix} \cos t - t\sin t \\ \sin t + t\cos t \end{pmatrix} \in \mathbb{R}^{2\times 1}$$

4.2 Differentiation Rules

Not all rules for taking derivatives that we know from one-dimensional theory can be guaranteed to hold in multivariable vector-valued theory.

Theorem

Let $f \in \mathbb{R}^n$ be open, $x^* \in f$, and let $f, g : x \to \mathbb{R}$ be differentiable at x^*.

(1) $(f+g)$ and (fg) are differentiable at x^* with

$$D(f+g)(x^*) = Df(x^*) + Dg(x^*).$$

$$D(f \cdot g)(x^*) = g(x^*) \cdot Df(x^*) + f(x^*) \cdot Dg(x^*)$$

(2) Let $Y \subseteq \mathbb{R}^m$ be open, $h : Y \to X$ be differentiable at $y^* \in Y$ where $h(y^*) = x^*$. Then $(f \circ h) : Y \to \mathbb{R}$ is differentiable at y^* and

$$D(f \circ h)(y^*) = Df(h(y^*)) \cdot Dh(y^*)$$

$$D(f \circ h)(y^*) = [\nabla f(h(y^*))]^T \cdot j_h(y^*)$$

$$\Rightarrow \nabla(f \circ h)(y^*) = [Df(h(y^*))]^T$$

$$\Rightarrow \nabla(f \circ h)(y^*) = [j_h(y^*)]^T \cdot \nabla f(h(y^*))$$

Examples

a) Directional Derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable, $v \in \mathbb{R}^n \setminus \{0\}$ a direction, and $t^* \in \mathbb{R}^n$.
Define $h : \mathbb{R} \to \mathbb{R}^n$ as $h(\lambda) = t^* + \lambda v$ (parametrization of the line in direction v through t^*).
The composition $f \circ h$ is the restriction of f to that line.
h is differentiable with

$$h(\lambda) = \begin{pmatrix} x_1^* + \lambda v_1 \\ \vdots \\ x_n^* + \lambda v_n \end{pmatrix} \Rightarrow Dh(\lambda) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = v.$$

$$\Rightarrow D(f \circ h)(\lambda) = Df(h(\lambda)) \cdot Dh(\lambda)$$

$$= [Df(h(\lambda))]^T \cdot v = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(h(\lambda)) \cdot v_i.$$

This is exactly the derivative of f along the line defined by h, i.e., the directional derivative of f at t^* in direction v.

$$\Rightarrow \partial_v f(t^*) = D(f \circ h)(0) = [Df(t^*)]^T \cdot v, \quad h(0) = t^*.$$

Thus, directional derivatives are just linear combinations of partial derivatives.

b) Example Calculation

Let $f(x_1, x_2) = 7 - x_1^2 - x_2^2 + 3x_1x_2$ and $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

$$Df(x_1, x_2) = \begin{pmatrix} -2x_1 + 3x_2 \\ -2x_2 + 3x_1 \end{pmatrix}.$$

$$\Rightarrow \partial_v f(x) = [Df(x_1, x_2)]^T \cdot v$$

$$= \begin{pmatrix} -2x_1 + 3x_2 \\ -2x_2 + 3x_1 \end{pmatrix}^T \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$= (-2x_1 + 3x_2) \cdot 1 + (-2x_2 + 3x_1) \cdot 2 = -2x_1 + 3x_2 - 4x_2 + 6x_1 = 4x_1 - x_2.$$

Problem

Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x) = x^\top A x$ with

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix}.$$

Determine a representation of the tangential plane P to the graph of f at the point $x^* = (1,1)^\top$ in:

- Implicit (normal) form, i.e., as the solution set of a system of linear equations.
- Explicit (parameter) form, i.e., as a linear/affine combination of some vectors.

Solution

By definition of the total differential, with $g = f'(x^*)$, the tangential plane to the graph of f at x^* is given by the linear approximation

$$x \mapsto f(x^*) + g \cdot (x - x^*).$$

First, we simplify f:

$$f(x) = x^\top A x = (x_1, x_2) \begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} 2x_1 + x_2 \\ 4x_2 \end{pmatrix} = 2x_1^2 + 4x_2^2 + x_1 x_2.$$

Next, we compute the gradient of f:

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 + x_2 \\ 8x_2 + x_1 \end{pmatrix}, \quad \text{thus} \quad \nabla f(1,1) = \begin{pmatrix} 5 \\ 9 \end{pmatrix}.$$

Implicit Form

A point $(x_1, x_2, x_3)^\top$ on the tangential plane satisfies:

$$x_3 = f(1,1) + (5,9)\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$$

$$\Leftrightarrow x_3 = 7 + 5(x_1 - 1) + 9(x_2 - 1)$$

$$\Leftrightarrow 5x_1 + 9x_2 - x_3 = 7.$$

Parameter Form

To find a parameter representation, we first determine two linearly independent vectors perpendicular to the normal vector $(5, 9, -1)^\top$. We choose:

$$\begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \\ 9 \end{pmatrix}.$$

A particular point on the plane is $(1,1,7)^\top$. Thus, the parameter form is:

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 7 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 9 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} = \begin{pmatrix} 1 \\ 1 \\ 7 \end{pmatrix} + \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 9 \end{pmatrix} \right\}.$$

Problem

$f : \mathbb{R}^3 \to \mathbb{R}^3$ and $g : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$f(x_1, x_2, x_3) = x_1^2 e^{x_1 \cdot x_2}$$

$$g(x_1, x_2) = \begin{pmatrix} 2x_1 + x_2 \\ x_1 - x_2 \\ x_1 \end{pmatrix}$$

Solution

$$\nabla f(x_1, x_2, x_3) = x \begin{pmatrix} 2x_1 e^{x_2 \cdot x_3} \\ x_1^2 x_3 e^{x_2 \cdot x_3} \\ x_1^2 x_2 e^{x_2 \cdot x_3} \end{pmatrix} \Rightarrow J_f(x_1, x_2, x_3) = x_1 e^{x_2 \cdot x_3} \cdot \begin{pmatrix} 2, & x_1 x_3, & x_1 x_2 \end{pmatrix}$$

$$J_g(x_1, x_2) = \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}$$

The differential of the function $f \circ g : \mathbb{R}^2 \to \mathbb{R}$ is a 1×2 vector (the transpose of the gradient of that function). It can be computed using the chain rule:

$$D(f \circ g)(x) = Df(g(x)) \cdot Dg(x).$$

Remember that $Df(x) = (\nabla f(x))^\top = J_f(x)$.

$$D(f \circ g)(x_1, x_2) = J_f(g(x_1, x_2)) \cdot J_g(x_1, x_2)$$

$$= J_f(2x_1 + x_2, x_1 - x_2, x_1) \cdot \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}$$

$$= (2x_1 + x_2) e^{(x_1 - x_2) \cdot x_1} \cdot \begin{pmatrix} 2 & (2x_1 + x_2)x_1 & (2x_1 + x_2)(x_1 - x_2) \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix}$$

$$= (2x_1 + x_2) e^{x_1 (x_1 - x_2)} \cdot \begin{pmatrix} 4 + (2x_1 + x_2)x_1 + (2x_1 + x_2)(x_1 - x_2) & 2 - (2x_1 + x_2)x_1 \end{pmatrix}$$

5 Polar Coordinates, Cylindrical and Spherical Coordinates

5.1 Polar Coordinates

A point in the polar coordinate system is described by:

- Distance ρ from the origin (radius)
- Angle φ between the x_1-axis and the vector pointing to the point

5.1.1 Relation to Cartesian Coordinates

For a point $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, the conversion is:

$$x_1 = \rho \cdot \cos\varphi$$

$$x_2 = \rho \cdot \sin\varphi$$

If $f : \mathbb{R}^2 \to \mathbb{R}$ is given in Cartesian coordinates, we can express it in polar coordinates using the mapping $h : \mathbb{R}^2 \to \mathbb{R}^2$:

$$h(\rho, \varphi) = \begin{pmatrix} \rho\cos\varphi \\ \rho\sin\varphi \end{pmatrix}$$

Then $f \circ h$ represents f in polar form.

5.1.2 Chain Rule Application

The derivative of the composition is:

$$D(f \circ h)(\rho, \varphi) = [Df(h(\rho, \varphi))]^T \cdot J_h(\rho, \varphi)$$

where the Jacobian matrix is:

$$J_h(\rho, \varphi) = \begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}$$

5.1.3 Example

Let $f(x_1, x_2) = x_1^2 + x_2^2$ with gradient:

$$Df(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

Then:

$$[D(f \circ h)(\rho, \varphi)]^T = \begin{pmatrix} 2\rho\cos\varphi \\ 2\rho\sin\varphi \end{pmatrix}^T \cdot \begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}$$

$$= \left(2\rho(\cos^2\varphi + \sin^2\varphi), -2\rho^2\sin\varphi\cos\varphi + 2\rho^2\sin\varphi\cos\varphi\right)$$

$$= (2\rho, 0)$$

This matches the polar form:

$$(f \circ h)(\rho, \varphi) = \rho^2$$

5.2 Cylindrical Coordinates

Extension of polar coordinates with height h:

$$T(\rho, \varphi, h) = \begin{pmatrix} \rho\cos\varphi \\ \rho\sin\varphi \\ h \end{pmatrix}$$

describes a point on a cylinder of radius ρ.

5.3 Spherical Coordinates

A point is described by:

- Radius S
- Azimuthal angle θ (in x_1-x_2 plane)
- Polar angle ψ (from x_3-axis)

$$T(\rho, \theta, \psi) = \begin{pmatrix} \rho \cos\theta \sin\psi \\ \rho \sin\theta \sin\psi \\ \rho \cos\psi \end{pmatrix}$$

6 Taylor Approximation

Definition
Let $X \subseteq \mathbb{R}^n$ be open, and $f : X \to \mathbb{R}$. Then f is called k-**times continuously differentiable** if and only if all partial derivatives of order k exist and are continuous on X, i.e.,

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} \quad \text{exist and are continuous on } X.$$

The set of all k-times continuously differentiable functions $f : X \to \mathbb{R}$ is denoted by $C^k(X)$.
$C^k(X)$ forms a vector space under pointwise addition and scalar multiplication.

Multi-Index Notation

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}_0^n$ be a multi-index. We define:

- Order: $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n$
- Factorial: $\alpha! := \alpha_1! \cdot \alpha_2! \cdots \alpha_n!$
- For $\boldsymbol{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$:
$$\boldsymbol{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$
- For $f \in C^{|\alpha|}(X)$, the partial derivative operator:

$$D^\alpha f(\boldsymbol{x}) := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}(\boldsymbol{x})$$

Example Calculation

Consider $f : \mathbb{R}^2 \to \mathbb{R}$ with:

$$f(x_1, x_2) = 3x_1x_2 + x_1^2 - 2x_2^2$$

We examine derivatives for different multi-indices $\alpha = (\alpha_1, \alpha_2)$:

6.0.1 Case $|\alpha| = 0$

$$\alpha = (0,0), \quad \alpha! = 0! \cdot 0! = 1, \quad \boldsymbol{x}^\alpha = 1$$
$$D^\alpha f(\boldsymbol{x}) = f(\boldsymbol{x}) = 3x_1x_2 + x_1^2 - 2x_2^2$$

6.0.2 Case $|\alpha| = 1$

- $\alpha = (1,0)$:
$$\alpha! = 1, \quad \boldsymbol{x}^\alpha = x_1, \quad D^\alpha f(\boldsymbol{x}) = \frac{\partial f}{\partial x_1} = 3x_2 + 2x_1$$
- $\alpha = (0,1)$:
$$\alpha! = 1, \quad \boldsymbol{x}^\alpha = x_2, \quad D^\alpha f(\boldsymbol{x}) = \frac{\partial f}{\partial x_2} = 3x_1 - 4x_2$$

6.0.3 Case $|\alpha| = 2$

- $\alpha = (2,0)$:
$$\alpha! = 2, \quad \boldsymbol{x}^\alpha = x_1^2, \quad D^\alpha f(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x_1^2} = 2$$
- $\alpha = (0,2)$:
$$\alpha! = 2, \quad \boldsymbol{x}^\alpha = x_2^2, \quad D^\alpha f(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x_2^2} = -4$$
- $\alpha = (1,1)$:
$$\alpha! = 1, \quad \boldsymbol{x}^\alpha = x_1x_2, \quad D^\alpha f(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x_1 \partial x_2} = 3$$

6.0.4 Case $|\alpha| = 3$

All third-order derivatives of f are identically zero.

Definitions

Let $X \subseteq \mathbb{R}^n$ be open, $x^* \in X$, and $f \in C^k(X)$. The k-**th order Taylor polynomial** of f at x^* is defined as:

$$T_k f(x; x^*) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x^*)}{\alpha!}(x - x^*)^\alpha$$

where $\alpha \in \mathbb{N}_0^n$ is a multi-index.

For $f \in C^\infty(X)$, the **Taylor series** of f at x^* is:

$$T_f(x; x^*) = \sum_{\alpha \in \mathbb{N}_0^n} \frac{D^\alpha f(x^*)}{\alpha!}(x - x^*)^\alpha$$

Example Calculation

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

$$f(x_1, x_2) = 3x_1x_2 + x_1^2 - 2x_2^2$$

The second-order Taylor polynomial at $x^* = (x_1^*, x_2^*)$ is:

$$T_2 f(x; x^*) = \sum_{|\alpha| \le 2} \frac{D^\alpha f(x^*)}{\alpha!}(x - x^*)^\alpha$$

Expanded form:

$$\begin{aligned} T_2 f(x; x^*) = f(x^*) & \\ & + \frac{\partial f(x^*)}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f(x^*)}{\partial x_2}(x_2 - x_2^*) \\ & + \frac{1}{2}\frac{\partial^2 f(x^*)}{\partial x_1^2}(x_1 - x_1^*)^2 + \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2}(x_1 - x_1^*)(x_2 - x_2^*) \\ & + \frac{1}{2}\frac{\partial^2 f(x^*)}{\partial x_2^2}(x_2 - x_2^*)^2 \end{aligned}$$

Substituting derivatives:

$$= \left[3x_1^*x_2^* + (x_1^*)^2 - 2(x_2^*)^2\right] + \left[3x_2^* + 2x_1^*\right](x_1 - x_1^*) + \left[3x_1^* - 4x_2^*\right](x_2 - x_2^*)$$

$$+\frac{1}{2} \cdot 2(x_1 - x_1^*)^2 + 3(x_1 - x_1^*)(x_2 - x_2^*) + \frac{1}{2} \cdot (-4)(x_2 - x_2^*)^2$$

Taylor's Theorem

Let $X \subseteq \mathbb{R}^n$ be open, $x^* \in X$, $f \in e^{k+1}(X)$, and $x \in X$ such that the line segment connecting x^* and x lies entirely in X. Then there exists $\tilde{x}$ on this line segment with:

$$f(x) = T_k f(x; x^*) + \sum_{|\alpha| = k+1} \frac{D^\alpha f(\tilde{x})}{\alpha!}(x - x^*)^\alpha$$

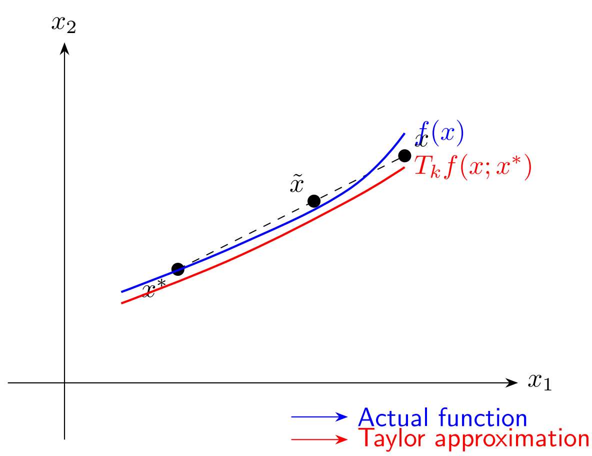


Figure 1: Visualization of Taylor's theorem in $\mathbb{R}^2$. The blue curve represents the actual function f, while the red curve shows its k-th order Taylor approximation $T_k f$ at x^*. The point $\tilde{x}$ lies on the line segment between x^* and x.

Problem

Compute the Taylor polynomial of order 3 at $x_0 = (1,0)^T$ for the function

$$f(x_1, x_2) = e^{x_1+x_2} + e^{x_1-x_2}.$$

Solution

We compute the gradient, the Hessian, and the partial derivatives of third order, evaluating them at the point x_0:

First Derivatives (Gradient)

$$\nabla f(x) = \begin{pmatrix} e^{x_1+x_2} + e^{x_1-x_2} \\ e^{x_1+x_2} - e^{x_1-x_2} \end{pmatrix}$$

At $x_0 = (1,0)$:

$$\nabla f(1,0) = \begin{pmatrix} e+1 \\ e-1 \end{pmatrix}$$

Second Derivatives (Hessian)

$$H_f(x) = \begin{pmatrix} e^{x_1+x_2} + e^{x_1-x_2} & e^{x_1+x_2} - e^{x_1-x_2} \\ e^{x_1+x_2} - e^{x_1-x_2} & e^{x_1+x_2} + e^{x_1-x_2} \end{pmatrix}$$

At $x_0 = (1,0)$:

$$H_f(1,0) = \begin{pmatrix} e+1 & e-1 \\ e-1 & e+1 \end{pmatrix}$$

Third Derivatives

$$\begin{aligned}
\frac{\partial^3 f}{\partial x_1^3}(x) &= e^{x_1+x_2} + e^{x_1-x_2} \\
\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(x) &= e^{x_1+x_2} - e^{x_1-x_2} \\
\frac{\partial^3 f}{\partial x_1 \partial x_2^2}(x) &= e^{x_1+x_2} + e^{x_1-x_2} \\
\frac{\partial^3 f}{\partial x_2^3}(x) &= e^{x_1+x_2} - e^{x_1-x_2}
\end{aligned}$$

At $x_0 = (1,0)$:

$$\begin{aligned}
\frac{\partial^3 f}{\partial x_1^3}(1,0) &= e+1 \\
\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(1,0) &= e-1 \\
\frac{\partial^3 f}{\partial x_1 \partial x_2^2}(1,0) &= e+1 \\
\frac{\partial^3 f}{\partial x_2^3}(1,0) &= e-1
\end{aligned}$$

Taylor Polynomial of Order 3

The Taylor polynomial is given by:

$$\begin{aligned}
T_3 f(x; x_0) = f(1,0) + \nabla f(1,0)^T \begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x_1 - 1 & x_2 \end{pmatrix} H_f(1,0) \begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} \\
+ \frac{1}{6} \left[\frac{\partial^3 f}{\partial x_1^3}(1,0)(x_1-1)^3 + 3 \frac{\partial^3 f}{\partial x_1^2 \partial x_2}(1,0)(x_1-1)^2 x_2 \right. \\
\left. + 3 \frac{\partial^3 f}{\partial x_1 \partial x_2^2}(1,0)(x_1-1) x_2^2 + \frac{\partial^3 f}{\partial x_2^3}(1,0) x_2^3 \right]
\end{aligned}$$

Substituting the computed values:

$$
\begin{aligned}
T_3 f(x; x_0) = (e+1) &+ [e(x_1-1) + (e-1)x_2] \\
&+ \frac{1}{2}\left[(e+1)(x_1-1)^2 + 2(e-1)(x_1-1)x_2 + (e+1)x_2^2\right] \\
&+ \frac{1}{6}\left[(e+1)(x_1-1)^3 + 3(e-1)(x_1-1)^2 x_2 + 3(e+1)(x_1-1)x_2^2 + (e-1)x_2^3\right]
\end{aligned}
$$

Problem

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function

$$f(x_1, x_2) = e^{x_1^2 \cdot x_2} + \cos(\pi \cdot x_1 \cdot x_2).$$

Determine a tangential plane to the graph of f at the point $(1, -1)^\top$ in the form

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \beta.$$

Solution

To find the tangential plane, we compute the first-order Taylor approximation (linear approximation) of f at $(1, -1)^\top$.

1. Compute the Gradient The gradient of f is:

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 x_2 e^{x_1^2 x_2} - \pi x_2 \sin(\pi x_1 x_2) \\ x_1^2 e^{x_1^2 x_2} - \pi x_1 \sin(\pi x_1 x_2) \end{pmatrix}.$$

2. Evaluate at $(1, -1)^\top$

$$
\begin{aligned}
f(1,-1) &= e^{-1} + \cos(-\pi) = e^{-1} - 1, \\
\nabla f(1,-1) &= \begin{pmatrix} -2e^{-1} + \pi \sin(-\pi) \\ e^{-1} - \pi \sin(-\pi) \end{pmatrix} = \begin{pmatrix} -2e^{-1} \\ e^{-1} \end{pmatrix}.
\end{aligned}
$$

3. Construct the Tangential Plane

The equation of the tangential plane is:

$$x_3 = f(1,-1) + \nabla f(1,-1)^\top \cdot \begin{pmatrix} x_1 - 1 \\ x_2 + 1 \end{pmatrix},$$

which simplifies to:

$$x_3 = e^{-1} - 1 - 2e^{-1}(x_1 - 1) + e^{-1}(x_2 + 1).$$

Rearranging terms:

$$
\begin{aligned}
x_3 = e^{-1} - 1 + 2e^{-1} + e^{-1} - 2e^{-1}x_1 + e^{-1}x_2, \\
2e^{-1}x_1 - e^{-1}x_2 + x_3 = 4e^{-1} - 1.
\end{aligned}
$$

4. Final Form

The coefficients are:

$$\alpha_1 = 2e^{-1}, \quad \alpha_2 = -e^{-1}, \quad \alpha_3 = 1, \quad \beta = 4e^{-1} - 1.$$

Thus, the tangential plane equation is:

$$2e^{-1}x_1 - e^{-1}x_2 + x_3 = 4e^{-1} - 1.$$

7 Local Optima of Multivariate Functions

Definition

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}$.

(1) A point $x^* \in X$ is called a *(strict) local minimum* of f if there is a δ-neighborhood $N_\delta(x^*)$ of x^* such that

$$f(x^*) \leq f(x), \quad \text{for all } x \in X \cap N_\delta(x^*).$$

(2) A point $x^* \in X$ is called a *(strict) local maximum* of f if there is a δ-neighborhood $N_\delta(x^*)$ of x^* such that

$$f(x^*) \geq f(x), \quad \text{for all } x \in X \cap N_\delta(x^*).$$

(3) A point x^* is called a *(strict) global maximizer/minimizer* of f if

$$f(x^*) \geq f(x) \text{ for all } x \in X$$

or

$$f(x^*) \leq f(x) \text{ for all } x \in X,$$

respectively.

(4) The respective function values are called *(strict) local/global minimum/maximum* of f.

(5) A local/global minimum or maximum is also referred to as an *extremum* or *optimum* of f.

7.1 Optimality Condition

If x^* is a local minimizer and x^* is an interior point, then it can be analyzed as follows:
Let

$$x := x^* - \epsilon \nabla f(x^*),$$

while

$$f(x) = f(x^*) + \nabla f(x^*)^T (x - x^*) + r(x - x^*).$$

Substituting x:

$$\begin{aligned} f(x) &= f(x^*) + Df(x^*)^T(-\varepsilon \nabla f(x^*)) + r(x - x^*) \\ &= f(x^*) - \epsilon \|\nabla f(x^*)\|^2 + r(x - x^*). \end{aligned}$$

As $r(x - x^*) \to 0$ when $x \to x^*$, this implies that for sufficiently small ϵ, the term $-\epsilon\|\nabla f(x^*)\|^2$ dominates, leading to:

$$f(x) < f(x^*),$$

unless $\nabla f(x^*) = 0$.

Theorem

Let $X \subseteq \mathbb{R}^n$ be open, $x^* \in X$, and $f : X \to \mathbb{R}$ differentiable at x^*. If x^* is a local optimum of f, then

$$\nabla f(x^*) = 0.$$

This is a necessary, but not sufficient, condition for optimality.

Definition

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}$ differentiable.

(1) A point $x^* \in X$ with $\nabla f(x^*) = 0$ is called a *stationary point* of f.

(2) If $x^* \in X$ is a stationary point but not a local extremum, it is called a *saddle point*.

7.2 Local Optima of Multivariate Functions

Problem Setup

Assume $f \in e^2(X)$ for $X \subseteq \mathbb{R}^n$ and $x^* \in X$ with $Df(x^*) = 0$. How can we determine whether x^* is a local minimum, local maximum, or a saddle point?

7.2.1 Taylor Expansion Approach

Using the second-order Taylor expansion around x^*:

$$f(x) = f(x^*) + Df(x^*)^T(x - x^*) + \frac{1}{2}(x - x^*)^T H_f(x^*)(x - x^*) + \cdots,$$

where $H_f(x^*)$ is the Hessian matrix of f at x^*. Since $Df(x^*) = 0$, the local behavior of f at x^* is governed by the quadratic form:

$$(x - x^*)^T H_f(x^*)(x - x^*).$$

7.2.2 Definition of Definite Matrices

Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then A is called:

1. *positive definite* if $v^T A v > 0$ for all $v \in \mathbb{R}^n \setminus \{0\}$;
2. *positive semidefinite* if $v^T A v \geq 0$ for all $v \in \mathbb{R}^n \setminus \{0\}$;
3. *negative definite* if $v^T A v < 0$ for all $v \in \mathbb{R}^n \setminus \{0\}$;
4. *negative semidefinite* if $v^T A v \leq 0$ for all $v \in \mathbb{R}^n \setminus \{0\}$;
5. *indefinite* if there exist $v, w \in \mathbb{R}^n \setminus \{0\}$ such that $v^T A v > 0$ and $w^T A w < 0$.

Remark

A matrix A is negative definite if and only if $-A$ is positive definite.

Theorem: Classification of Stationary Points

Let $X \subseteq \mathbb{R}^n$ be open, $f \in C^2(X)$, and $x^* \in X$ a stationary point (i.e., $Df(x^*) = 0$). Then:

1. If x^* is a local minimizer of f, then $H_f(x^*)$ is positive semidefinite.
2. If $H_f(x^*)$ is positive definite, then x^* is a local minimizer of f.
3. If x^* is a local maximizer of f, then $H_f(x^*)$ is negative semidefinite.
4. If $H_f(x^*)$ is negative definite, then x^* is a local maximizer of f.
5. If $H_f(x^*)$ is indefinite, then x^* is a saddle point.

Example (a)

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

$$f(x_1, x_2) = 2 - (x_1 - 1)^2 - (x_2 - 1)^2.$$

Finding Stationary Points

The gradient is:

$$\nabla f(x) = \begin{pmatrix} -2(x_1 - 1) \\ -2(x_2 - 1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{cases} x_1 = 1, \\ x_2 = 1. \end{cases}$$

Thus, the stationary point is $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Classification

The Hessian matrix is:

$$H_f(x) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

For any $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \neq 0$:

$$v^T H_f(1,1) v = -2v_1^2 - 2v_2^2 < 0.$$

Since the Hessian is *negative definite*, x^* is a *local maximum*.

Example (b)

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

$$f(x_1, x_2) = 7 - x_1^2 - x_2^2 + 3x_1x_2.$$

Finding Stationary Points

The gradient is:

$$\nabla f(x) = \begin{pmatrix} -2x_1 + 3x_2 \\ -2x_2 + 3x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving the system:

$$\begin{cases} -2x_1 + 3x_2 = 0, \\ 3x_1 - 2x_2 = 0, \end{cases}$$

we find $x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Classification

The Hessian matrix is:

$$H_f(x) = \begin{pmatrix} -2 & 3 \\ 3 & -2 \end{pmatrix}.$$

For $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$:

$$v^T H_f(0,0) v = -2 < 0.$$

For $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$:

$$v^T H_f(0,0) v = -2 - 2 + 6 = 2 > 0.$$

Since the Hessian is *indefinite*, x^* is a *saddle point*.

Example (c)

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

$$f(x_1, x_2) = x_1x_2.$$

Finding Stationary Points

The gradient is:

$$\nabla f(x) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Classification

The Hessian matrix is:

$$H_f(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

For $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$:

$$v^T H_f(0,0) v = 2 > 0.$$

For $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$:

$$v^T H_f(0,0) v = -2 < 0.$$

Since the Hessian is *indefinite*, x^* is a *saddle point*.

Sylvester's Determinant Criterion for Definiteness

Theorem (Sylvester's Criterion)

For a symmetric matrix $A \in \mathbb{R}^{n \times n}$, $A = [a_{ij}]_{i,j=1}^n$, define the leading principal submatrices:

$$A_k := [a_{ij}]_{i,j=1}^k \quad \text{for} \quad k = 1, \ldots, n.$$

(1) A is *positive definite* if and only if all leading principal minors $\det(A_k) > 0$ for $k = 1, \ldots, n$.

(2) A is *negative definite* if and only if:

- $\det(A_k) < 0$ for all odd k,
- $\det(A_k) > 0$ for all even k.

Remark

A is negative definite if and only if $-A$ is positive definite. This implies:

$$\det(-A_k) = (-1)^k \det(A_k) > 0 \quad \text{for all} \quad k.$$

Examples

Example (a)

Consider the matrix:

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

- $\det(A_1) = \det(2) = 2 > 0$,
- $\det(A_2) = \det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4 > 0$.

Since all leading principal minors are positive, A is *positive definite*.

Example (b)

Consider the matrix:

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

- $\det(A_1) = \det(-2) = -2 < 0$,
- $\det(A_2) = \det \begin{pmatrix} -2 & 3 \\ 3 & -2 \end{pmatrix} = 4 - 9 = -5 < 0$.

Since $\det(A_1) < 0$ (odd k) but $\det(A_2) < 0$ (even k), the conditions for negative definiteness are *not* satisfied. Instead, A is *indefinite*.

Example (c)

Consider the matrix:

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

- $\det(A_1) = \det(2) = 2 > 0$,
- $\det(A_2) = \det\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 4 - 1 = 3 > 0$,
- $\det(A_3) = \det\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

  To compute $\det(A_3)$:

$$\begin{aligned}\det(A_3) &= 2 \cdot \det\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} - (-1) \cdot \det\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} + (-1) \cdot \det\begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix} \\ &= 2(2 \cdot 1 - 1 \cdot 1) + 1(-1 \cdot 1 - 1 \cdot (-1)) - 1(-1 \cdot 1 - 2 \cdot (-1)) \\ &= 2(1) + 1(0) - 1(1) = 2 + 0 - 1 = 1 > 0.\end{aligned}$$

Since all leading principal minors are positive, A is *positive definite*.

8 Numerical Optimization Methods

determining stationary points of functions may be very difficult or maybe not even possible - think of a function implicitly defined as the solution of a convoluted differential equation system. Hence, in practice numerical methods that approximate an optimal solution are used frequently

Interpretation of the Gradient

Consider $f : X \to \mathbb{R}$ differentiable at $x^* \in X$. For a direction $v \in \mathbb{R}^n \setminus \{0\}$ with $\|v\| = 1$, the directional derivative $\partial_v f(x^*)$ represents the local rate of change of f at x^* in direction v. It satisfies:

$$|\partial_v f(x^*)| = |(\nabla f(x^*))^T v| \leq \|\nabla f(x^*)\| \cdot \|v\| = \|\nabla f(x^*)\|$$

Steepest Ascent/Descent Directions

The norm of the gradient $\|\nabla f(x^*)\|$ measures the maximum local rate of change of f at x^*.

For the normalized direction:

$$v = \frac{\nabla f(x^*)}{\|\nabla f(x^*)\|}$$

we obtain the maximal directional derivative:

$$\partial_v f(x^*) = \nabla f(x^*)^T \cdot \frac{\nabla f(x^*)}{\|\nabla f(x^*)\|} = \|\nabla f(x^*)\|$$

- $\nabla f(x^*)$ points in the direction of *steepest ascent*
- $-\nabla f(x^*)$ points in the direction of *steepest descent*

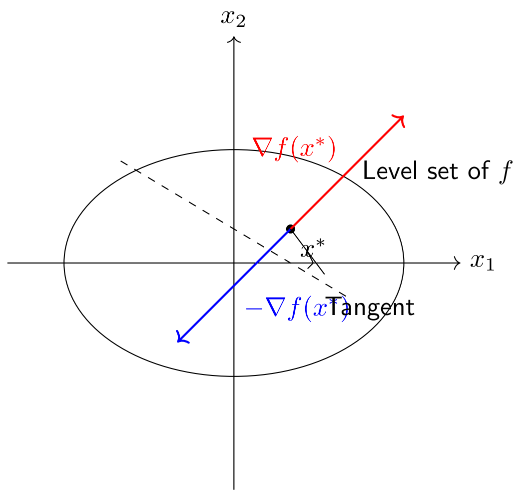
Definition

Let $X \subseteq \mathbb{R}^n$ be open, $f \in C^1(X)$, and $x^* \in X$ with $\nabla f(x^*) \neq 0$. The normalized vectors:

$$\frac{\nabla f(x^*)}{\|\nabla f(x^*)\|} \quad \text{(steepest ascent)}, \quad -\frac{\nabla f(x^*)}{\|\nabla f(x^*)\|} \quad \text{(steepest descent)}$$

give the directions of maximal increase/decrease of f at x^*.

Geometric Interpretation



Key observations:

- The gradient $\nabla f(x^*)$ is orthogonal to the tangent of the level set at x^*
- It points in the direction of greatest increase of f
- The negative gradient points in the direction of greatest decrease

8.1 Algorithm for Gradient Descent Method

1. **Initialize:**
   - Set $k := 0$ (iteration counter)
2. **While** $\nabla f(x^{(k)}) \neq 0$ **or** $\|Df(x^{(k)})\| > \epsilon$:
   - Set search direction: $v := \dfrac{-Df(x^{(k)})}{\|Df(x^{(k)})\|}$
   - Find step size $\lambda \in [0, \infty)$ such that:
   $$f(x^{(k)} + \lambda v) < f(x^{(k)})$$
   - Update: $x^{(k+1)} := x^{(k)} + \lambda v$
   - Increment: $k := k + 1$
3. **End While**

Step Size Selection Methods

- **Line Search:**
  - Determine λ that minimizes $f(x^{(k)} + \lambda v)$
- **Heuristic Approach:**

- Start with $\lambda := 1$
- If $f(x^{(k)} + \lambda v) \geq f(x^{(k)})$:
  * Set $\lambda := \frac{1}{2}\lambda$
  * Repeat until $f(x^{(k)} + \lambda v) < f(x^{(k)})$
  * Keep final λ for next step
- Else:
  * Try $\tilde{\lambda} := 2\lambda$
  * If $f(x^{(k)} + \tilde{\lambda} v) < f(x^{(k)})$:
    · Use $\tilde{\lambda}$ as new step size
  * Else keep original λ

Algorithm Specification

Input:

- $f \in C^1(\mathbb{R}^n)$ (continuously differentiable function)
- Initial value $x \in \mathbb{R}^n$

Output:

- Point $x^* \in \mathbb{R}^n$ approximating a local minimum of f

Gradient descent uses a linear approximation of function f. Can we use a quadratic approximation instead?

$$T_2 f(x; x^{(k)}) = f(x^{(k)}) + Df(x^{(k)})^T (x - x^{(k)}) + \frac{1}{2}(x - x^{(k)})^T H_f(x^{(k)})(x - x^{(k)})$$

Find a stationary point/minimum of $T_2 f(x; x^{(k)})$:

$$0 = \nabla T_2 f(x; x^{(k)}) = \nabla f(x^{(k)}) + H_f(x^{(k)})(x - x^{(k)})$$
$$\Leftrightarrow H_f(x^{(k)})(x - x^{(k)}) = -\nabla f(x^{(k)})$$

$$\underbrace{\text{This is Newton's equation}}_{\text{Update rule: } x^{(k+1)} = x^{(k)} + (x - x^{(k)})}$$

Works well if $H_f(x^{(k)})$ is positive definite $\rightarrow$ local minimum approximation.

Algorithm: Newton's Method

Input:

- $f \in C^2(\mathbb{R}^n)$ (twice differentiable)
- Initial value $x \in \mathbb{R}^n$

Output:

- x^* approximating a local minimizer of f
- Or indication that Newton's equation cannot be solved

1. Initialize $k := 0$
2. While $\|Df(x^{(k)})\| \neq 0$ (or $\|Df(x^{(k)})\| \geq \epsilon$):
   - Find $v \in \mathbb{R}^n$ that solves:
     $$H_f(x^{(k)})v = -Df(x^{(k)})$$

- If no solution exists: terminate with error
- Update:
$$x^{(k+1)} := x^{(k)} + v$$
(Alternative: $x^{(k+1)} := x^{(k)} + \alpha v$ for some step size α)
- Increment: $k := k + 1$

3. Return $x^* := x^{(k)}$

9 Constrained Optimization & the KKT theorem

Problem Setting

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t. } h(x) = 0 \end{aligned}$$

where:

- $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function
- $h : \mathbb{R}^n \to \mathbb{R}^m$ defines constraints

Definitions

- **Feasible set:** $F := \{x \in \mathbb{R}^n : h(x) = 0\}$
- **Lagrange function:** For $z \in \mathbb{R}^m$,
$$L(x, z) := f(x) + h(x)^T z$$

9.1 Theorem (Lagrange Multipliers)

Let $X \subseteq \mathbb{R}^n$ be open, $f : X \to \mathbb{R}$ and $h : X \to \mathbb{R}^k$ be continuously differentiable. Consider:

$$\begin{aligned} &\min_x\ f(x) \\ &\text{s.t. } h(x) = 0 \end{aligned}$$

If $x^* \in F$ is an extremum point of f over F and rank$(Jh(x^*)) = \min\{n, m\}$, then there exist Lagrange multipliers $z^* \in \mathbb{R}^k$ such that:

$$Df(x^*) + Jh(x^*)^T z^* = 0$$
$$\Rightarrow Df(x^*) + \sum_{i=1}^{k} Dh_i(x^*)\zeta_i^* = 0$$

9.2 KKT Conditions for Constrained Optimization

General Problem Setting

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t. } g(x) \le 0 \\ &\qquad h(x) = 0 \end{aligned}$$

where:

- $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function
- $g : \mathbb{R}^n \to \mathbb{R}^m$ defines inequality constraints
- $h : \mathbb{R}^n \to \mathbb{R}^\ell$ defines equality constraints

Definition (KKT Point)

Let $X \subseteq \mathbb{R}^n$ be open, with $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^m$, and $h : X \to \mathbb{R}^\ell$ all continuously differentiable. A point $(x^*, y^*, z^*) \in X \times \mathbb{R}^m \times \mathbb{R}^\ell$ is called a **KKT point** (Karush-Kuhn-Tucker point) if:

KKT Conditions

(1) **Stationarity**:

$$\nabla f(x^*) + \sum_{i=1}^{m} y_i^* \nabla g_i(x^*) + \sum_{j=1}^{\ell} z_j^* \nabla h_j(x^*) = 0$$

(2) **Primal feasibility**:

$$g(x^*) \leq 0 \quad \text{and} \quad h(x^*) = 0$$

(3) **Dual feasibility**:

$$y^* \geq 0$$

(4) **Complementary slackness**:

$$(y^*)^T g(x^*) = 0 \quad \Leftrightarrow \quad y_i^* \cdot g_i(x^*) = 0 \ \forall i \in \{1, \ldots, m\}$$

9.3 Theorem (Necessary Optimality Conditions)

Let $X \subseteq \mathbb{R}^n$ be open, with $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^m$, and $h : X \to \mathbb{R}^\ell$ continuously differentiable. If $x^* \in X$ is a local minimizer and the **constraint qualification** holds:

$$\{\nabla g_i(x^*) : g_i(x^*) = 0\} \cup \{\nabla h_j(x^*)\} \text{ are linearly independent}$$

then there exist $y^* \in \mathbb{R}^m$, $z^* \in \mathbb{R}^\ell$ such that (x^*, y^*, z^*) is a KKT point.

Problem

The constraint $x \in C$ may be expressed as

$$\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|^2 = 9$$

$$\Leftrightarrow (x_1 - 1)^2 + x_2^2 - 9 = 0$$

The optimization problem to solve is therefore given by

$$\min f(x_1, x_2) = 2x_1 + x_2$$

subject to $\quad h(x_1, x_2) = (x_1 - 1)^2 + x_2^2 - 9 = 0.$

Let us first state that the objective is a continuous function and the feasible set is clearly compact, so we can be sure that a minimizer of f over the circle actually exists.

The objective function is convex, so we can try to apply the method of Lagrange multipliers. To do that, we first need the differentials of f and h:

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$J_h(x_1, x_2) = \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix}^\top$$

With Lagrange multiplier $\zeta \in \mathbb{R}$, we get the following condition for a minimizer x^*:

$$\nabla f(x) + J_h(x)^\top \cdot \zeta = 0$$

$$\Leftrightarrow \begin{pmatrix} 2 + 2\zeta(x_1 - 1) \\ 1 + 2\zeta x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Together with the condition $x \in C$ we get the following nonlinear equation system:

$$2\zeta x_1 = -2 + 2\zeta$$
$$2\zeta x_2 = -1$$
$$(x_1 - 1)^2 + x_2^2 = 9$$

Let us first investigate the case $\zeta \neq 0$. We can then solve the first two equations for x_1 and x_2, respectively:

$$x_1 = \frac{-1+\zeta}{\zeta} \quad \text{for } \zeta \neq 0$$
$$x_2 = \frac{-1}{2\zeta} \quad \text{for } \zeta \neq 0$$

Substituting these into the third equation yields

$$\left(\frac{-1+\zeta}{\zeta} - 1\right)^2 + \left(\frac{-1}{2\zeta}\right)^2 = 9$$
$$\Leftrightarrow \left(\frac{-1}{\zeta}\right)^2 + \left(\frac{-1}{2\zeta}\right)^2 = 9$$
$$\Leftrightarrow \frac{1}{\zeta^2} + \frac{1}{4\zeta^2} = 9$$
$$\Leftrightarrow \frac{5}{4\zeta^2} = 9$$
$$\Leftrightarrow \frac{5}{36} = \zeta^2$$

This in turn yields the following candidates for a minimizer:

$$x^* = \begin{pmatrix} \frac{-1+\zeta}{\zeta} \\ \frac{-1}{2\zeta} \end{pmatrix} = \begin{pmatrix} \frac{-1}{\zeta} + 1 \\ \frac{-1}{2\zeta} \end{pmatrix} = \begin{pmatrix} \frac{-6}{\sqrt{5}} + 1 \\ \frac{-1}{2\cdot\frac{\sqrt{5}}{6}} \end{pmatrix} = \begin{pmatrix} \frac{-6+\sqrt{5}}{\sqrt{5}} \\ \frac{-3}{\sqrt{5}} \end{pmatrix} = -\frac{1}{\sqrt{5}} \begin{pmatrix} 6-\sqrt{5} \\ 3 \end{pmatrix}$$
$$x^{**} = \begin{pmatrix} \frac{-1+\zeta}{\zeta} \\ \frac{-1}{2\zeta} \end{pmatrix} = \begin{pmatrix} \frac{-1}{\zeta} + 1 \\ \frac{-1}{2\zeta} \end{pmatrix} = \begin{pmatrix} \frac{6}{\sqrt{5}} + 1 \\ \frac{-1}{2\cdot\frac{-\sqrt{5}}{6}} \end{pmatrix} = \begin{pmatrix} \frac{6+\sqrt{5}}{\sqrt{5}} \\ \frac{3}{\sqrt{5}} \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 6+\sqrt{5} \\ 3 \end{pmatrix}$$

A quick check verifies that both candidate points are contained in the feasible set C:

$$\left\| x^* - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|^2 = \left\| \frac{1}{\sqrt{5}} \begin{pmatrix} 6-\sqrt{5}+\sqrt{5} \\ 3 \end{pmatrix} \right\|^2 = \frac{1}{5}(36+9) = 9 = 3^2$$
$$\left\| x^{**} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|^2 = \left\| \frac{1}{\sqrt{5}} \begin{pmatrix} 6+\sqrt{5}-\sqrt{5} \\ 3 \end{pmatrix} \right\|^2 = \frac{1}{5}(36+9) = 9 = 3^2$$

To find the minimizer of our objective function, we simply compare the objective values of x^* and x^{**}:

$$f(x^*) = \frac{-2\cdot(6-\sqrt{5})}{\sqrt{5}} + \frac{-3}{\sqrt{5}} = \frac{-15+2\sqrt{5}}{\sqrt{5}}$$
$$f(x^{**}) = \frac{2\cdot(6+\sqrt{5})}{\sqrt{5}} + \frac{3}{\sqrt{5}} = \frac{15+2\sqrt{5}}{\sqrt{5}}$$

Clearly, the value for x^* is lower, so this is the desired minimizer of f over the circle C.
Remember, we still have to take care of the case $\zeta = 0$ in our equation system

$$2\zeta x_1 = -2 + 2\zeta$$
$$2\zeta x_2 = -1$$
$$(x_1 - 1)^2 + x_2^2 = 9$$

Substituting that value for ζ yields the system

$$0 = -2$$
$$0 = -1$$
$$(x_1 - 1)^2 + x_2^2 = 9$$

That clearly is a contradiction, thus there is no solution of the equation system for $\zeta = 0$. The above minimizer x^* is therefore unique

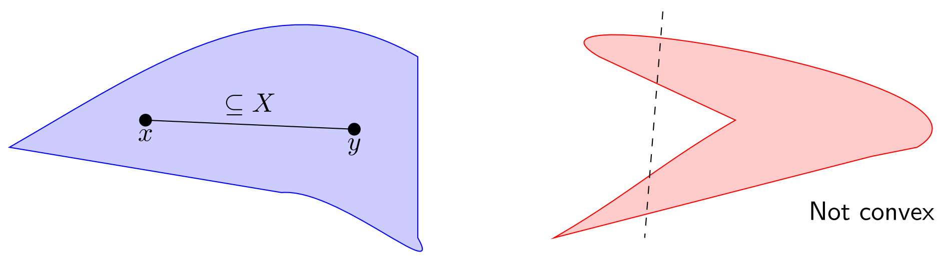
10 Convex Optimization

10.1 Convex Functions and Convex Sets

Definitions

(1) A set $X \subseteq \mathbb{R}^n$ is **convex** if $\forall x, y \in X$, the line segment between them is also contained in X:

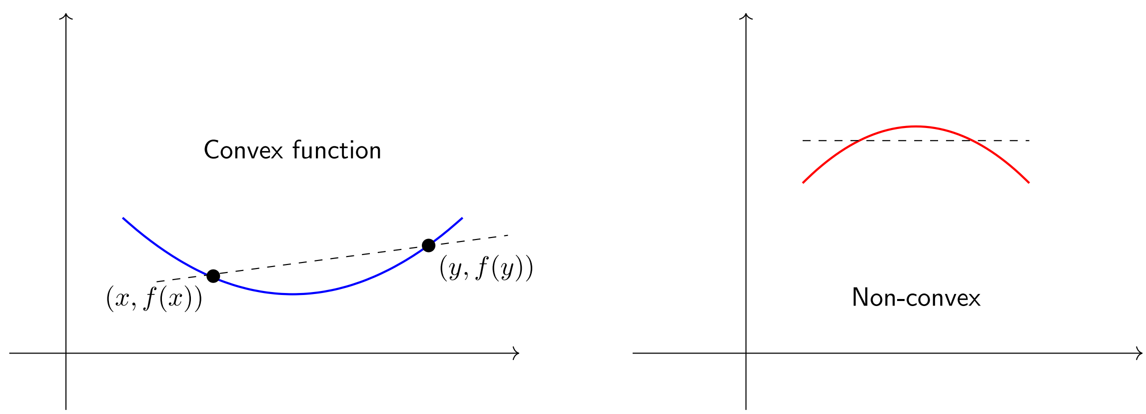
$$\{\lambda x + (1-\lambda)y \mid \lambda \in [0,1]\} \subseteq X$$



(2) For convex X, $f : X \to \mathbb{R}$ is **(strictly) convex** if $\forall x, y \in X$ and $\lambda \in (0,1)$:

$$f(\lambda x + (1-\lambda)y) \ (<) \ \leq \ \lambda f(x) + (1-\lambda) f(y)$$

(3) A function $f : X \to \mathbb{R}$ on a convex set $X \subseteq \mathbb{R}^n$ is called **concave** if $(-f)$ is convex



Theorem (Convexity and Hessian)

Let $X \subseteq \mathbb{R}^n$ be open and convex, $f \in C^2(X)$.

(1) f is convex $\Leftrightarrow$ $H_f(x)$ is positive semidefinite $\forall x \in X$

(2) If $H_f(x)$ is positive definite $\forall x \in X$, then f is strictly convex

Examples

- **Convex:** $f(x) = x^2$, $f(x,y) = x^2 + y^2$
- **Strictly convex:** $f(x) = e^x$, $f(x,y) = x^2 + 2y^2$
- **Non-convex:** $f(x) = \sin x$, $f(x,y) = x^2 - y^2$

10.2 Theorem

Let $X \subseteq \mathbb{R}^n$ be convex and $f : X \to \mathbb{R}$ a convex function.

(1) The set of global minimizers $F := \{x^* \in X : f(x^*) = \min_{x \in X} f(x)\}$ is convex.

(2) Every local minimizer of f is also a global minimizer.

(3) If f is differentiable at $x^* \in X$, then x^* is a minimizer if and only if $\nabla f(x^*) = 0$.

(4) If f is strictly convex, it has at most one (unique) minimizer.

10.3 Theorem

Let $X \subseteq \mathbb{R}^n$ be open and convex, $f : X \to \mathbb{R}$ convex and differentiable, and $g : X \to \mathbb{R}^m$ with each $g_i : X \to \mathbb{R}$ convex and differentiable.

- **Slater Condition**: If $\exists \bar{x} \in X$ such that $g(\bar{x}) < 0$ (strict feasibility), then: $x^* \in X$ is optimal for

$$\min f(x) \quad \text{s.t.} \quad g(x) \leq 0$$

  if and only if $\exists y^* \in \mathbb{R}^m_+$ such that (x^*, y^*) is a KKT point.

Example: Linear Programming

Consider the LP:

$$\begin{aligned} &\max c^T x \\ &\text{s.t. } Ax \leq b \end{aligned}$$

Equivalent convex formulation:

$$\begin{cases} f(x) = -c^T x & \text{(convex)} \\ g(x) = Ax - b \leq 0 & \text{(convex constraints)} \end{cases}$$

KKT conditions:

1. Stationarity: $\nabla f(x) + \nabla g(x)^T y = 0$
   $\Rightarrow -c + A^T y = 0$
   $\Rightarrow A^T y = c$
2. Primal feasibility: $g(x) \leq 0$
   $\Rightarrow Ax \leq b$
3. Dual feasibility: $y \geq 0$
4. Complementary slackness: $y^T(Ax - b) = 0$

Key Observations

1. Affine functions are convex
2. For LPs, KKT conditions reduce to:

$$\begin{cases} A^T y = c \\ Ax \leq b \\ y \geq 0 \\ y^T(Ax - b) = 0 \end{cases}$$

3. Complementary slackness shows which constraints are active
4. LP duality is a special case of KKT!

11 Riemann Integration

11.1 Introduction to Integration

1. Volume between graph(f) and x_1/x_2-plane?
2. Volume over area in x_1/x_2-plane?
3. Surface isolated between curve in x_1/x_2-plane and its image under f?
4. Integrals along a curve?
   (generalization of 1D integrals)
   e.g. energy spent for a hike in the mountains
   e.g. length of a curve?
5. inversion of of differentiation? antiderivative (for a given function $f : \mathbb{R}^n \to \mathbb{R}^n$ we can ask: "is there a function $F : \mathbb{R}^n \to \mathbb{R}$ such that grad$(F) = f$?")

$$f : \mathbb{R}^n \to \mathbb{R}$$
$$\Rightarrow D_f(x) \in \mathbb{R}^n$$
$$\Rightarrow \text{grad}(f) : \mathbb{R}^n \to \mathbb{R}^n$$
$$x \mapsto D_f(x)$$

11.2 Riemann Integrals over Boxes

Recall definition of integral on $\mathbb{R}^n$:

- Found area from above and from below except subtypes
- Decorate initial body
- If both bounds converge to the same value
- Integral of f

Instead of intervals, we will use boxes in $\mathbb{R}^n$

Definition

(1) A set $B \subseteq \mathbb{R}^n$ is called a *box* if there exist vectors

$$a = [\alpha_1, \alpha_2, \dots, \alpha_n]^T \in \mathbb{R}^n$$

and

$$b = [\beta_1, \beta_2, \dots, \beta_n]^T \in \mathbb{R}^n$$

with $a \neq b$ (i.e., $\alpha_i \neq \beta_i$ for all i), such that

$$B = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times \dots \times [\alpha_n, \beta_n]$$

Note that: $B = B(a, b)$.

(2) For a box $B(a, b)$, the volume is

$$\text{vol}(B(a, b)) = \prod_{i=1}^{n} (\beta_i - \alpha_i)$$

and the diameter by

$$\text{diam}(B(a, b)) = \|b - a\|$$

(3) Let $B \subseteq \mathbb{R}^n$ be a box. A *(rectangular) partition* P of B is a finite set of boxes $P = \{B_1, \dots, B_m\}$ such that:

- $\bigcup_{k=1}^{m} B_k = B$
- $\text{int}(B_k) \cap \text{int}(B_j) = \emptyset$ for all $j \neq k$

(4) The *diameter* of a partition $P = \{B_1, \dots, B_m\}$ is defined as

$$\text{diam}(P) = \max_k \text{diam}(B_k)$$

For a box $B \subseteq \mathbb{R}^n$, the set of all partitions of B is denoted by $\mathcal{P}(B)$.

Definition

Let $B \subseteq \mathbb{R}^n$ be a box and $f : B \to \mathbb{R}$ a function.

(1) For $P \in \mathcal{P}(B)$, the *upper* and *lower Riemann sums* are defined as

$$\sum^{+}(f, P) := \sum_{B_k \in P} \sup_{x \in B_k} f(x) \cdot \mathrm{vol}(B_k)$$

and

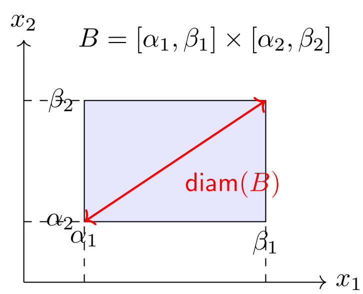
$$\sum^{-}(f, P) := \sum_{B_k \in P} \inf_{x \in B_k} f(x) \cdot \mathrm{vol}(B_k)$$

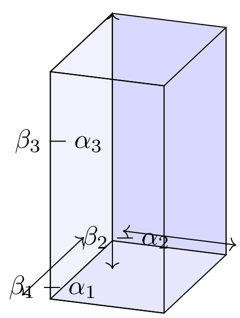
if these values exist.

(2) The function f is called *Riemann-integrable* on B if

$$\sup_{P \in \mathcal{P}(B)} \sum^{-}(f, P) = \inf_{P \in \mathcal{P}(B)} \sum^{+}(f, P)$$

and that value is called *integral* of f on B, denoted by $\int_B f(x)dx$.





Theorem

Let $B \subseteq \mathbb{R}^n$ be a box, $\alpha, \beta \in \mathbb{R}$ and $f, g : B \to \mathbb{R}$ Riemann-integrable.

1. Linearity: $\alpha f + \beta g$ is Riemann-integrable with

$$\int_B (\alpha f + \beta g)(x)dx = \alpha \cdot \int_B f(x)dx + \beta \cdot \int_B g(x)dx$$

2. Monotonicity: if $f(x) \leq g(x)$ for all $x \in B$, then

$$\int_B f(x)dx \leq \int_B g(x)dx$$

3. The function $x \mapsto |f(x)|$ is Riemann-integrable with

$$\left\| \int_B f(x)dx \right\| \leq \int_B |f(x)|dx$$

4. Every continuous function $h : B \to \mathbb{R}$ is Riemann-integrable

12 Fubini's Theorem (Satz von Fubini)

Definition

Let $B = B(a,b) \subseteq \mathbb{R}^n$ be a box with $a = (\alpha_1, \ldots, \alpha_n)^T \in \mathbb{R}^n$ and $b = (\beta_1, \ldots, \beta_n)^T \in \mathbb{R}^n$ and let $f : B \to \mathbb{R}$ be a function on B. For $i \in \{1, \cdots, n\}$ let a_{-i}, b_{-i} denote the vectors

$$a_{-i} := (\alpha_1, \alpha_2, \cdots, \alpha_n)^T \in \mathbb{R}^{n-1}$$
$$b_{-i} := (\beta_1, \beta_2, \cdots, \beta_n)^T \in \mathbb{R}^{n-1}$$

Then the function

$$F_i : B(a_{-i}, b_{-i}) \to \mathbb{R}$$

defined by

$$F_{i+1} := \int_{x_i=\alpha_i}^{\beta_i} f(x_1, \ldots, x_n)\, dx_i$$

if x_i satisfies

$$x_i \in A$$

is called the **i-th parameter** integral of F if it exists.

Theorem

Let $B(a,b) \subseteq \mathbb{R}^n$ be a box and $f : B(a,b) \to \mathbb{R}$ be continuous on $B(a,b)$.

(1) The parameter integrals F_i are continuous on $B(a_{-i}, b_{-i})$ for $i \in \{1, \cdots, n\}$.

(2) Let $\pi : \{1, \cdots, n\} \to \{1, \cdots, n\}$ be bijective ("punctuations"), then

$$\int_{B(a,b)} f(x)dx = \int_{\alpha(\pi(1))}^{\beta(\pi(1))} \left(\int_{\alpha(\pi(2))}^{\beta(\pi(2))} \cdots \left(\int_{\alpha(\pi(n))}^{\beta(\pi(n))} f(x_1, x_2, \ldots, x_n) dx_{\pi(n)} \right) \cdots dx_{\pi(2)} \right) dx_{\pi(1)}$$

for $x_1, x_2, \ldots, x_n \in B(a,b)$.

For continuous function f on a box B the integral $\int_B f(x)dx$ can be computed through successive parametric integrals in arbitrary order.

12.1 Integration over Arbitary Sets

Let $O \subseteq \mathbb{R}^n$, $f : O \to \mathbb{R}$ and $B \subseteq \mathbb{R}^n$ with $O \subseteq B$ (i.e., O is bounded). Set

$$f_O : B \to \mathbb{R}$$

$$f_O(X) := \begin{cases} f(X), & x \in O; \\ 0, & x \in B \setminus O. \end{cases}$$

The function f is *Riemann-integrable* over O if f_O is *Riemann-integrable* over B, and we set

$$\int_O f(x)\, dx = \int_B f_O(x)\, dx.$$

A set $D \subseteq \mathbb{R}^n$ is called *Jordan-measurable* if the constant function $x \mapsto 1$ is integrable over D; in that case, we denote the n-dimensional volume of D by

$$\mathrm{vol}(D) := \int_D 1\, dx.$$

Example

Consider the set of rational points in $[0,1]^2$:

$$Q := \{x \in [0,1]^2 : x \in \mathbb{Q}^2\}.$$

Volume of Q?

- Every box of positive diameter contains a rational point (because rationals are dense in the reals) and also an irrational point.
- Consider the characteristic function $\chi_Q : [0,1]^2 \to \mathbb{R}$,

$$\chi_Q(x) = \begin{cases} 1, & x \in Q; \\ 0, & x \notin Q. \end{cases}$$

- For every partition, the supremum of χ_Q over any sub-box is 1, and the infimum is 0.
- Thus, the upper Riemann sum is 1, and the lower Riemann sum is 0.
- χ_Q is not Riemann-integrable over $[0,1]^2$.
- Q is not Jordan-measurable.

12.2 Normal Domains

Consider sets with a "nice" description for integration. A set $O \subseteq \mathbb{R}^n$ is called a *normal domain* if there exist two numbers $\alpha \leq \beta$ and for each $i \in \{1,\dots,n-1\}$ continuous functions $\alpha_i, \beta_i : \mathbb{R}^i \to \mathbb{R}$, such that

$$O = \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : \alpha \leq x_1 \leq \beta \wedge \alpha_1(x_1) \leq x_2 \leq \beta_1(x_1) \wedge \dots \wedge \alpha_{n-1}(x_1,\dots,x_{n-1}) \leq x_n \leq \beta_{n-1}(x_1,\dots,x_{n-1}) \right\}.$$

Examples

1. $$D = \left\{ x = (x_1,x_2) \in \mathbb{R}^2 : -\frac{\sqrt{3}}{2} \leq x_1 \leq \frac{\sqrt{3}}{2} \wedge x_1^2 \leq x_2 \leq 7 - x_1^2 \right\}.$$
   This is a region bounded by two parabolas.
2. $$D = \{x \in \mathbb{R}^2 : \|x\|^2 \leq 7\}$$
   This is a disc centered at the origin with radius $\sqrt{7}$.
3. $$D = \{x \in \mathbb{R}^3 : \|x\|^2 \leq 7\}$$
   This is a ball centered at the origin with radius $\sqrt{7}$.

12.3 Theorem

Let $O \subseteq \mathbb{R}^n$ be a normal domain described by continuous functions $\alpha, \beta \in \mathbb{R}$ and $\alpha_i, \beta_i : \mathbb{R}^i \to \mathbb{R}$ for $i \in \{1,\dots,n-1\}$. Then, a continuous function $f : O \to \mathbb{R}$ is Riemann-integrable over O with

$$\int_O f(x)\,dx = \int_\alpha^\beta \int_{\alpha_1(x_1)}^{\beta_1(x_1)} \cdots \int_{\alpha_{n-1}(x_1,\dots,x_{n-1})}^{\beta_{n-1}(x_1,\dots,x_{n-1})} f(x_1,\dots,x_n)\,dx_n \cdots dx_1.$$

12.4 Change of Variables

Let $X \subseteq \mathbb{R}^n$ be open and $g : X \to \mathbb{R}^n$ an injective, continuously differentiable function with

$$\det(J_g(x)) \neq 0 \quad \text{for all } x \in X,$$

where $J_g(x)$ is the Jacobian matrix of g at x.

(1) If $D \subseteq X$ is closed, bounded, and Jordan-measurable, then $g(D)$ is also closed, bounded, and Jordan-measurable.

(2) Let $D \subseteq X$ be closed, bounded, and Jordan-measurable, and let $f : g(D) \to \mathbb{R}$ be continuous. Then f is Riemann-integrable over $g(D)$, and

$$\int_{g(D)} f(y)\,dy = \int_D f(g(x)) \cdot |\det(J_g(x))|\,dx.$$

Examples of Transformations (Polar-, Kugel-, Zylinderkoordinaten)

12.4.1 Half-Ring in Polar Coordinates

$$X = \{x \in \mathbb{R}^2 : x_1 > 0,\, 1 < \|x\|^2 < 4\},$$
$$f(x_1, x_2) = x_1 \cdot (x_1^2 + x_2^2).$$

This is a half-ring with inner radius 1 and outer radius 2.

In polar coordinates:

$$X = \left\{(x_1, x_2) = (r\cos t, r\sin t) : 1 < r < 2,\, -\frac{\pi}{2} < t < \frac{\pi}{2}\right\}.$$

The variables r and t are independent, making X a "box" in polar coordinates.

12.4.2 Transformation and Jacobian

To evaluate $\int_X x\,dx$ using polar coordinates:

$$g(r, t) = \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix},$$

with domain $D = \{(r, t) : r \in (1, 2),\, t \in (-\frac{\pi}{2}, \frac{\pi}{2})\}$.

The Jacobian matrix of g is:

$$J_g(r, t) = \begin{pmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{pmatrix},$$

and its determinant is $\det(J_g(r, t)) = r$.

12.4.3 Octant of a Ball in Spherical Coordinates

$$X = \{x \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0,\, x_1^2 + x_2^2 + x_3^2 \leq 1\}.$$

This is the first octant of a ball with radius $\sqrt{7}$.

In spherical coordinates:

$$g(r, \rho, \psi) = \begin{pmatrix} r\cos\rho\cos\psi \\ r\sin\rho\cos\psi \\ r\sin\psi \end{pmatrix},$$

with $D = \{(r, \rho, \psi) : r \in (0, \sqrt{1}),\, \rho \in (0, \frac{\pi}{2}),\, \psi \in (0, \frac{\pi}{2})\}$.

The Jacobian determinant is:

$$\det(J_g(r, \rho, \psi)) = r^2\cos\psi.$$

The volume integral becomes:

$$\int_X 1\,dx = \int_{r=0}^{\sqrt{1}} \int_{\rho=0}^{\pi/2} \int_{\psi=0}^{\pi/2} r^2\cos\psi\,d\psi\,d\rho\,dr.$$

12.5 Parametrized Curves

Definition

A *parametrized curve* in $\mathbb{R}^n$ is a continuous, piecewise continuously differentiable mapping $\gamma : I \to \mathbb{R}^n$, where $I \subseteq \mathbb{R}$ is an interval. The *graph* (or trace) of γ is the set:

$$\text{graph}(\gamma) = \{\gamma(t) : t \in I\}.$$

For $t \in I$ where γ is differentiable:

- The derivative $\gamma'(t)$ is called the *tangent vector* of γ at t.
- The norm $\|\gamma'(t)\|$ is called the *speed* of γ at t.

12.6 Examples

1. Linear path:
$$\gamma : [0,7] \to \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} t \\ t \end{pmatrix}.$$
2. Quadratic path:
$$\gamma : [0,7] \to \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} t^2 \\ t^2 \end{pmatrix}.$$
3. Unit circle:
$$\gamma : [0,2\pi] \to \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$
The speed is constant: $\|\gamma'(t)\| = 1$.
4. Decaying spiral:
$$\gamma : [0,\infty) \to \mathbb{R}^2, \quad \gamma(t) = e^{-t} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$
5. Graph of a function: For $f : [\alpha,\beta] \to \mathbb{R} \Rightarrow \gamma_f : [\alpha,\beta] \to \mathbb{R}^2$,
$$\gamma_f(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix}.$$
6. Helix:
$$\gamma : [0,\infty) \to \mathbb{R}^3, \quad \gamma(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}.$$

12.7 Curve Length

For a parametrized curve $\gamma : I \to \mathbb{R}^n$, we assume:

$$\dot\gamma(t) \neq 0 \text{ for all } t \in I.$$

- Let $\gamma : [\alpha,\beta] \to \mathbb{R}^n$ be a curve.
- Partition the interval $[\alpha,\beta]$ with points:
$$\alpha = t_0 < t_1 < \cdots < t_{n-1} < t_n = \beta.$$
- Compute the length of the polygonal path through $\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_n)$:
$$L_n = \sum_{k=1}^{n} \|\gamma(t_k) - \gamma(t_{k-1})\|.$$

- This approximates the true curve length $s(\gamma)$.
- Rewrite using the mean value theorem:

$$L_n = \sum_{k=1}^{n} (t_k - t_{k-1}) \cdot \left\| \frac{\gamma(t_k) - \gamma(t_{k-1})}{t_k - t_{k-1}} \right\|.$$

- As the partition becomes finer ($|t_k - t_{k-1}| \to 0$):

$$s(\gamma) = \int_\alpha^\beta \|\dot{\gamma}(t)\| \, dt.$$

12.7.1 Theorem

Let $\gamma : [\alpha, \beta] \to \mathbb{R}^n$ be a continuously differentiable curve with $\dot{\gamma}(t) \neq 0$ for all $t \in [\alpha, \beta]$. Then the length $s(\gamma)$ of the curve is:

$$s(\gamma) = \int_\alpha^\beta \|\dot{\gamma}(t)\| \, dt.$$

Example: Logarithmic Spiral
Consider the logarithmic spiral:

$$\gamma : [0, \infty) \to \mathbb{R}^2, \quad \gamma(t) = e^{-t} \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$

$$\dot{\gamma}(t) = e^{-t} \begin{pmatrix} -\cos t - \sin t \\ -\sin t + \cos t \end{pmatrix}.$$

$$\begin{aligned}
\|\dot{\gamma}(t)\| &= e^{-t} \sqrt{(-\cos t - \sin t)^2 + (-\sin t + \cos t)^2} \\
&= e^{-t} \sqrt{(\cos^2 t + 2\cos t \sin t + \sin^2 t) + (\cos^2 t - 2\cos t \sin t + \sin^2 t)} \\
&= e^{-t} \sqrt{2(\cos^2 t + \sin^2 t)} \\
&= e^{-t} \sqrt{2}.
\end{aligned}$$

$$s(\gamma) = \int_0^\infty \|\dot{\gamma}(t)\| \, dt = \sqrt{2} \int_0^\infty e^{-t} \, dt = \sqrt{2} \left[-e^{-t} \right]_0^\infty = \sqrt{2}.$$

We want to define functions using polar coordinates in the plane:

- Domain: Angle θ
- Function value: Radius $r = r(\theta)$
- Typically, domain is an interval of $[0, 2\pi)$ (or a translation)
- Recall the polar-to-Cartesian transformation:

$$T(r, \theta) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$$

Definition

Let $I \subseteq [0, 2\pi)$ and let $r : I \to \mathbb{R}_{\geq 0}$ be a continuous function. The curve $\gamma : I \to \mathbb{R}^2$ defined by

$$\gamma(\theta) := T(r(\theta), \theta)$$

is the *polar curve* associated to r, where T is the transformation from polar to Cartesian coordinates.

Examples

(a) A circle with radius $r^* \in \mathbb{R}_{>0}$ is a polar curve:

$$\gamma : [0, 2\pi) \to \mathbb{R}^2, \quad \gamma(\theta) = \begin{pmatrix} r^* \cos\theta \\ r^* \sin\theta \end{pmatrix} = T(r^*, \theta)$$

or equivalently:

$$r(\theta) := r^*$$

(b) Rose petal curve:

$$r(\theta) = 3\sin(2\theta), \quad \theta \in \left[0, \frac{\pi}{2}\right)$$

For a circle:

$$\text{Area} = \frac{r^2}{2}(\theta_2 - \theta_1)$$

This represents the ratio of the sector angle to the full circle:

$$\frac{\theta_2 - \theta_1}{2\pi} \times \pi r^2 = \frac{r^2}{2}(\theta_2 - \theta_1)$$

12.8 General Polar Curve

For a polar curve with radius $r = r(\theta)$:

- Subdivide the angle interval $[\theta_1, \theta_2]$
- Approximate area using circular sectors:

$$\text{Area} \approx \sum_{i=1}^{n-1} \frac{[r(\theta_i)]^2}{2} \cdot (\theta_{i+1} - \theta_i)$$

- Take the limit as the partition becomes finer (Riemann sum):

$$\text{Area} = \int_{\theta_1}^{\theta_2} \frac{[r(\theta)]^2}{2} \, d\theta$$

For a planar curve in polar coordinates, the sector area between angles θ_1 and θ_2 is given by:

$$A = \frac{1}{2} \int_{\theta_1}^{\theta_2} [r(\theta)]^2 d\theta$$

Consider a curve $\gamma : I \to \mathbb{R}^2$ given in Cartesian coordinates:

$$\gamma(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

$$\int_{t_1}^{t_2} f(g(t)) \cdot g'(t) dt = \int_{g(t_1)}^{g(t_2)} f(x)$$

area can be written as

$$\int_{t_1}^{t_2} \frac{1}{2} [r(\rho(t))]^2 \cdot \dot{\rho}(t) dt$$

Using the polar transformation:

$$T(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

We can extract polar coordinates from Cartesian coordinates:

$$r(t)^2 = x_1(t)^2 + x_2(t)^2$$
$$\theta(t) = \arctan\left(\frac{x_2(t)}{x_1(t)}\right)$$

The derivative of $\theta(t)$ is:

$$\begin{aligned}\frac{d}{dt}\theta(t) &= \frac{d}{dt}\left(\arctan\frac{x_2(t)}{x_1(t)}\right)\\ &= \frac{1}{1+\left(\frac{x_2(t)}{x_1(t)}\right)^2}\cdot\frac{\dot{x}_2(t)x_1(t)-x_2(t)\dot{x}_1(t)}{x_1(t)^2}\\ &= \frac{\dot{x}_2(t)x_1(t)-x_2(t)\dot{x}_1(t)}{x_1(t)^2+x_2(t)^2}\end{aligned}$$

For a general parametric curve, the sector area between parameters t_1 and t_2 is:

$$\begin{aligned}A &= \frac{1}{2}\int_{t_1}^{t_2}[r(\theta(t))]^2\dot{\theta}(t)dt\\ &= \frac{1}{2}\int_{t_1}^{t_2}\left(x_1(t)^2+x_2(t)^2\right)\cdot\frac{\dot{x}_2(t)x_1(t)-x_2(t)\dot{x}_1(t)}{x_1(t)^2+x_2(t)^2}dt\\ &= \frac{1}{2}\int_{t_1}^{t_2}\left(\dot{x}_2(t)x_1(t)-x_2(t)\dot{x}_1(t)\right)dt\end{aligned}$$

This represents the sector area of the planar curve $\gamma(t) = \begin{pmatrix} x_1(t) \\ x_2(t)\end{pmatrix}$ between parameters t_1 and t_2.

13 Line Integrals

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ and a curve

$$\gamma : [\alpha, \beta] \to \mathbb{R}^n.$$

We would like to compute the integral of f along γ - called the *line integral*.

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar function and $\gamma : [\alpha, \beta] \to \mathbb{R}^n$ be a curve. The line integral of f along γ is defined as:

$$\int_\gamma f\,ds := \int_\alpha^\beta f(\gamma(t))\cdot\|\dot{\gamma}(t)\|\,dt.$$

- ds represents the arc length element
- If γ is closed (i.e., $\gamma(\alpha) = \gamma(\beta)$), we denote the integral with a circle: $\oint_\gamma f\,ds$

Approximate the curve with a polygonal path through points:

$$\gamma(t_1), \gamma(t_2), \ldots, \gamma(t_n)$$

where $\alpha = t_0 < t_1 < \cdots < t_n = \beta$.

The Riemann sum approximation is:

$$\begin{aligned}&\sum_{k=1}^n f(\gamma(t_k))\cdot\|\gamma(t_k)-\gamma(t_{k-1})\|\\ &= \sum_{k=1}^n f(\gamma(t_k))\cdot\left\|\frac{\gamma(t_k)-\gamma(t_{k-1})}{t_k-t_{k-1}}\right\|\cdot(t_k-t_{k-1})\end{aligned}$$

Taking the limit as the partition becomes finer yields the integral:

$$\int_\alpha^\beta f(\gamma(t))\cdot\|\dot{\gamma}(t)\|\,dt$$

Example

Consider:

$$f : \mathbb{R}^3 \to \mathbb{R}$$
$$f(x_1, x_2, x_3) = x_1^2 + x_2 \cdot x_3$$

and the helical curve:

$$\gamma : [0, 2\pi] \to \mathbb{R}^3, \quad \gamma(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}$$

1. Compute the derivative:

$$\dot{\gamma}(t) = \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix}$$
$$\|\dot{\gamma}(t)\| = \sqrt{(-\sin t)^2 + \cos^2 t + 1^2} = \sqrt{2}$$

2. Compute the line integral:

$$\begin{aligned} \int_\gamma f\, ds &= \int_0^{2\pi} f(\gamma(t)) \cdot \|\dot{\gamma}(t)\|\, dt \\ &= \sqrt{2} \int_0^{2\pi} (\cos^2 t + t \sin t)\, dt \\ &= \sqrt{2} \left(\int_0^{2\pi} \cos^2 t\, dt + \int_0^{2\pi} t \sin t\, dt \right) \end{aligned}$$

3. Evaluate each integral:

- For $\cos^2 t$:

$$\int \cos^2 t\, dt = \frac{1}{2}(\cos t \sin t + t) + C$$
$$\frac{1}{2}(\cos t \sin t + t)\Big|_0^{2\pi} = \pi$$

- For $t \sin t$:

$$\int t \sin t\, dt = -t \cos t + \sin t + C$$
$$(-t \cos t + \sin t)\big|_0^{2\pi} = -2\pi$$

4. Final result:

$$\int_\gamma f\, ds = \sqrt{2}(\pi - 2\pi) = -\pi\sqrt{2}$$

13.1 Vector Fields

Vector fields describe quantities with both magnitude and direction that may vary depending on location in space.

13.1.1 Examples

- Varying winds in the atmosphere
- Ocean currents
- Gravitational field acting on a satellite
- Motion of charged particles in electromagnetic fields

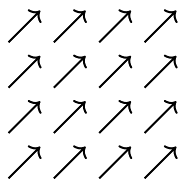
13.1.2 Definitions

A set $O \subseteq \mathbb{R}^n$ is called *connected* if for any two points $x, y \in O$, there exists a continuous curve connecting x to y that lies entirely within O.

Let $O \subseteq \mathbb{R}^n$ be connected. A function $v : O \to \mathbb{R}^n$ is called an *n-dimensional vector field* over O. In contrast, $f : O \to \mathbb{R}$ is called a *scalar field.*

Vector fields can be visualized by plotting vectors $v(X_i)$ anchored at sample points $X_i \in D$ (typically a regular grid).

Examples of Vector Fields

1. **Constant Field**:
$$v : \mathbb{R}^2 \to \mathbb{R}^2, \quad v(x_1, x_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$


2. **Central Field**:
$$V : \mathbb{R}^2 \to \mathbb{R}^2, \quad V\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{\|x\|^2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
where $\|x\| = \sqrt{x_1^2 + x_2^2}$.

3. **Rotational Field**:
$$V : \mathbb{R}^2 \to \mathbb{R}^2, \quad V\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$
Note that $V(x) \perp x$ for all x.

4. **Gradient Field**: Let $F : \mathbb{R}^n \to \mathbb{R}$ be differentiable. The gradient field is:
$$v(x) = \nabla F(x) = \begin{pmatrix} \frac{\partial F}{\partial x_1} \\ \vdots \\ \frac{\partial F}{\partial x_n} \end{pmatrix}$$
Example:
$$F(x_1, x_2) = x_1^2 + x_2^2$$
$$\nabla F(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = 2x$$

13.2 Rotation and Divergence

Definition

Let $v : \mathbb{R}^3 \to \mathbb{R}^3$ be differentiable, then the **rotation (rotor) of v** is the vector field **rot**(v) defined as:

$$\text{rot } v = \begin{pmatrix} \dfrac{\partial v_3}{\partial x_2} - \dfrac{\partial v_2}{\partial x_3} \\ \dfrac{\partial v_1}{\partial x_3} - \dfrac{\partial v_3}{\partial x_1} \\ \dfrac{\partial v_2}{\partial x_1} - \dfrac{\partial v_1}{\partial x_2} \end{pmatrix}$$

memory aid:

$$\text{rot } v = \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \dfrac{\partial}{\partial x_2} \\ \dfrac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

rot v is the measure for local rotation of a vector field as experienced by a partial moving along the field

Definition

Let $v : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable, then the **diversion of v** is the function:

$$\text{div}v(x) = \sum_{i=1}^{n} \frac{\partial v_i}{\partial x_i}(x)$$

memory aid:

$$\text{div } v = \langle \nabla, v \rangle = \langle \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \rangle$$

divergence is a measure of local density of sources/sinks

14 Line Integrals over Vector Fields

A vector field $v : \mathbb{R}^n \to \mathbb{R}^n$ (e.g., gravitational, electric/magnetic field, wind, etc.) is given.

- A curve $\gamma : [a, b] \to \mathbb{R}^n$ describes a motion through v under the influence of the field described by v (e.g., movement of an airplane through atmospheric wind).
- What is the amount of energy necessary for that movement?
- What is the amount of work expended at this point (amount of field v that winds in air)?

14.1 Line Integral

The line integral is the projection of $v(x)$ onto the direction of movement $\dot{\gamma}(t)$ where $\gamma(t) = x$.

- Effective strength of the force at point x that is experienced when moving along curve γ at point x.
- Integrable projection of $v(\gamma(t))$ onto $\dot{\gamma}(t)$ along the curve γ.
- Projection (line integral): $\langle v(\gamma(t)), \frac{\dot{\gamma}(t)}{\|\dot{\gamma}(t)\|} \rangle$ for absolute t (includes sign).

  The line integral can be written as:

$$\int_\gamma f \, ds = \int_a^b \langle v(\gamma(t)), \frac{\dot{\gamma}(t)}{\|\dot{\gamma}(t)\|} \rangle \|\dot{\gamma}(t)\| \, dt = \int_a^b \langle v(\gamma(t)), \dot{\gamma}(t) \rangle \, dt$$

Definition

Let $v : D \to \mathbb{R}^n$ be a vector field on a connected set $D \subseteq \mathbb{R}^n$, and let $\gamma : [a, b] \to D$ be a curve contained in D. The line integral of v along γ, denoted by $\int_\gamma v \, ds$, is defined as:

$$\int_\gamma v \, ds = \int_a^b \langle v(\gamma(t)), \dot{\gamma}(t) \rangle \, dt$$

Example

Let $v : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by:

$$v \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 x_2^2 \\ x_1 x_3 \\ x_3 \end{pmatrix}$$

Let $\gamma : [0, 1] \to \mathbb{R}^3$ be defined by:

$$\gamma(t) = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}$$

The line integral of v along γ is:

$$\int_\gamma v\,ds = \int_0^1 \langle v(\gamma(t)), \dot{\gamma}(t)\rangle\,dt$$

First, compute $v(\gamma(t))$:

$$v(\gamma(t)) = v\begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix} = \begin{pmatrix} t\cdot(t^2)^2 \\ t\cdot t^3 \\ t^3 \end{pmatrix} = \begin{pmatrix} t^5 \\ t^4 \\ t^3 \end{pmatrix}$$

Next, compute $\dot{\gamma}(t)$:

$$\dot{\gamma}(t) = \begin{pmatrix} 1 \\ 2t \\ 3t^2 \end{pmatrix}$$

Now, compute the dot product:

$$\langle v(\gamma(t)), \dot{\gamma}(t)\rangle = t^5\cdot 1 + t^4\cdot 2t + t^3\cdot 3t^2 = t^5 + 2t^5 + 3t^5 = 6t^5$$

Finally, integrate:

$$\int_0^1 6t^5\,dt = \left[t^6\right]_0^1 = 1$$

14.2 Gradient Fields

For a scalar field $F : D \to \mathbb{R}$, the gradient defines a vector field:

$$\text{grad } F : D \to \mathbb{R}^n, \quad \text{grad } F(x) = \nabla F(x).$$

One could say that F is an antiderivative of the vector field grad F. Vector fields that can be represented as the gradient of some scalar field have a special name.

Definition

A vector field $v : D \to \mathbb{R}^n$ on a connected set $D \subseteq \mathbb{R}^n$ is called a **gradient field**, **potential field**, or **conservative** if there exists a scalar field $F : D \to \mathbb{R}$ such that:

$$v = \text{grad } F.$$

Such a function F is called a **potential** or an **antiderivative** of v.

Example 1

Consider the vector field:

$$v\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 e^{x_1x_2+x_3} \\ x_1 e^{x_1x_2} \\ x_1 + 2x_3 \end{pmatrix}.$$

Is v a gradient field?

Assume there exists a potential $F : \mathbb{R}^3 \to \mathbb{R}$ such that $v = \text{grad } F$. Then:

$$\frac{\partial F}{\partial x_1}(x) = x_2 e^{x_1x_2+x_3}, \quad \frac{\partial F}{\partial x_2}(x) = x_1 e^{x_1x_2}, \quad \frac{\partial F}{\partial x_3}(x) = x_1 + 2x_3.$$

Integrate the first equation with respect to x_1:

$$F(x) = \int x_2 e^{x_1x_2+x_3}\,dx_1 = e^{x_1x_2+x_3} + \Phi_1(x_2, x_3).$$

But from the second equation:

$$\frac{\partial F}{\partial x_2}(x) = x_1 e^{x_1x_2} = x_1 e^{x_1x_2} + \frac{\partial}{\partial x_2}\Phi_1(x_2, x_3).$$

This implies:

$$\frac{\partial}{\partial x_2}\Phi_1(x_2, x_3) = 0 \Rightarrow \Phi_1 \text{ does not depend on } x_2.$$

Thus, $\Phi_1(x_2, x_3) = \Phi_2(x_3)$, and:

$$F(x) = e^{x_1x_2+x_3} + \Phi_2(x_3).$$

Now, from the third equation:

$$\frac{\partial F}{\partial x_3}(x) = e^{x_1x_2+x_3} + \frac{d}{dx_3}\Phi_2(x_3) = x_1 + 2x_3.$$

This requires:

$$\frac{d}{dx_3}\Phi_2(x_3) = x_1 + 2x_3 - e^{x_1x_2+x_3},$$

which is impossible unless $x_1x_2 + x_3$ is constant. Therefore, v is **not** a gradient field.

Example 2

Consider the vector field:

$$v\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2e^{x_1} + x_2x_3 \\ e^{x_1} + x_1x_3 \\ e^{x_1} + x_1x_2 \end{pmatrix}.$$

Assume there exists a potential $F : \mathbb{R}^3 \to \mathbb{R}$ for v. Then:

$$\frac{\partial F}{\partial x_1}(x) = x_2e^{x_1} + x_2x_3 \Rightarrow F(x) = x_2e^{x_1} + x_1x_2x_3 + \Phi_1(x_2, x_3).$$

From the second equation:

$$\frac{\partial F}{\partial x_2}(x) = e^{x_1} + x_1x_3 + \frac{\partial}{\partial x_2}\Phi_1(x_2, x_3) = e^{x_1} + x_1x_3.$$

This implies:

$$\frac{\partial}{\partial x_2}\Phi_1(x_2, x_3) = 0 \Rightarrow \Phi_1 \text{ does not depend on } x_2.$$

Thus, $\Phi_1(x_2, x_3) = \Phi_2(x_3)$, and:

$$F(x) = x_2e^{x_1} + x_1x_2x_3 + \Phi_2(x_3).$$

Now, from the third equation:

$$\frac{\partial F}{\partial x_3}(x) = x_1x_2 + \frac{d}{dx_3}\Phi_2(x_3) = e^{x_1} + x_1x_2.$$

This requires:

$$\frac{d}{dx_3}\Phi_2(x_3) = e^{x_1},$$

which is impossible because Φ_2 depends only on x_3, not on x_1. Therefore, v is **not** conservative.

14.3 Line Integrals in Gradient Fields

Consider a gradient field $v : \mathbb{R}^n \to \mathbb{R}^n$ with potential $F : \mathbb{R}^n \to \mathbb{R}$ and a curve $\gamma : [\alpha, \beta] \to \mathbb{R}^n$ with $a = \gamma(\alpha)$ and $b = \gamma(\beta)$. Then:

$$\text{grad } F(x) = v(x) \quad \forall x \in \mathbb{R}^n.$$

Define the function $f(t) := F(\gamma(t))$, the restriction of F to the curve γ. Then:

$$f'(t) = \nabla F(\gamma(t)) \cdot \dot\gamma(t) = \langle v(\gamma(t)), \dot\gamma(t)\rangle.$$

Thus, the line integral of v along γ is:

$$\int_\gamma v\,ds = \int_\alpha^\beta \langle v(\gamma(t)), \dot\gamma(t)\rangle\,dt = \int_\alpha^\beta f'(t)\,dt = f(\beta) - f(\alpha) = F(b) - F(a).$$

This shows that the line integral depends only on the potential F and its values at the endpoints of γ.

14.4 Theorem

Let $D \subseteq \mathbb{R}^n$ be open, $v : D \to \mathbb{R}^n$ a continuous vector field on D and $\gamma : (\alpha, \beta) \to D$ a continously differentiable curve. If v is a gradient field, then the value of the line integral $\int_{gamma} vds$ only depends only on the start- and endpoint of γ:

$$\int_\gamma vds = F(\gamma(\beta)) - F(\gamma(\alpha))$$

where $F : D \to \mathbb{R}$ is a potential of v
In particular, if γ is a closed curve, i.e. $\gamma(\beta) = \gamma(\alpha)$, then

$$\int_\gamma vds = 0$$

15 Integrability Conditions (Integrablitätsbedingungen)

Let $F : D \to \mathbb{R}$ be a potential for a differentiable vector field $v : D \to \mathbb{R}^n$.
If $F \in C^2$, then by Schwarz's Theorem:

$$\frac{\partial^2 F}{\partial x_i \partial x_j}(x) = \frac{\partial^2 F}{\partial x_j \partial x_i}(x) \quad \text{for all } i \neq j$$

This implies the integrability conditions for v:

$$\frac{\partial v_j}{\partial x_i}(x) = \frac{\partial v_i}{\partial x_j}(x) \quad \text{for all } i \neq j$$

Theorem

Let $v : D \to \mathbb{R}^n$ be a continuously differentiable vector field. If v is a gradient field, then:

$$\frac{\partial v_j}{\partial x_i}(x) = \frac{\partial v_i}{\partial x_j}(x) \quad \text{for all } i, j \in \{1, \ldots, n\} \text{ with } i \neq j.$$

Simply Connected Domains

A connected set $D \subseteq \mathbb{R}^n$ is called **simply connected** if every closed curve in D can be continuously contracted to a single point in D. Every convex set is simply connected.

Example 1

Consider $v : \mathbb{R}^2 \to \mathbb{R}^2$:

$$v(x_1, x_2) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

Check integrability conditions:

$$\frac{\partial v_1}{\partial x_2} = -1, \quad \frac{\partial v_2}{\partial x_1} = +1$$

Since $-1 \neq +1$, v is **not** a gradient field on $\mathbb{R}^2$.

Example 2

Consider $v : \mathbb{R}^2 \to \mathbb{R}^2$:

$$v(x_1, x_2) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Integrability conditions:

$$\frac{\partial v_1}{\partial x_2} = 0 = \frac{\partial v_2}{\partial x_1}$$

Since $\mathbb{R}^2$ is simply connected, v **is** a gradient field.

Example 3

Consider $v : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2$:

$$v(x_1, x_2) = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

Integrability conditions:

$$\frac{\partial v_1}{\partial x_2} = \frac{-x_1^2 + x_2^2}{(x_1^2 + x_2^2)^2} = \frac{\partial v_2}{\partial x_1}$$

The domain $\mathbb{R}^2 \setminus \{0\}$ is **not** simply connected, so we cannot conclude that v is a gradient field. However, if we restrict to a simply connected subdomain (not containing the origin), then v has a potential.

Example 4

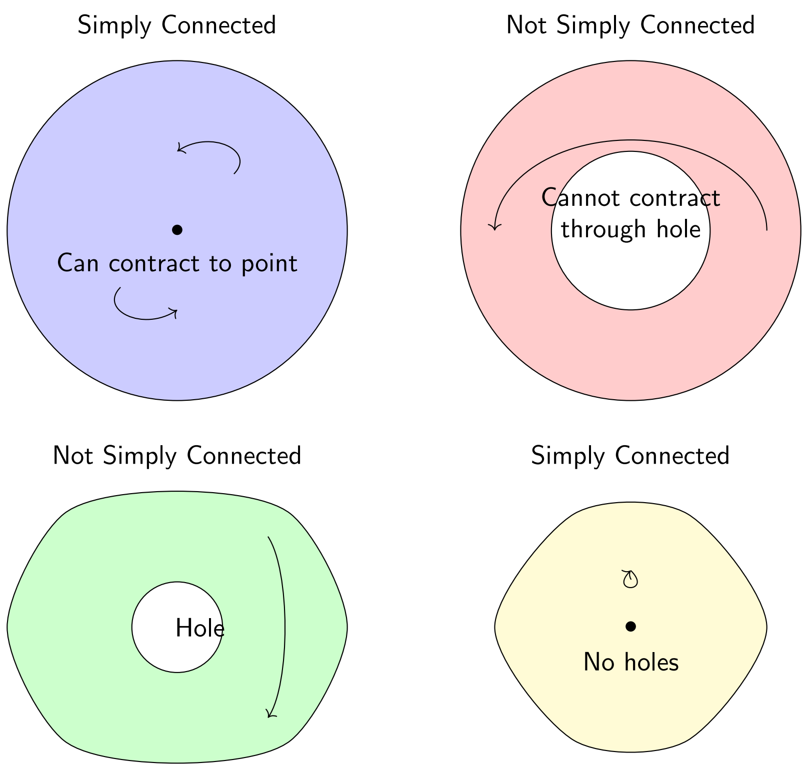
Consider $v : \mathbb{R}^3 \to \mathbb{R}^3$:

$$v(x_1, x_2, x_3) = \begin{pmatrix} e^{x_1} x_2 \\ e^{x_1} x_1 \\ x_3 \end{pmatrix}$$

Integrability conditions:

$$\frac{\partial v_1}{\partial x_2} = e^{x_1} = \frac{\partial v_2}{\partial x_1}, \quad \frac{\partial v_1}{\partial x_3} = 0 = \frac{\partial v_3}{\partial x_1}, \quad \frac{\partial v_2}{\partial x_3} = 0 = \frac{\partial v_3}{\partial x_2}$$

Since $\mathbb{R}^3$ is simply connected, v **is** a gradient field.



16 Green's Theorem

16.1 Theorem

Let $D \subseteq \mathbb{R}^2$ be open, and $v : D \to \mathbb{R}^2$ a continuously differentiable vector field. Furthermore, let $B \subseteq D \subseteq \mathbb{R}^2$ be a compact set (closed + bounded) with $\text{int}(B) \neq \emptyset$, and let $\text{bd}(B)$ consist of finitely many closed, piecewise smooth curves such that B is to the left of these curves with respect to the direction of traversal of the respective curves. Then:

$$\int_{\text{bd}(B)} v\,ds = \int_B \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) d(x_1, x_2).$$

16.2 Example*

Consider the vector field:

$$v\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix},$$

and the set:

$$B = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 7,\ x_2 \geq 0 \right\}.$$

The boundary $\text{bd}(B)$ consists of two parts: γ_1 and γ_2, where:

$$\gamma_1 : [-\sqrt{7}, \sqrt{7}] \to \mathbb{R}^2,\ \gamma_1(t) = \begin{pmatrix} t \\ 0 \end{pmatrix},$$

$$\gamma_2 : [0, \pi] \to \mathbb{R}^2,\ \gamma_2(t) = \begin{pmatrix} \sqrt{7}\cos t \\ \sqrt{7}\sin t \end{pmatrix}.$$

16.3 Application of Green's Theorem

For the given vector field:

$$\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} = 1 - (-1) = 2.$$

Thus:

$$\int_B \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) d(x_1, x_2) = \int_B 2\,d(x_1, x_2) = 2 \cdot \text{vol}(B).$$

On the other hand, the line integral over $\text{bd}(B)$ is:

$$\int_{\text{bd}(B)} v\,ds = \int_{\gamma_1} v\,ds + \int_{\gamma_2} v\,ds.$$

Calculating each part:

$$\int_{\gamma_1} v\,ds = \int_{-1}^{1} \langle v(\gamma_1(t)), \dot{\gamma}_1(t)\rangle\,dt = \int_{-1}^{1} \langle(-0, t), (1, 0)\rangle\,dt = 0,$$

$$\int_{\gamma_2} v\,ds = \int_0^{\pi} \langle v(\gamma_2(t)), \dot{\gamma}_2(t)\rangle\,dt = \int_0^{\pi} \langle(-\sin t, \cos t), (-\sin t, \cos t)\rangle\,dt = \int_0^{\pi} (\sin^2 t + \cos^2 t)\,dt = \pi.$$

Therefore:

$$2 \cdot \text{vol}(B) = \pi \implies \text{vol}(B) = \frac{\pi}{2}.$$

Problem

Consider the points $a = (1, 0)$ and $b = (-1, 0)$ in $\mathbb{R}^2$.

1. Determine a parameterization for the curve γ_1 with endpoints a and b (starting at a) that corresponds to the straight line connecting these two points.

   Solution: The line connecting a and b can be parameterized as:

$$\gamma_1(t) = t \cdot b + (1 - t) \cdot a = \begin{pmatrix} -2t + 1 \\ 0 \end{pmatrix}, \quad t \in [0, 1].$$

2. Determine a parameterization for the curve γ_2 with endpoints a and b (starting at a) that corresponds to a line along the unit circle with $x_2 \geq 0$ connecting these two points.

   Solution: The unit circle can be parameterized as:

$$\gamma_2(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, \pi].$$

3. Consider the vector field $v : \mathbb{R}^2 \to \mathbb{R}^2$ given by:

$$v(x_1, x_2) = \begin{pmatrix} x_1 \\ x_1^2 + x_2^2 \end{pmatrix}.$$

   Compute the line integrals of v along the curves γ_1 and γ_2.

   Solution: First, compute the derivatives of the parameterizations:

$$\gamma_1'(t) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \gamma_2'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}.$$

   The line integrals are:

$$\int_{\gamma_1} v(x)\, dx = \int_0^1 v(\gamma_1(t))^\top \gamma_1'(t)\, dt = \int_0^1 (-2)(-2t + 1)\, dt = 0,$$

$$\int_{\gamma_2} v(x)\, dx = \int_0^\pi (-\sin(t)\cos(t) + \cos(t))\, dt = 0.$$

4. Is v a conservative field?

   Solution: No, because the Jacobian matrix of v:

$$J_v(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ 2x_1 & 2x_2 \end{pmatrix},$$

   is not symmetric. Thus, v is not conservative

5. Use Green's theorem to compute the integral of the vector field:

$$v(x_1, x_2) = \begin{pmatrix} x_1 x_2 \\ x_1^2 x_2^3 \end{pmatrix},$$

   along the curve γ defined as the boundary of the triangle B with vertices $(0, 0)^\top$, $(1, 0)^\top$, and $(1, 2)^\top$ (in counter-clockwise direction).
   By Green's theorem:

$$\int_\gamma v(x)\, dx = \int_B \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) d(x_1, x_2),$$

   where:

$$\frac{\partial v_2}{\partial x_1} = 2x_1 x_2^3, \quad \frac{\partial v_1}{\partial x_2} = x_1.$$

   The integral becomes:

$$\int_B (2x_1 x_2^3 - x_1)\, d(x_1, x_2) = \int_0^1 \int_0^{2x_1} (2x_1 x_2^3 - x_1)\, dx_2\, dx_1 = \frac{2}{3}.$$

6. Consider the conservative vector field $v : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$v(x_1, x_2) = \begin{pmatrix} 2x_1x_2 \cdot e^{x_1^2+x_2} \\ e^{x_2} \cdot \left(1 + e^{x_1^2} + x_2 e^{x_1^2}\right) \end{pmatrix}.$$

a) Determine an antiderivative $F : \mathbb{R}^2 \to \mathbb{R}$ of v, i.e., a function with $\nabla F(x) = v(x)$.

b) Let γ be a curve in $\mathbb{R}^2$ that starts at $(1, 0)$ and ends at $(0, 1)$. Determine the line integral $\int_\gamma v(x)\,\mathrm{d}x$.

Solution

a) For an antiderivative F of v, the following relations need to hold:

$$\frac{\partial F}{\partial x_1}(x_1, x_2) = v_1(x_1, x_2) = 2x_1x_2 \cdot e^{x_1^2+x_2}$$

$$\frac{\partial F}{\partial x_2}(x_1, x_2) = v_2(x_1, x_2) = e^{x_2} \cdot \left(1 + e^{x_1^2} + x_2 e^{x_1^2}\right)$$

We can thus use successive integration to determine F. First, we integrate $v_1(x_1, x_2)$ with respect to x_1 and obtain

$$F(x_1, x_2) = \int v_1(x_1, x_2)\,\mathrm{d}x_1 = \int 2x_1x_2 \cdot e^{x_1^2+x_2}\,\mathrm{d}x_1 = x_2 \cdot e^{x_1^2+x_2} + G(x_2),$$

where G is a function of x_2 alone. Differentiating this result with respect to x_2 yields

$$\frac{\partial F(x_1, x_2)}{\partial x_2} = e^{x_1^2+x_2} + x_2 \cdot e^{x_1^2+x_2} + G'(x_2) = e^{x_2} \cdot \left(e^{x_1^2} + x_2 e^{x_1^2}\right) + G'(x_2).$$

Comparison with v_2 now results in $G'(x_2) = e^{x_2}$, so we get

$$G(x_2) = \int G'(x_2)\,\mathrm{d}x_2 = \int e^{x_2}\,\mathrm{d}x_2 = e^{x_2} + C$$

for a constant C and thus

$$F(x_1, x_2) = x_2 \cdot e^{x_1^2+x_2} + e^{x_2}$$

as one possible antiderivative of v (we choose to set $C = 0$ here because the question only asks for one possible antiderivative of f).

b) As v is conservative with antiderivative F, the value of a line integral over v does not depend on the curve that connects the endpoints. Hence, we can easily compute the integral using

$$\int_\gamma v(x)\,\mathrm{d}x = F(0, 1) - F(1, 0) = e^1 + e^1 - 0 - e^0 = 2e - 1.$$

7. An ellipse $E(a; b) = \{(x_1, x_2) \in \mathbb{R}^2 : \left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 \leq 1\}$ with semiaxes a and b (where $a, b > 0$) can be considered as the image of two parameters, radius r and angle ϕ, under the transformation $T : [0; 1] \times [0; 2\pi) \to \mathbb{R}^2$ (elliptic polar coordinates) defined by

$$T(r, \phi) = \begin{pmatrix} a \cdot r\cos(\phi) \\ b \cdot r\sin(\phi) \end{pmatrix}.$$

a) Compute the determinant of the Jacobian of T.

b) Use the elliptic polar coordinates defined by T above to compute the integral

$$\int_{E(2,1)} x_1 \sqrt{\left(\frac{x_1}{2}\right)^2 + x_2^2}\,\mathrm{d}(x_1, x_2).$$

Solution

a) We start with the Jacobian of T:

$$J_T(r,\phi) = \begin{pmatrix} a\cdot\cos(\phi) & -a\cdot r\sin(\phi) \\ b\cdot\sin(\phi) & b\cdot r\cos(\phi) \end{pmatrix}.$$

Consequently, the determinant of the Jacobian of T is

$$\det(J_T(r,\phi)) = abr\cdot\cos^2(\phi) + abr\sin^2(\phi) = abr.$$

b) For this part, we integrate over the elliptical area $E(2,1)$, which means we use $a=2$ and $b=1$ for the parameter values. Using the transformation rule, we can compute the integral as follows:

$$\begin{aligned}
&\int_{E(2,1)} x_1\sqrt{\left(\frac{x_1}{2}\right)^2 + x_2^2}\,\mathrm{d}(x_1,x_2) \\
&= \int_{E(2,1)} 2\cdot r\cdot\cos(\phi)\cdot\sqrt{\left(\frac{2\cdot r\cdot\cos(\phi)}{2}\right)^2 + (1\cdot r\cdot\sin(\phi))^2}\cdot 2\cdot 1\cdot r\,\mathrm{d}(r,\phi) \\
&= \int_{\phi=0}^{2\pi}\left(\int_{r=0}^{1} 2r\cdot\cos(\phi)\cdot\sqrt{r^2\cos^2(\phi)+r^2\sin^2(\phi)}\cdot 2r\,\mathrm{d}r\right)\mathrm{d}\phi \\
&= \int_{\phi=0}^{2\pi}\left(\int_{r=0}^{1} 2r\cdot\cos(\phi)\cdot|r|\cdot 2r\,\mathrm{d}r\right)\mathrm{d}\phi = \int_{\phi=0}^{2\pi}\left(\int_{r=0}^{1} 4r^3\cdot\cos(\phi)\,\mathrm{d}r\right)\mathrm{d}\phi \\
&= \int_{\phi=0}^{2\pi}\left[r^4\cdot\cos(\phi)\right]_{r=0}^{1}\,\mathrm{d}\phi = \int_{\phi=0}^{2\pi}\cos(\phi)\,\mathrm{d}\phi = [\sin(\phi)]_{\phi=0}^{2\pi} = 0.
\end{aligned}$$