32AH Notes

Brendan Connelly

December 2023

Linear Algebra

Vector Space Axioms

i Additive Associativity: $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

ii Additive Identity: $\boldsymbol{v} + \mathbf{0} = \mathbf{0} + \boldsymbol{v} = \boldsymbol{v}$

iii Additive Inverse: *For all $\boldsymbol{v} \in V$ there exists a $\boldsymbol{w} \in V$ such that $\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}$*

iv Additive Commutativity: $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

v Scalar Associativity: $\lambda(\alpha\boldsymbol{v}) = (\lambda\alpha)\boldsymbol{v}$

vi Scalar Identity: $1\boldsymbol{v} = \boldsymbol{v}$

vii Distribution of Scalar Addition: $(\lambda + \alpha)\boldsymbol{u} = \lambda\boldsymbol{u} + \alpha\boldsymbol{u}$

viii Distribution of Vector Addition: $\lambda(\boldsymbol{u} + \boldsymbol{v}) = \lambda\boldsymbol{u} + \lambda\boldsymbol{v}$

Vector Subspace

i Non-empty $\Rightarrow$ contains the zero vector

ii Closed under vector addition $\Rightarrow \boldsymbol{u} + \boldsymbol{v} \in W$

iii Closed under scalar multiplication $\Rightarrow \lambda(\boldsymbol{v}) \in W$

Pointwise addition and scalar multiplication of continuous functions $f : \mathbb{R} \to \mathbb{R}$

i $(f + g)(x) := f(x) + g(x)$

ii $(\lambda f)(x) := \lambda(f(x))$

Basis of a Vector Space

An ordered set of vectors B is a basis of V if

i $\mathcal{B} \subset$ V

ii span$(\mathcal{B})$ = V

iii $\mathcal{B}$ is linearly independent

Linear Independence/Dependence

A set of vectors $A \subset V$ is said to be linearly dependent if for every nonempty finite subset of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subset A$, there exist scalars α_i, not all zero, such that

$$\alpha_i \boldsymbol{v}_i + \ldots + \alpha_k \boldsymbol{v}_k = \mathbf{0}$$

Otherwise, the set of vectors A is linearly independent

Linear Maps

A linear map $T : V \to W$ is defined as follows for all $k \in \mathbb{N}, \alpha_i \in \mathbb{R}$, and all vectors $x_i \in V$

$$T\left(\sum_i^k \alpha_i x_i\right) = \sum_i^k \alpha_i T(x_i)$$

Equivalently, a linear map will satisfy the following:

i $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$

ii $T(\lambda \boldsymbol{u}) = \lambda T(\boldsymbol{u})$

Standard Matrix

Given a basis $\mathcal{B} = \{\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}\}$, the standard matrix A of a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ is given by

$$\big[A\big] = \begin{bmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{bmatrix} \in M_{m \times n}(\mathbb{R})$$

Determinant Formulas

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

Inverse of a matrix $\in M_{2\times 2}(\mathbb{R})$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Dot Product

The dot product of two vectors can be defined in two primary ways:

1. **Algebraic Definition:**
   Given two vectors $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ in an n-dimensional space, their dot product is the sum of the products of their corresponding components:

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

2. **Geometric Definition:**
   The dot product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ can also be defined as the product of their magnitudes and the cosine of the angle θ between them:

$$\boldsymbol{a} \cdot \boldsymbol{b} = \|\boldsymbol{a}\| \|\boldsymbol{b}\| \cos \theta$$

   where $\|\boldsymbol{a}\|$ and $\|\boldsymbol{b}\|$ are the magnitudes of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively.

Cross Product

The cross product of two vectors in three-dimensional space can be defined in three ways:

1. **Determinant Definition:**
   Given two vectors $\boldsymbol{a} = (a_1, a_2, a_3)$ and $\boldsymbol{b} = (b_1, b_2, b_3)$, their cross product can be expressed using the determinant of a matrix:
   $$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$
   where $\hat{i}, \hat{j}, \hat{k}$ are the unit vectors in the direction of the x, y, and z axes, respectively.

2. **Magnitude and Direction Definition:**
   The magnitude of the cross product is given by the product of the magnitudes of the two vectors and the sine of the angle θ between them:
   $$\|\boldsymbol{a} \times \boldsymbol{b}\| = \|\boldsymbol{a}\|\|\boldsymbol{b}\| \sin\theta$$
   The direction of $\boldsymbol{a} \times \boldsymbol{b}$ is perpendicular to the plane formed by $\boldsymbol{a}$ and $\boldsymbol{b}$, following the right-hand rule.

3. **Algebraic Definition:**
   Given two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$, their cross product $\boldsymbol{u} \times \boldsymbol{v}$ is the unique vector in $\mathbb{R}^3$ defined by the property:
   $$(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = \det \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \\ \boldsymbol{w} \end{bmatrix}$$
   for all $\boldsymbol{w} \in \mathbb{R}^3$.

Properties of the Dot Product:

1. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$ (Commutativity).
2. $\lambda(\boldsymbol{u} \cdot \boldsymbol{v}) = (\lambda \boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \lambda(\boldsymbol{v})$ (Compatibility with Scalars).
3. $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$ (Distribution).
4. $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$, equality only when $\boldsymbol{v} = \boldsymbol{0}$ (Positive Definite).
5. Cauchy-Schwarz Inequality: $|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$.
6. Triangle Inequality: $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$.

Properties of the Cross Product:

1. $\boldsymbol{u} \times \boldsymbol{v} = -(\boldsymbol{v} \times \boldsymbol{u})$ (Anti-commutativity).
2. $\boldsymbol{u} \times \boldsymbol{v}$ is orthogonal to both $\boldsymbol{u}$ and $\boldsymbol{v}$.
3. The cross product $\times : \mathbb{R}^3 \to \mathbb{R}^3 \to \mathbb{R}^3$ is bilinear.
4. $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{0}$ if and only if $\boldsymbol{u}$ and $\boldsymbol{v}$ are parallel.

Orthogonal/Orthonormal

A subset of vectors $S = \{v_1, v_2, \ldots, v_k\} \subseteq \mathbb{R}^n$ is said to be orthogonal if
$$v_i \cdot v_j = 0 \quad \text{for all } i \neq j.$$
Furthermore, if $\|v_i\| = 1$ for all $1 \leq i \leq k$, we say that the subset $S = \{v_1, v_2, \ldots, v_k\} \subseteq \mathbb{R}^n$ is orthonormal.

Projection of a $\boldsymbol{u}$ along a $\boldsymbol{v}$

Assume $\boldsymbol{v} \neq 0$. The projection of $\boldsymbol{u}$ along $\boldsymbol{v}$ is the vector
$$\boldsymbol{u}_{\|\boldsymbol{v}} = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{v} \cdot \boldsymbol{v}}\right) \boldsymbol{v} = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{v}\|^2}\right) \boldsymbol{v} = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{v}\|}\right) \hat{e}_{\boldsymbol{v}}$$

where $\hat{e}_{\boldsymbol{v}}$ is the unit vector in the direction of $\boldsymbol{v}$.

This vector is sometimes denoted as $\text{proj}_{\boldsymbol{v}}\boldsymbol{u}$.

The scalar $\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{v}\|}$ is called the scalar component of $\boldsymbol{u}$ along $\boldsymbol{v}$.

Parameterization of a Line

The line L in $\mathbb{R}^n$, passing through the point $P = (x_1, \ldots, x_n)$, in the direction of the vector $\boldsymbol{v} = \langle v_1, \ldots, v_n\rangle$, can be described by the vector-valued function $\boldsymbol{r}(t)\colon \mathbb{R} \to \mathbb{R}^n$ defined by

$$\boldsymbol{r}(t) = \boldsymbol{r}_0 + t\boldsymbol{v}$$

where $\boldsymbol{r}_0$ is the vector $\boldsymbol{r}_0 = \overrightarrow{OP} = \langle x_1, \ldots, x_n\rangle$. We call $\boldsymbol{r}(t)$ the vector parametrization of L.

Parameterization of a Plane in $\mathbb{R}^n$

The plane P through the point $P = (x_1, \ldots, x_n)$ and determined by two non-parallel vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, can be described by the vector function $\boldsymbol{r}(s,t)\colon \mathbb{R}^2 \to \mathbb{R}^n$ defined by

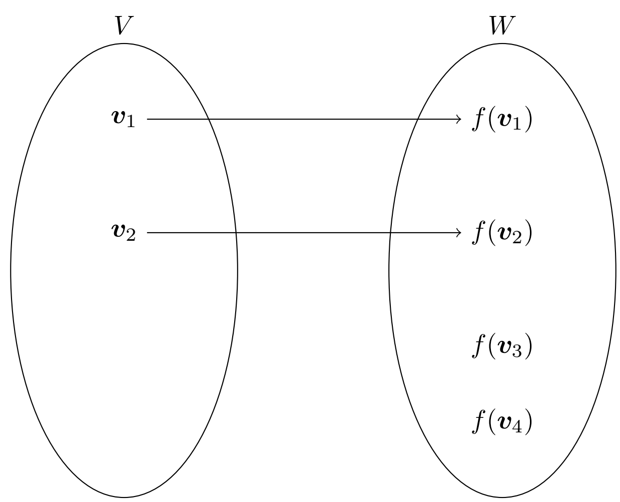
$$\boldsymbol{r}(s,t) = \boldsymbol{r}_0 + s\boldsymbol{u} + t\boldsymbol{v}$$

where $\boldsymbol{r}_0$ is the vector $\boldsymbol{r}_0 = \overrightarrow{OP} = \langle x_1, \ldots, x_n\rangle$. We call $\boldsymbol{r}(s,t)$ the parametrization of P.

Injective

Let $f : V \to W$ be a linear map. We say that f is injective or one-to-one (or sometimes, f is an injection) if the following holds: For all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$, if $f(\boldsymbol{v}_1) = f(\boldsymbol{v}_2)$, then $\boldsymbol{v}_1 = \boldsymbol{v}_2$.

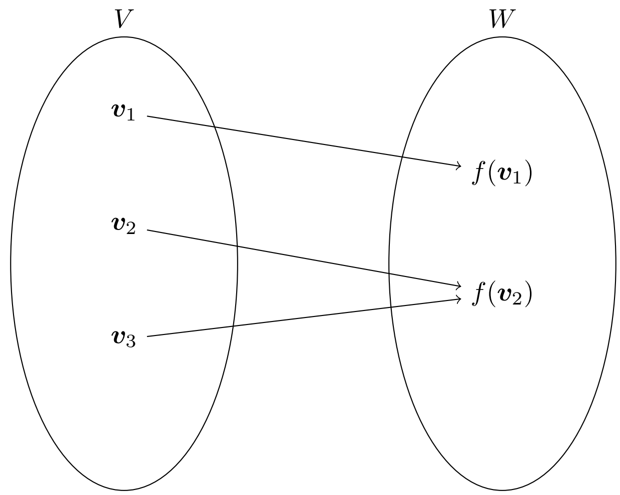
That is, a map f is injective if any element in the codomain of f is the image of at most one element in its domain.



Surjective

Let $f : V \to W$ be a linear map. We say that f is surjective or onto (or sometimes, f is a surjection) if the following holds: For all $w \in W$, there exists a $\boldsymbol{v} \in V$ such that $f(\boldsymbol{v}) = w$.

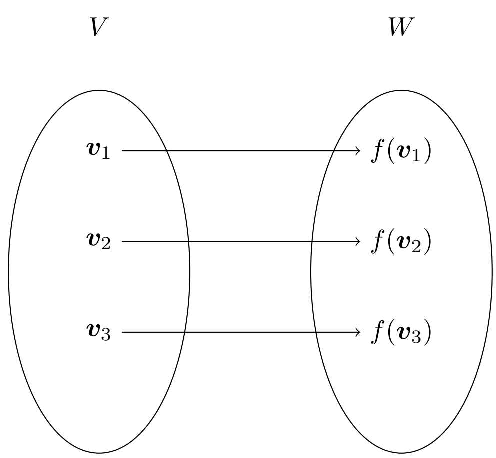
That is, any element in the codomain of f is the image of at least one element in its domain.



Bijective

Let $f : V \to W$ be a linear map. We say that f is bijective (or sometimes, f is a bijection) if f is both injective and surjective.

That is, any element in the codomain of f is the image of exactly one element in its domain. This implies that for all $\boldsymbol{w} \in W$, there exists exactly one $\boldsymbol{v} \in V$ such that $f(\boldsymbol{v}) = \boldsymbol{w}$.



Invertiblility

A linear transformation $T : V \to W$ is invertible if there exists a linear transformation $S : W \to V$ such that $S \circ T = \mathrm{id}_V$ and $T \circ S = \mathrm{id}_W$, where id_V and id_W are the identity maps on V and W, respectively.

Recall that linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be written as matrices. Thus, a matrix $A \in M_{n \times n}(\mathbb{R})$ is invertible if there exists a matrix $B \in M_{n \times n}(\mathbb{R})$ such that $AB = BA = I_n$. Here, B is called the inverse of A.

Isomorphism

A linear transformation $T : V \to W$ is an isomorphism of vector spaces if T satisfies any of the following equivalent conditions:

1. T is invertible.
2. T is bijective.

If $T : V \to W$ is an isomorphism, we say that V and W are isomorphic vector spaces.

We can check if a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$ is an isomorphism by checking if the determinant of the matrix representing the linear transformation is nonzero. This comes from the properties of matrix multiplication and the definition of invertibility.

Formula for a Plane

The plane P in $\mathbb{R}^3$ determined by a point $P_0 = (x_0, y_0, z_0)$ and a normal vector $\boldsymbol{n} = \langle a, b, c\rangle$ is described by the equation:

$$\boldsymbol{n} \cdot \langle x, y, z\rangle = d$$

where we set $d = ax_0 + by_0 + cz_0$.

Hyperplane

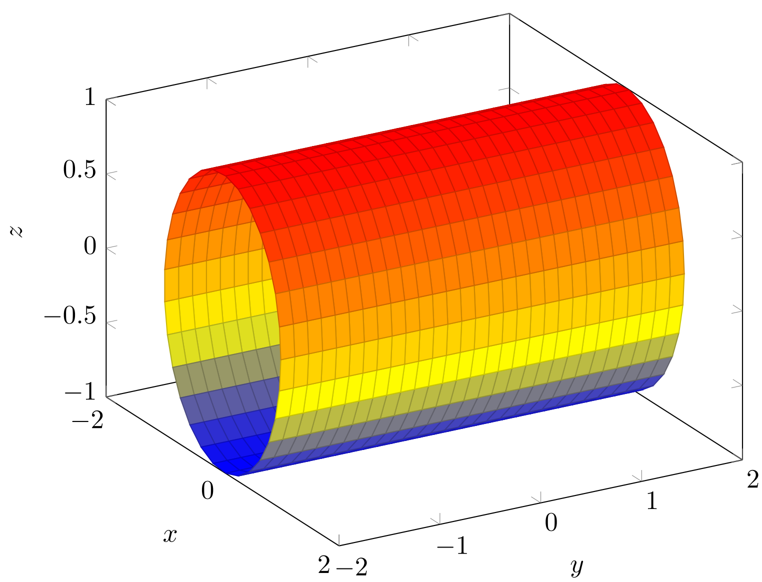
Let $\boldsymbol{n} \in V$, with $\boldsymbol{n} \neq 0$. The hyperplane W normal to $\boldsymbol{n}$ (passing through the origin) is the subspace defined as

$$W = \{\boldsymbol{v} \in V \mid \boldsymbol{n} \cdot \boldsymbol{v} = 0\}$$

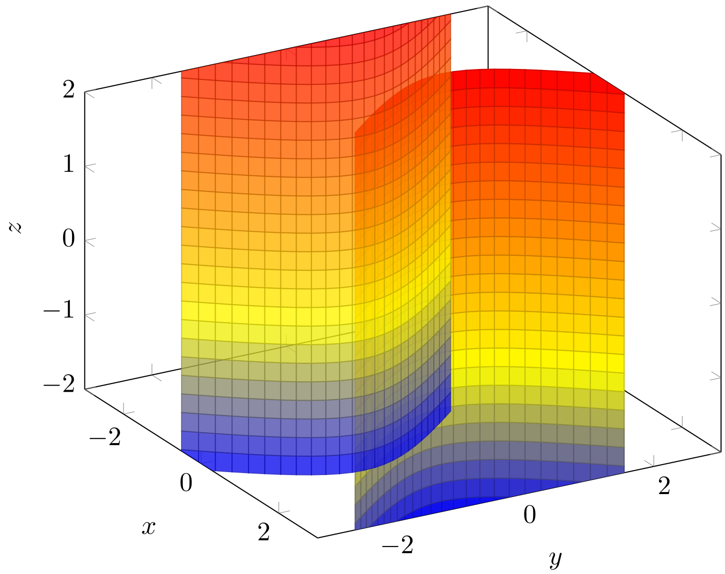
We say that $\boldsymbol{n}$ is a normal vector of W.

Quadric Surfaces

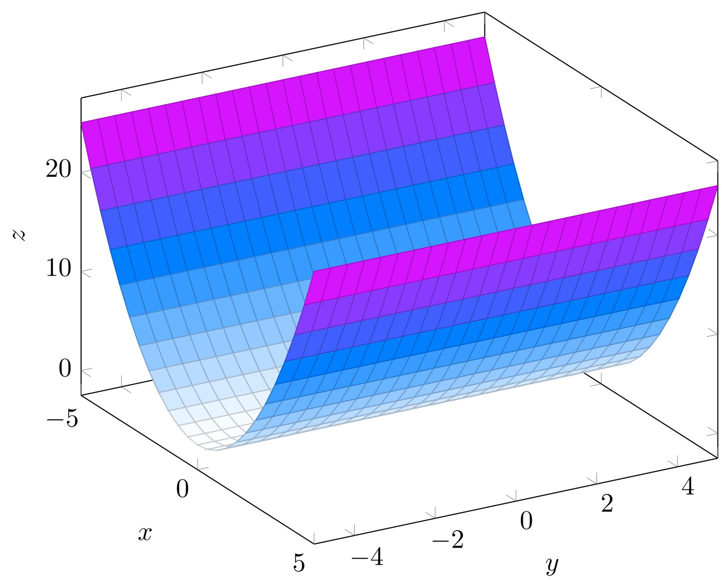
1. Elliptic Cylinder: $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$



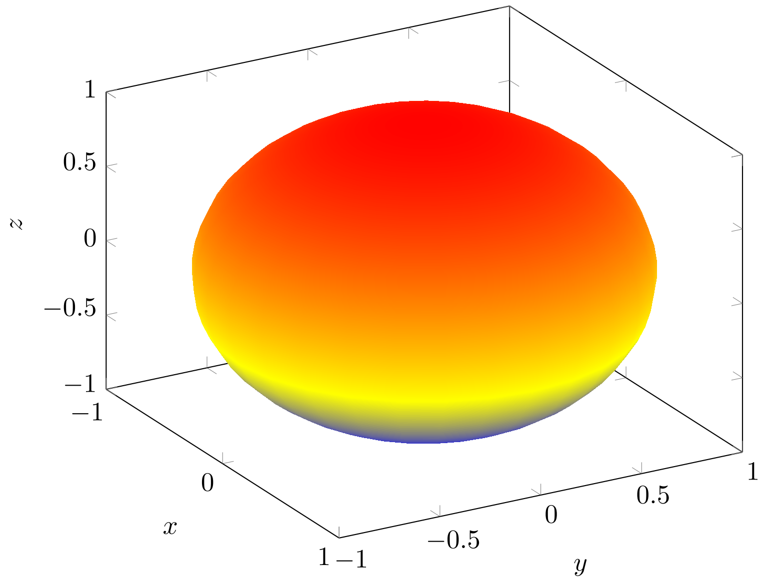
2. Hyperbolic Cylinder: $\left(\frac{y}{b}\right)^2 - \left(\frac{x}{a}\right)^2 = 1$



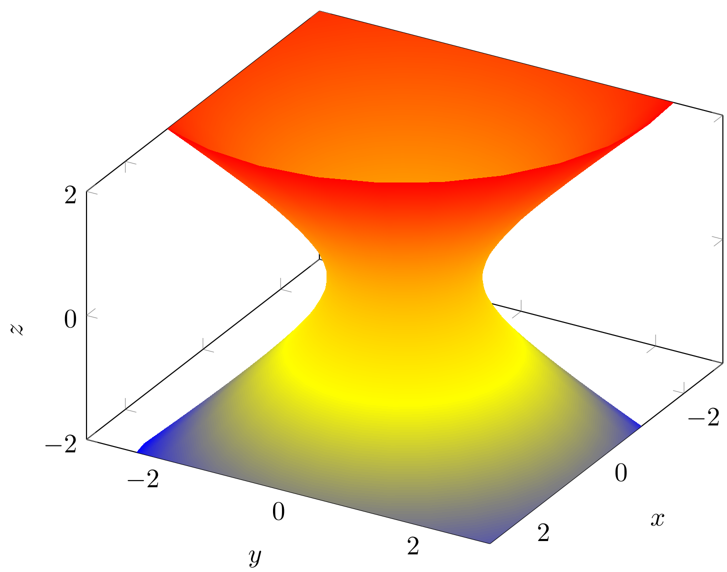
3. Parabolic Cylinder: $y = ax^2$



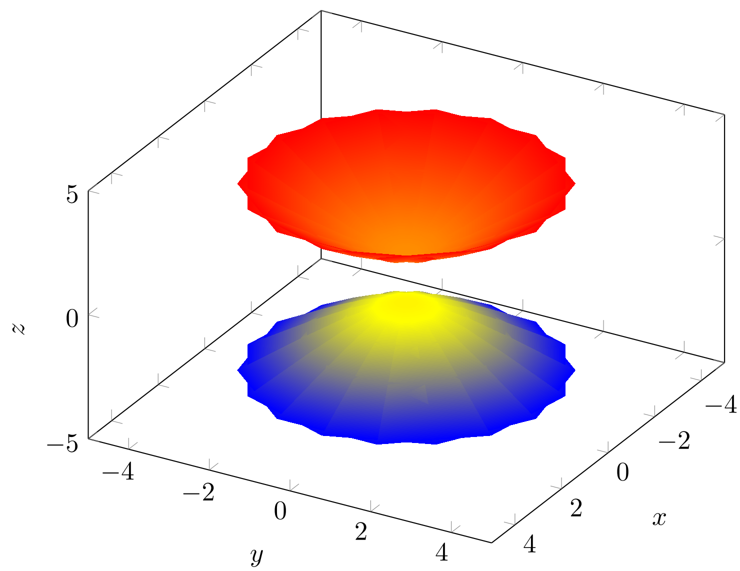
4. Ellipsoid: $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1$



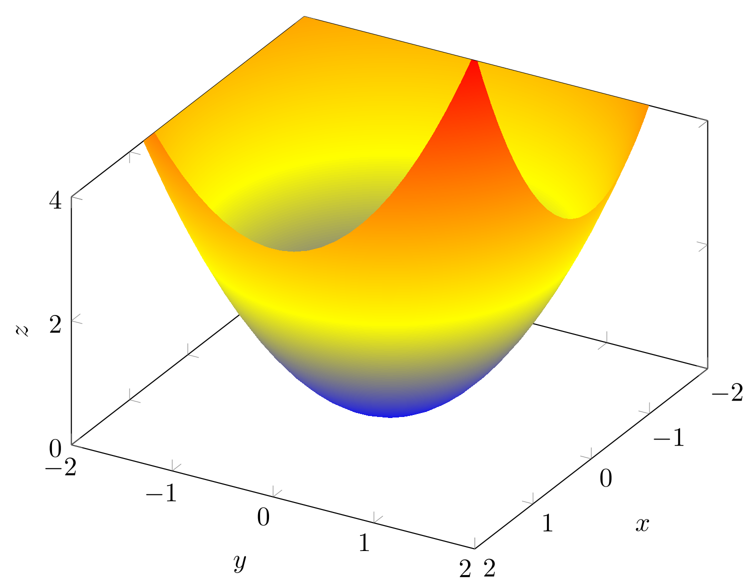
5. Hyperboloid One Sheet: $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = \left(\frac{z}{c}\right)^2 + 1$



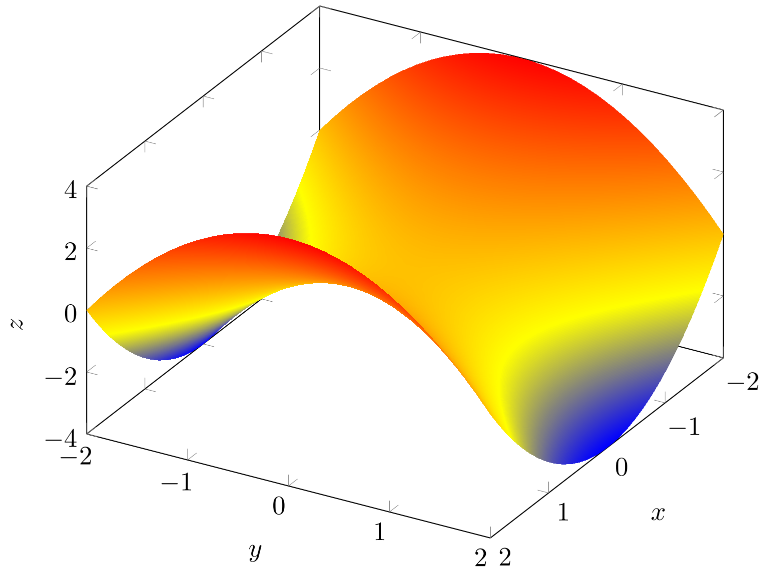
6. Hyperboloid Two Sheets: $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = \left(\frac{z}{c}\right)^2 - 1$



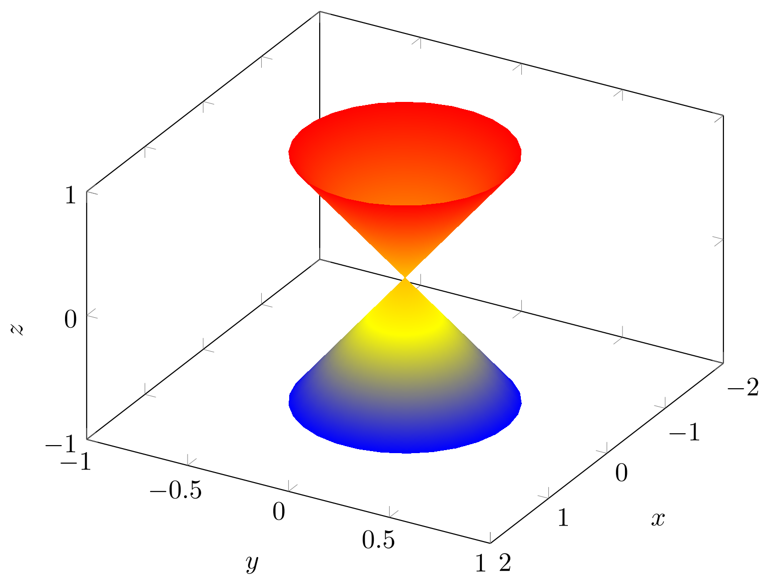
7. Elliptic Paraboloid: $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = z$



8.Hyperbolic Paraboloid: $\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = z$



9. Cone (Elliptical): $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = \left(\frac{z}{c}\right)^2$



Graphs

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, its graph is the following subset of $\mathbb{R}^{n+1}$:

$$\Gamma_f := \{(x_1, \ldots, x_n, f(x_1, \ldots, x_n))\} \subset \mathbb{R}^{n+1}$$

In other words, the graph is given by the equation

$$x_{n+1} = f(x_1, \ldots, x_n)$$

in $\mathbb{R}^{n+1}$.

Traces

The trace in the plane P of a graph $\Gamma \subset \mathbb{R}^3$ is the intersection of Γ with P. That is,

$$\Gamma \cap P = \{x \in \mathbb{R}^3 \mid x \in \Gamma \text{ and } x \in P\}$$

Level Curves

The level curves (isoclines, contour map) of a function of two variables $f(x, y)$ are the z-traces of the graph $z = f(x, y)$.

Vanishing Locus

Given a multivariable function $G(x_1, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R}$, its vanishing locus is the set of points

$$\{(x_1, \ldots, x_n) \mid G(x_1, \ldots, x_n) = 0\}$$

All quadric surfaces are the vanishing loci of the general quadratic equation

$$Q(x, y, z) = Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + ax + by + cz + d$$

Limits

Limit of Sequence Definition

Let $\{\boldsymbol{a_n}\}$ be a sequence of vectors in $\mathbb{R}^k$. We say that the sequence $\{\boldsymbol{a_n}\}$ converges to the vector $\boldsymbol{L} \in \mathbb{R}^k$ if the following holds:

For all $\varepsilon > 0$, there exists an M such that for all $m > M$, $\|\boldsymbol{a_m} - \boldsymbol{L}\| < \varepsilon$.

We say $\boldsymbol{L}$ is the *limit* of the sequence $\{\boldsymbol{a_n}\}$. If no such $\boldsymbol{L}$ exists, we say that $\{\boldsymbol{a_n}\}$ *diverges.*

Definition of a Ball

Let $P \in \mathbb{R}^n$. The open ball of radius ε around P, denoted $B_\varepsilon(P)$, is the set of points defined by

$$B_\varepsilon(P) := \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{P}\| < \varepsilon\}.$$

Subsequences

Let $\{\boldsymbol{a_n}\}$ be a sequence of vectors in $\mathbb{R}^k$. A subsequence of $\{\boldsymbol{a_n}\}$ is a sequence $\{\boldsymbol{b_i}\}$, where

$$\boldsymbol{b_i} = \boldsymbol{a_{n_i}}$$

such that $n_1 < n_2 < \cdots < n_i < \cdots$.

Let $\{\boldsymbol{a_n}\}$ be a sequence of vectors in $\mathbb{R}^k$. If $\{\boldsymbol{a_n}\}$ has a subsequence $\{\boldsymbol{a_{n_i}}\}$ that diverges, then $\{\boldsymbol{a_n}\}$ diverges.

Delta-Epsilon Limit Definition

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ has the limit $\boldsymbol{b}$ at $\boldsymbol{a}$ if the following holds:

For all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $\boldsymbol{x} \in \mathbb{R}^n$,
$0 < \|\boldsymbol{x} - \boldsymbol{a}\| < \delta$ implies $\|f(\boldsymbol{x}) - \boldsymbol{b}\| < \varepsilon$.

Properties of Limits of a function from $\mathbb{R}^n \to \mathbb{R}$

Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be functions of n variables. Suppose that $\lim_{\boldsymbol{x} \to \boldsymbol{P}} f(\boldsymbol{x})$ and $\lim_{\boldsymbol{x} \to \boldsymbol{P}} g(\boldsymbol{x})$ exist. Then

a. Sum Law:

$$\lim_{\boldsymbol{x} \to \boldsymbol{P}} (f(\boldsymbol{x}) + g(\boldsymbol{x})) = \lim_{\boldsymbol{x} \to \boldsymbol{P}} f(\boldsymbol{x}) + \lim_{\boldsymbol{x} \to \boldsymbol{P}} g(\boldsymbol{x})$$

b. Scalar Multiple Law:

$$\lim_{\boldsymbol{x}\to\boldsymbol{P}} \lambda f(\boldsymbol{x}) = \lambda \lim_{\boldsymbol{x}\to\boldsymbol{P}} f(\boldsymbol{x})$$

c. Product Law:

$$\lim_{\boldsymbol{x}\to\boldsymbol{P}} (f(\boldsymbol{x})g(\boldsymbol{x})) = \left(\lim_{\boldsymbol{x}\to\boldsymbol{P}} f(\boldsymbol{x})\right)\left(\lim_{\boldsymbol{x}\to\boldsymbol{P}} g(\boldsymbol{x})\right)$$

d. Quotient Law: If $\lim_{\boldsymbol{x}\to\boldsymbol{P}} g(\boldsymbol{x}) \neq 0$,

$$\lim_{\boldsymbol{x}\to\boldsymbol{P}} \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \frac{\lim_{\boldsymbol{x}\to\boldsymbol{P}} f(\boldsymbol{x})}{\lim_{\boldsymbol{x}\to\boldsymbol{P}} g(\boldsymbol{x})}$$

Limit Point

Let $X \subset \mathbb{R}^n$. We say that a point $p \in \mathbb{R}^n$ is a *limit point* of X if there is a sequence $\{a_n\}$ contained inside X such that $\{a_n\}$ converges to p.

Paths to show a limit does not exist

Let $X \subset \mathbb{R}^n$, let $f : X \to \mathbb{R}^m$ be a function, and let a be a limit point of X. Then the following statements are equivalent:

a. $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}$

b. For every sequence $\{\boldsymbol{a_n}\}$ converging to $\boldsymbol{a}$ (with $\boldsymbol{a_n} \neq \boldsymbol{a}$), the sequence $\{f(\boldsymbol{a_n})\}$ converges to $\boldsymbol{b}$.

In other words, in order for a limit of a multivariable function to exist, it must yield the same value along all possible approaches.

Squeeze Theorem

Let $f(\boldsymbol{x})$, $g(\boldsymbol{x})$, and $h(\boldsymbol{x})$ be functions of n variables such that

$$\lim_{\boldsymbol{x}\to\boldsymbol{P}} f(\boldsymbol{x}) = L = \lim_{\boldsymbol{x}\to\boldsymbol{P}} h(\boldsymbol{x}).$$

If there exists $\delta > 0$ such that for all $\boldsymbol{x} \in B_\delta(\boldsymbol{P}) \setminus \{\boldsymbol{P}\}$, we have that

$$f(\boldsymbol{x}) \leq g(\boldsymbol{x}) \leq h(\boldsymbol{x}).$$

Limits using Polar Coordinates

Let $f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables, which we can express in polar coordinates as $g(r, \theta) := f(r\cos(\theta), r\sin(\theta))$. Then

$$\lim_{(x,y)\to(0,0)} f(x, y) = L$$

if and only if there exists $\delta > 0$ and a function $h : \mathbb{R} \to \mathbb{R}$ such that

1. If $0 < r < \delta$, then $|g(r, \theta) - L| \leq h(r)$ for all θ, **AND**
2. $\lim_{r\to 0} h(r) = 0$.

Corollary 2.3.14. If $\lim_{r\to 0} g(r, \theta)$ depends on θ, then the value of the limit will differ for different straight line paths. Thus, $\lim_{(x,y)\to(0,0)} f(x, y)$ does not exist.

Derivatives

Limit Definition of the Derivative

A multivariable function $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$ is *differentiable* at an interior point $\boldsymbol{x_0}$ of A if there exists a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{\|f(\boldsymbol{x_0}+\boldsymbol{h}) - f(\boldsymbol{x_0}) - T(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = 0.$$

The derivative of f at $\boldsymbol{x_0}$ is the linear transformation $Df(\boldsymbol{x_0}) := T$. By our characterization of linear transformations, $Df(\boldsymbol{x_0}) : \mathbb{R}^m \to \mathbb{R}^n$ corresponds to a matrix $[Df(\boldsymbol{x_0})] \in M_{n\times m}(\mathbb{R})$.

Chain Rule

Let $f : \mathbb{R}^n \to \mathbb{R}^m$, and let $g : \mathbb{R}^m \to \mathbb{R}^k$ be multivariable functions such that f is differentiable at $\boldsymbol{x_0} \in \mathbb{R}^n$, and g is differentiable at $f(\boldsymbol{x_0}) \in \mathbb{R}^m$. Then $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ is differentiable at $\boldsymbol{x_0} \in \mathbb{R}^n$, and

$$D(g\circ f)(\boldsymbol{x_0}) = Dg(f(\boldsymbol{x_0})) \circ Df(\boldsymbol{x_0})$$

We can prove this using the definition of the derivative. However, since we know that the derivative can be computed in terms of the Jacobian, we equivalently have

In Coordinates: Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\boldsymbol{x_0} \in \mathbb{R}^n$, and $g : \mathbb{R}^m \to \mathbb{R}^k$ is differentiable at $f(\boldsymbol{x_0}) \in \mathbb{R}^m$. Then

$$[J_{g\circ f}(\boldsymbol{x_0})] = [J_g(f(\boldsymbol{x_0}))][J_f(\boldsymbol{x_0})]$$

For Paths: Let $f(x_1, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function, and let $\boldsymbol{r}(t) = \langle x_1(t), \ldots, x_n(t)\rangle : \mathbb{R} \to \mathbb{R}^n$ be a vector-valued function. Then $f(\boldsymbol{r}(t)) : \mathbb{R} \to \mathbb{R}$ is a single-variable function, and the derivative of f at t_0 along the path $\boldsymbol{r}(t)$ is given by

$$\frac{d}{dt} f(\boldsymbol{r}(t_0)) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{r}(t_0))x_i'(t_0)$$

where $x_i'(t_0)$ is the derivative of the i-th component of $\boldsymbol{r}(t)$ at t_0. This measures the rate of change of f along the path $\boldsymbol{r}(t)$.

$$f(\mathbf{r}(t_0)) = \nabla f(\mathbf{r}(t_0)) \cdot \mathbf{r}'(t_0)$$

The Jacobian

Let $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$ be a multivariable function defined by $f^i : A \subset \mathbb{R}^m \to \mathbb{R}$:

$$f(\boldsymbol{x}) = \begin{bmatrix} f^1(\boldsymbol{x}) \\ \vdots \\ f^n(\boldsymbol{x}) \end{bmatrix}.$$

The Jacobian matrix of f at $\boldsymbol{x_0}$ is

$$[J_f(\boldsymbol{x_0})] = \begin{bmatrix} D_1 f^1(\boldsymbol{x_0}) & D_2 f^1(\boldsymbol{x_0}) & \cdots & D_m f^1(\boldsymbol{x_0}) \\ D_1 f^2(\boldsymbol{x_0}) & D_2 f^2(\boldsymbol{x_0}) & \cdots & D_m f^2(\boldsymbol{x_0}) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f^n(\boldsymbol{x_0}) & D_2 f^n(\boldsymbol{x_0}) & \cdots & D_m f^n(\boldsymbol{x_0}) \end{bmatrix}$$

if the partial derivatives exist.

Directional Derivative

If $\mathbf{u} = \langle u_1, \ldots, u_n \rangle$ is a unit vector in $\mathbb{R}^n$, then the directional derivative of a function f at the point $\mathbf{x_0} \in \mathbb{R}^n$ in the direction of $\mathbf{u}$ is defined as

$$D_{\mathbf{u}} f(\mathbf{x_0}) = u_1 \frac{\partial f}{\partial x_1}(\mathbf{x_0}) + \cdots + u_n \frac{\partial f}{\partial x_n}(\mathbf{x_0}).$$

Gradient

If $f(x_1, \ldots, x_n)$ is a function of n variables, then the *gradient* of f is the vector-valued function given by

$$\nabla f = \left\langle \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle.$$

That is, ∇f is the transpose of the matrix of partial derivatives of f,

$$\nabla f = \left[D(f(\boldsymbol{x_0})) \right]^\top,$$

where $[A]^\top$ indicates the transpose matrix.
Thinking of z as the height of $z = f(x, y)$, the gradient ∇f points in the direction of steepest ascent.
The opposite of the gradient, $-\nabla f$, points in the direction of steepest descent.

Linear Approximation

If $f : A \subset \mathbb{R}^m \to \mathbb{R}$ is differentiable at a point $\boldsymbol{a} = (a_1, \ldots, a_n)$, and $\boldsymbol{x} = (x_1, \ldots, x_n)$ is close to $\boldsymbol{a}$, then

$$\begin{aligned} f(\boldsymbol{x}) &\approx f(\boldsymbol{a}) + \begin{bmatrix} D_1 f(\boldsymbol{a}) & D_2 f(\boldsymbol{a}) & \cdots & D_n f(\boldsymbol{a}) \end{bmatrix} (\boldsymbol{x} - \boldsymbol{a}) \\ &= f(\boldsymbol{a}) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(\boldsymbol{a}) \right) (x_i - a_i). \end{aligned}$$

Critical Point

A point $\boldsymbol{P} \in \mathbb{R}^n$ is said to be a *critical point* of a function $f : \mathbb{R}^n \to \mathbb{R}$ if either

a. $Df(\boldsymbol{P}) = 0$, **OR**

b. $Df(\boldsymbol{P})$ does not exist.

Hessian Matrix

The Hessian matrix of $f : \mathbb{R}^n \to \mathbb{R}$ at $\boldsymbol{x_0}$ is

$$[H_f(\boldsymbol{x_0})] = \begin{bmatrix} D_1 D_1 f(\boldsymbol{x_0}) & D_2 D_1 f(\boldsymbol{x_0}) & \cdots & D_n D_1 f(\boldsymbol{x_0}) \\ D_1 D_2 f(\boldsymbol{x_0}) & D_2 D_2 f(\boldsymbol{x_0}) & \cdots & D_n D_2 f(\boldsymbol{x_0}) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 D_n f(\boldsymbol{x_0}) & D_2 D_n f(\boldsymbol{x_0}) & \cdots & D_n D_n f(\boldsymbol{x_0}) \end{bmatrix},$$

where $D_i D_j f(\boldsymbol{x_0})$ denotes the second partial derivative of f with respect to x_i and then x_j at $\boldsymbol{x_0}$.

Clairaut's Theorem.

Let $f : \mathbb{R}^n \to \mathbb{R}$. Suppose that $D_i f$, $D_j f$, and $D_i D_j f$ exist and are continuous on an open disk $D \subset \mathbb{R}^n$. Then $D_j D_i f$ exists on D, and moreover, $D_i D_j f = D_j D_i f$ on the disk D.

Second Derivative Test

Let $\boldsymbol{x_0} \subset U$ be a critical point of $f(x, y) : U \to \mathbb{R}$, and suppose that f is in $C^2(U)$. Let us write $D = \det[H_f(\boldsymbol{x_0})]$.

a. If $D > 0$ and $f_{xx}(\boldsymbol{x_0}) > 0$, then there is a local minimum at $\boldsymbol{x_0}$.

b. If $D > 0$ and $f_{xx}(\boldsymbol{x_0}) < 0$, then there is a local maximum at $\boldsymbol{x_0}$.

c. If $D < 0$, then f has a saddle point at $\boldsymbol{x_0}$.

d. If $D = 0$ or does not exist, then the test is inconclusive.

Convex Subset

A subset $A \subset \mathbb{R}^n$ is said to be a convex subset of $\mathbb{R}^n$ if it contains the line segment joining any two points of A. That is, for all $\boldsymbol{a}, \boldsymbol{b} \in A$, and for all $t \in [0, 1]$, then $\boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a}) \in A$.

Bounded

A subset $D \subset \mathbb{R}^n$ is *bounded* if there exists some $r > 0$ such that

$$D \subset B_r(\mathbf{0}).$$

Boundary Point

A point $\boldsymbol{x_0} \in \mathbb{R}^n$ is a *boundary point* of $D \subset \mathbb{R}^n$ if: for all $\varepsilon > 0$,

a. $B_\varepsilon(\boldsymbol{x_0}) \cap D$ is non-empty, and

b. $B_\varepsilon(\boldsymbol{x_0}) \cap D^c$ is non-empty,

where D^c is the complement of D in $\mathbb{R}^n$.

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

Langrange Multipliers

Assume that $f(x, y)$ and $g(x, y)$ are differentiable functions. If

a. $f(x, y)$ has a local maximum or minimum subject to the constraint $g(x, y) = 0$ at a point (a, b), **AND**

b. $\nabla g(a, b) \neq 0$

then there is a scalar λ such that

$$\nabla f(a, b) = \lambda \nabla g(a, b).$$

We can use the Lagrange equations

$$f_x(a, b) = \lambda g_x(a, b)$$

and

$$f_y(a, b) = \lambda g_y(a, b).$$

Global Max/Min

If D is a closed and bounded subset of $\mathbb{R}^n$, and f is a continuous function on D, then f has a global maximum and a global minimum in D. That is, there exists an $M \in D$ and an $m \in D$ such that

$$f(m) \leq f(x) \leq f(M)$$

for all $x \in D$.

Frenet Frame Formulas

$$\boldsymbol{T}(t) = \frac{1}{||\boldsymbol{r}'(t)||}\boldsymbol{r}'(t)$$

$$\boldsymbol{N}(t) = \frac{1}{||\boldsymbol{T}'(t)||}\boldsymbol{T}'(t)$$

$$\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$$

$$\boldsymbol{B}(t) = \frac{\boldsymbol{r}'(t) \times \boldsymbol{r}''(t)}{||\boldsymbol{r}'(t) \times \boldsymbol{r}''(t)||}$$

$$\boldsymbol{N} = \boldsymbol{B} \times \boldsymbol{T}$$

$$\kappa(t) = \frac{1}{||\boldsymbol{r}'(t)||}||\boldsymbol{T}'(t)||$$

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

$$\kappa(t) = \frac{1}{R}$$