Notes on the transition from $\mathbb{R}$ to $\mathbb{R}^n$

A. Brilli[1]

These notes are a summary of the material used for the Mathematical Programming course for the Bachelor program in Computer Engineering[2]

---

[1] Dip. Ingegneria Informatica Automatica e Gestionale, "Sapienza" Univ. di Roma

[2] Acknowledgement: the material was developed by revisiting Prof. Liuzzi's lecture notes.

Recalls on $\mathbb{R}$

- **Open spherical neighborhood** of $x_o \in \mathbb{R}$

$$B(x_0, \rho) = \{x \in \mathbb{R} : |x - x_0| < \rho\}, \quad \rho > 0$$

$B(x_0, \rho) = (x_0 - \rho, x_0 + \rho)$ (open interval centered at x_0)

$$B(3, 2) = ?$$

- **Open neighborhood** of $x_0 \in \mathbb{R}$

$$I(x_0) = (a, b)$$

any open interval containing $B(x_0, \rho)$ for some $\rho > 0$.

An open neighborhood of 3?

Accumulation points on $\mathbb{R}$

Let $A \subseteq \mathbb{R}$ be a subset of $\mathbb{R}$

- $y \in \mathbb{R}$ is an accumulation point for A when

$$(B(y,\rho) \cap A) \setminus \{y\} \neq \emptyset, \qquad \forall \rho > 0$$

$$A = (1,7]$$

which of these points are accumulation points for A?

$$y = 0, \qquad y = 1, \qquad y = 3, \qquad y = 7$$

which accumulation points of A belong to A?

Derived set and isolated points

Given $A \subseteq \mathbb{R}^*$

- **Derived set** of A,
$$\mathcal{D}A = \{x \in \mathbb{R}^* : \ x \text{ accumulation point of } A\}$$
- $x \in A$ is **isolated** in A when
$$x \in A \setminus \mathcal{D}A$$
- $x \in A$ is **non-isolated** in A when
$$x \in A \cap \mathcal{D}A$$

$$A = (1,7] \qquad \mathcal{D}A?$$
$$A = (1,7] \cup \{0\} \qquad \mathcal{D}A?$$
$$A = [1,7] \cup \{0\} \qquad \mathcal{D}A?$$

Cartesian product

- Given two sets A and B, **Cartesian product**

$$A \times B = \{(x, y) : x \in A,\ \ y \in B\}$$

- Given three or more sets $A_1, A_2, \dots, A_k$

$$A_1 \times A_2 \times \cdots \times A_k = \{(x_1, x_2, \dots, x_k) : x_i \in A_i, i = 1, \dots, k\}$$

- Given a set A

$$A \times A \triangleq A^2 = \{(x, y) : x, y \in A\}$$

- What does $\mathbb{R}^n$ mean for $n \in \mathbb{N}, n \geq 2$?

$\mathbb{R}^n$

$$\mathbb{R}^n \triangleq \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}} = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R}\}$$

- $\mathbb{R}^n$ is the Cartesian product of **totally ordered** sets ($\mathbb{R}$)
- $\mathbb{R}^n$ **is ordered**, that is it is possible to define on $\mathbb{R}^n$ an order relation (reflexive, antisymmetric, transitive). **Which one?**
- Is $\mathbb{R}^n$ totally ordered?

Functions on $\mathbb{R}^n$

Definition (Real-valued)

A function $f : \mathbb{R}^n \to \mathbb{R}$ *is a rule that assigns a value* $f(x) \in \mathbb{R}$ *to an element* $x \in \mathbb{R}^n$.

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

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ *is a map that assigns an element* $f(x) \in \mathbb{R}^m$ *to an element* $x \in \mathbb{R}^n$.

Operations on $\mathbb{R}^n$

- **sum** of two elements $x, y \in \mathbb{R}^n$

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

- scalar-vector product $\alpha \in \mathbb{R}$, $x \in \mathbb{R}^n$

$$\alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$

$\mathbb{R}^n$ is **closed** with respect to the two operations above

$\mathbb{R}^n$ with the operations of addition and scalar multiplication is a **vector space**

Operations on $\mathbb{R}^n$

- product between vectors? given two vectors $x, y \in \mathbb{R}^n$
  1. **Hadamard product** returns a vector
  2. **inner or scalar product** returns a number

Hadamard product

$$x \circ y = (x_1 y_1, x_2 y_2, \ldots, x_n y_n)$$

The following properties hold:

1. *commutative* property: given $x, y \in \mathbb{R}^n$, $x \circ y = y \circ x$;
2. *associative* property: given $x, y, v \in \mathbb{R}^n$, $x \circ y \circ v = (x \circ y) \circ v = x \circ (y \circ v)$;
3. *distributive* property with respect to addition: given $x, y, v \in \mathbb{R}^n$, $x \circ (y + v) = (x \circ y) + (x \circ v)$;
4. *homogeneity* property: given $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, $\alpha(x \circ y) = (\alpha x) \circ y$;
5. *existence of the identity element*: the vector $\mathbf{1} = \underbrace{(1, 1, \ldots, 1)}_{n \text{ times}}$ is such that, for every $x \in \mathbb{R}^n$,
$$\mathbf{1} \circ x = x \circ \mathbf{1} = x$$
6. *existence of the zero element*: the origin of $\mathbb{R}^n$ is such that, for every $x \in \mathbb{R}^n$,
$$\mathbf{0} \circ x = x \circ \mathbf{0} = \mathbf{0}.$$

Scalar (or inner) product

$$\langle x, y\rangle = x^\top y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

enjoys the properties

1. *commutative* property: given $x, y \in \mathbb{R}^n$: $x^\top y = y^\top x$;
2. *distributive* property with respect to addition: if $x, y, v \in \mathbb{R}^n$: $x^\top (y + v) = (x^\top y) + (x^\top v)$
3. *homogeneity* property: given $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, $\alpha(x^\top y) = (\alpha x)^\top y$;
4. *existence of the zero element*: the origin of $\mathbb{R}^n$ is such that, for every $x \in \mathbb{R}^n$,
   $$\mathbf{0}^\top x = x^\top \mathbf{0} = 0.$$
5. moreover:
   $$\langle x, x\rangle \geq 0, \ \langle x, x\rangle = 0 \ \Leftrightarrow \ x = \mathbf{0}.$$

Vector space $\mathbb{R}^n$

Let $\mathbb{R}^n$ be the set of real *n*-tuples, or vectors with *n* real components

A vector $x \in \mathbb{R}^n$ will **always** be for us a **column vector** with *n* components

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

or alternatively

$$x = (x_1,\ x_2,\ \ldots,\ x_n)^\top$$

Distances in $\mathbb{R}$

On the real line, given $x \in \mathbb{R}$,

$$|x| \equiv \text{distance of } x \text{ from the origin}$$

The function $|x|$ satisfies the following

i) $\boxed{|x| \geq 0}$, that is the absolute value of any real number is a positive number, at most zero;

ii) $\boxed{|x| = 0 \text{ if and only if } x = 0}$, that is the origin has zero absolute value, and the only real number whose absolute value is zero is the origin;

iii) $\boxed{|x + y| \leq |x| + |y|}$, that is the absolute value of the sum of two numbers is always less than, or at most equal to, the sum of their absolute values;

iv) $\boxed{|\alpha x| = |\alpha| \cdot |x|}$, that is the absolute value of the product of two real numbers equals the product of their absolute values.

Distances in $\mathbb{R}^2$

In the Cartesian plane $\mathbb{R}^2$, given $(x, y) \in \mathbb{R}^2$

$$\|(x, y)\| \triangleq \sqrt{x^2 + y^2}$$

Do the properties of $|\cdot|$ hold?

i) $\boxed{\|(x, y)\| \geq 0}$?

ii) $\boxed{\|(x, y)\| = 0 \text{ if and only if } (x, y) = (0, 0)}$?

iii) $\boxed{\|(x, y) + (v, w)\| \leq \|(x, y)\| + \|(v, w)\|}$?

iv) $\boxed{\|\alpha(x, y)\| = |\alpha| \cdot \|(x, y)\|}$?

… in $\mathbb{R}^3$

In Cartesian space $\mathbb{R}^3$, given $(x, y, z) \in \mathbb{R}^3$

$$\|(x, y, z)\| \triangleq ?$$

Do the properties of $|\cdot|$ hold?

i) $\boxed{\|(x, y, z)\| \geq 0}$?

ii) $\boxed{\|(x, y, z)\| = 0 \text{ if and only if } (x, y, z) = (0, 0, 0)}$?

iii) $\boxed{\|(x, y, z) + (v, w, u)\| \leq \|(x, y, z)\| + \|(v, w, u)\|}$?

iv) $\boxed{\|\alpha(x, y, z)\| = |\alpha| \cdot \|(x, y, z)\|}$?

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

What if we wanted to extend the concept of $|x|$ in $\mathbb{R}$ to $\mathbb{R}^n$?

$(x, y) \in \mathbb{R}^2$ point in the plane with respect to an orthogonal Cartesian reference frame.

$$\|(x,y)\| \equiv \text{distance of } (x,y) \text{ from the origin} = \sqrt{x^2 + y^2}$$

More generally, in $\mathbb{R}^n$, given $x \in \mathbb{R}^n$,

$$\|x\| \equiv \text{distance of } x \text{ from the origin} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

N.B. it satisfies properties (i)–(iv) of the absolute value in $\mathbb{R}$

Norm in $\mathbb{R}^n$

Definition (Norm)

A "norm" on $\mathbb{R}^n$ is a function that associates to each point $x \in \mathbb{R}^n$ a number $\|x\|$ enjoying the properties

- i) $\|x\| \geq 0$;
- ii) $\|x\| = \mathbf{0}$ *if and only if* $x = \mathbf{0}$;
- iii) $\|x + y\| \leq \|x\| + \|y\|$;
- iv) $\|\alpha x\| = |\alpha| \cdot \|x\|$.

For example, for $p \geq 1$, $p \in \mathbb{N}$, the following is a norm

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Property

Considering a norm $\|\cdot\|$ and two vectors $x, y \in \mathbb{R}^n$,

$$\|x - y\| \geq \Big| \|x\| - \|y\| \Big|$$

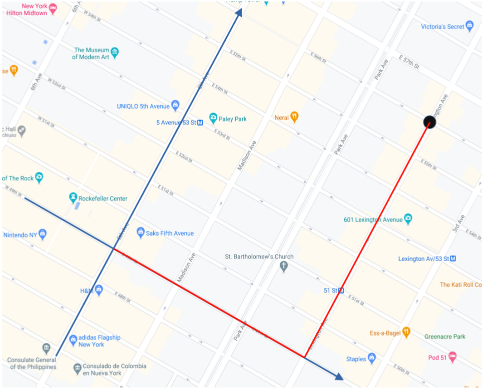
Other norms

The Euclidean norm (or **2-norm** or ℓ_2 norm) is not the only norm that can be defined on $\mathbb{R}^n$; there are many others

The most common ones

- 1-norm or ℓ_1, also known as the Manhattan or Taxicab norm

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$



Other norms

The Euclidean norm (or **2-norm** or ℓ_2 norm) is not the only norm that can be defined on $\mathbb{R}^n$; there are many others

The most common ones

- 1-norm or ℓ_1, also known as the Manhattan or Taxicab norm

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

- ∞-norm or ℓ_∞, also known as the Chebyshev norm

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$$

- Euclidean norm
- 3-norm or ℓ_3

$$\|x\|_3 = \sqrt[3]{|x_1|^3 + |x_2|^3 + \cdots + |x_n|^3}$$

- ...
- p-norm or ℓ_p

$$\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}$$

Neighborhoods, interior, exterior, and boundary points

Definition

Given $x \in \mathbb{R}^n$ *and a scalar* $\epsilon \in \mathbb{R}^+$ *(*$\epsilon > 0$*), the* **open spherical neighborhood** *of center* x *and radius* ϵ *is the set*

$$B_\epsilon(x) = B(x;\epsilon) = \{y \in \mathbb{R}^n : \; d(x,y) < \epsilon\} = \{y \in \mathbb{R}^n : \; \|y - x\| < \epsilon\}$$

Given $E \subseteq \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$:

1. there exists $\epsilon > 0$:
$$B_\epsilon(x) \subseteq E \quad \Longrightarrow \quad \textbf{interior point of } E$$
2. there exists $\epsilon > 0$:
$$B_\epsilon(x) \subseteq co(E) \quad \Longrightarrow \textbf{exterior point of } E$$
3. for every $\epsilon > 0$:
$$\begin{matrix} B_\epsilon(x) \cap E \neq \emptyset \\ B_\epsilon(x) \cap co(E) \neq \emptyset \end{matrix} \quad \Longrightarrow \quad \textbf{boundary point of } E \text{ and } co(E)$$

Interior, boundary, and closure of a set

Definition (Interior of E)

The **interior** *of E is the set of all interior points of E*

$$\overset{\circ}{E} = \{x \in \mathbb{R}^n : x \text{ is interior to } E\}$$

Definition (Boundary of E)

The **boundary** *of E is the set of all boundary points of E*

$$\partial E = \{x \in \mathbb{R}^n : x \text{ is a boundary point of } E\}$$

Therefore, by definition, $\partial E = \partial Co(E)$

Definition (Closure of E)

The **closure** *of E is the set* $\bar{E} = E \cup \partial E$

Remarks

For every $x \in E$, either $x \in \overset{\circ}{E}$ or $x \in \partial E$. Moreover, $\overset{\circ}{E} \cap \partial E = \emptyset$

Therefore (it can be shown)

- $\overset{\circ}{E} \cap \partial E = \emptyset$
- $\overset{\circ}{E} \subseteq E \subseteq \bar{E}$
- $\bar{E} = \overset{\circ}{E} \cup \partial E$
- $\partial E = \bar{E} \setminus \overset{\circ}{E}$

Bounded set

Let $E \subseteq \mathbb{R}^n$

Definition

E is called **bounded** *if there exists a number $r > 0$ such that*

$$E \subseteq B_r(\mathbf{0})$$

Equivalently, E is bounded if there exists $r > 0$ such that $\|x\| < r$ for every $x \in E$

The Bolzano-Weierstrass Theorem

Theorem

If E is a **bounded** *and* **infinite** *subset of $\mathbb{R}^n$, then there exists at least one accumulation point x for E in $\mathbb{R}^n$*

Open sets, interior, boundary

$x \in \mathbb{R}^n$ **interior**	$x \in \mathbb{R}^n$ **exterior**	$x \in \mathbb{R}^n$ **boundary**
$\exists\ \epsilon > 0 : B(x, \epsilon) \subseteq E$	$\exists\ \epsilon > 0 : B(x, \epsilon) \subseteq co(E)$	$\forall\ \epsilon > 0 : \begin{array}{l} B(x,\epsilon) \cap E \neq \emptyset \\ B(x,\epsilon) \cap co(E) \neq \emptyset \end{array}$

$x \in \mathbb{R}^n$ **accumulation point of** E	$x \in E$ **isolated in** E
$\forall\ \epsilon > 0 : (B(x, \epsilon) \cap E) \setminus \{x\} \neq \emptyset$	$\exists\ \epsilon > 0 : (B(x, \epsilon) \cap E) \setminus \{x\} = \emptyset$

E **open** (by definition) when $\forall\ x \in E$, x is interior to E

E **closed** (by definition) when $co(E)$ is open

E closed $\Leftrightarrow \partial E \subseteq E \Leftrightarrow E$ contains all its accumulation points

$$E \text{ neither closed nor open} \Leftrightarrow \begin{cases} \exists\ x \in \partial E \cap E \text{ at least one point is not interior} \\ \exists\ x \in \partial E \setminus E \text{ at least one boundary point of } \partial E \text{ is not in } E \end{cases}$$

Topological properties of points in $\mathbb{R}^n$

Let $E \subseteq \mathbb{R}^n$

1. $x \in \mathbb{R}^n$ is **interior** to E when $\exists\ \epsilon > 0 :\ B(x, \epsilon) \subseteq E$
    - **interior** of E, $int(E) = \overset{\circ}{E}$, set of all interior points of E
    - E is **open** when $E = \overset{\circ}{E}$
    - E is **closed** when $co(E)$ is open
2. $x \in \mathbb{R}^n$ is **exterior** to E when $\exists\ \epsilon > 0 :\ B(x, \epsilon) \subseteq co(E)$
3. $x \in \mathbb{R}^n$ that is neither interior nor exterior is a **boundary point** of E
    - that is when $\forall\ \epsilon > 0 : B(x, \epsilon) \cap E \neq \emptyset$ and $B(x, \epsilon) \cap co(E) \neq \emptyset$
    - **boundary** of E, ∂E, set of all boundary points of E
    - by definition it follows that $\partial E = \partial co(E)$

NB given $E \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$, either x is interior to E, or x is exterior to E, or x is a boundary point of E, i.e., exactly one of the following holds

1. $x \in \overset{\circ}{E}$, or
2. $x \in\ int(co(E))$, or
3. $x \in\ \partial E$

therefore $\mathbb{R}^n = \overset{\circ}{E} \cup int(co(E)) \cup \partial E$

Properties of open and closed sets

1. $\mathbb{R}^n$ and $\emptyset$ are both open and closed, and they are the only sets with this property
2. the union of a finite or infinite number of open sets is always open
3. the intersection of a finite number of open sets is open
4. the intersection of an infinite number of open sets may fail to be open
5. the intersection of a finite or infinite number of closed sets is always closed
6. the union of a finite number of closed sets is closed
7. the union of an infinite number of closed sets may fail to be closed

Limits in $\mathbb{R}$

Let $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, $x_0 \in \mathbb{R}^*$ be an accumulation point of D. When do we say that

$$\lim_{x \to x_0} f(x) = \ell \in \mathbb{R}^*$$

Definition

Let $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, $x_0 \in \mathbb{R}^*$ be an accumulation point of D.

$$\lim_{x \to x_0} f(x) = \ell \in \mathbb{R}^*$$

when **for every neighborhood $\mathcal{U}(\ell)$** **there exists a neighborhood $\mathcal{N}(x_0)$** *such that*

$$x \in (\mathcal{N}(x_0) \cap D) \setminus \{x_0\} \;\Rightarrow\; f(x) \in \mathcal{U}(\ell)$$

Limits in $\mathbb{R}^n$

Let $E \subseteq \mathbb{R}^n$, $f\colon E \to \mathbb{R}$, and $x_0 \in \mathbb{R}^n_*$ be an accumulation point of E, $\ell \in \mathbb{R}^*$

Definition (Limit)

We say that

$$\lim_{x \to x_0} f(x) = \ell$$

when **for every neighborhood $\mathcal{U}(\ell)$** **there exists a neighborhood $\mathcal{N}(x_0)$** *such that*

$$x \in (\mathcal{N}(x_0) \cap D) \setminus \{x_0\} \;\Rightarrow\; f(x) \in \mathcal{U}(\ell)$$

Limits of vector-valued functions

Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$, that is

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

Let $f\colon D \to \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, $x \in \mathbb{R}^n_*$ ($x_0 \in \mathbb{R}^n$ or $x_0 = \infty$) be an accumulation point of D

$$\lim_{x \to x_0} f(x) = \ell \in \mathbb{R}^m_* \quad (\ell \in \mathbb{R}^m \text{ or } \ell = \infty)$$

when (by definition) **for every** neighborhood $\mathcal{U}(\ell)$ **there exists a** neighborhood $\mathcal{N}(x_0)$ such that

$$x \in (\mathcal{N}(x_0) \cap D) \setminus \{x_0\} \;\Rightarrow\; f(x) \in \mathcal{U}(\ell)$$

that is

$$\text{if } x \in (\mathcal{N}(x_0) \cap D) \setminus \{x_0\} \text{ then } f(x) \in \mathcal{U}(\ell)$$

Limits for $f\colon \mathbb{R}^n \to \mathbb{R}^m$ – properties

Proposition

Let $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, $x_0 \in \mathbb{R}^n_*$ *be an accumulation point of* D, *and* $\ell \in \mathbb{R}^m$

$$\lim_{x \to x_0} f(x) = \ell \quad \Longleftrightarrow \quad \begin{cases} \lim\limits_{x \to x_0} f_1(x) = \ell_1 \\ \lim\limits_{x \to x_0} f_2(x) = \ell_2 \\ \quad\vdots \\ \lim\limits_{x \to x_0} f_m(x) = \ell_m \end{cases}$$

Properties of limits

Definition

Let $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, $x_0 \in \mathbb{R}^n_*$ be an accumulation point of D. We say that f has a certain property $\mathcal{P}$ **eventually** *as x tends to x_0 when:*
there exists a neighborhood $\mathcal{U}(x_0)$ such that f has property $\mathcal{P}$ for every point $x \in (U(x_0) \cap D) \setminus \{x_0\}$

1. **uniqueness of the limit**: if $\lim_{x\to x_0} f(x) = \ell_1$ and $\lim_{x\to x_0} f(x) = \ell_2$ then $\ell_1 = \ell_2$
2. **sign-preservation**: if $\lim_{x\to x_0} f(x) = \ell > 0$ then $f(x) > 0$ eventually as $x \to x_0$

Limit operations

Let $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$, $g : B \subseteq \mathbb{R}^n \to \mathbb{R}$, and let x_0 be an accumulation point of $A \cap B$, with $\lim_{x\to x_0} f(x) = \ell_1 \in \mathbb{R}$ and $\lim_{x\to x_0} g(x) = \ell_2 \in \mathbb{R}$

3. $\lim_{x\to x_0} cf(x) = c\ell_1,\ \forall\ c \in \mathbb{R}$
4. $\lim_{x\to x_0} f(x) + g(x) = \ell_1 + \ell_2$
5. $\lim_{x\to x_0} f(x)g(x) = \ell_1\ell_2$
6. $\lim_{x\to x_0} 1/g(x) = 1/\ell_2$ if $\ell_2 \neq 0$
7. $\lim_{x\to x_0} f(x)/g(x) = \ell_1/\ell_2$ if $\ell_2 \neq 0$

8. if $\lim_{x\to x_0} f(x) = +\infty$ and $g(x)$ is eventually bounded below as $x \to x_0$, then $\lim_{x\to x_0} f(x) + g(x) = +\infty$
9. if $\lim_{x\to x_0} f(x) = +\infty$ and $g(x) > \ell > 0$ eventually as $x \to x_0$, then $\displaystyle\lim_{x\to x_0} f(x)g(x) = +\infty$
10. if $\lim_{x\to x_0} f(x) = 0$ and $g(x)$ is eventually bounded as $x \to x_0$, then $\lim_{x\to x_0} f(x)g(x) = 0$
11. if $\lim_{x\to x_0} f(x) = 0^+$ then $\lim_{x\to x_0} 1/f(x) = +\infty$
12. if $\lim_{x\to x_0} f(x) = +\infty$ then $\lim_{x\to x_0} 1/f(x) = 0^+$

Properties of limits

Let $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$, $g : B \subseteq \mathbb{R}^n \to \mathbb{R}$, $h : C \subseteq \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n_*$ be an accumulation point of $A \cap B \cap C$

⑬ **squeeze theorem** (or sandwich theorem): if $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} g(x) = \ell$ and $f(x) \leq h(x) \leq g(x)$ eventually as $x \to x_0$, then $\lim\limits_{x \to x_0} h(x) = \ell$

Let $f\colon X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, $g : Y \to \mathbb{R}$, $Y \subseteq \mathbb{R}^m$, with X and Y such that $f(x) \in Y$ for every $x \in X$, and let $x_0 \in \mathbb{R}^n_*$ and $\ell \in \mathbb{R}^m_*$ be accumulation points of X and Y.

⑭ **limit of a composite function**: if $\lim\limits_{x \to x_0} f(x) = y_0$ and $\lim\limits_{y \to y_0} g(y) = \ell$ and $f(x) \neq y_0$ eventually as $x \to x_0$, then $\lim\limits_{x \to x_0} g(f(x)) = \ell$

Sequences in $\mathbb{R}$

Definition

A sequence of numbers, or in $\mathbb{R}$*, is a function* $f\colon \mathbb{N} \to \mathbb{R}$ *(or* $f\colon \{k \in \mathbb{N} : k \geq \bar{k}\} \to \mathbb{R}$*) that assigns to every non-negative integer* $k \in \mathbb{N}$ *a number*

$$f(k) = x(k) = x_k \in \mathbb{R}$$

A sequence in $\mathbb{R}$ *is denoted* $\{x_k\}_{k\in\mathbb{N}}$, $\{x_k\}_{\mathbb{N}}$, *or* $\{x_k\}$

Example

$\{1/k\}$, $\{k^2/(k+1)\}$

Sequences in $\mathbb{R}^n$

Definition

A sequence of ***vectors****, or in $\mathbb{R}^n$, is a function $f\colon \mathbb{N} \to \mathbb{R}^n$ (or $f\colon \{k \in \mathbb{N} : k \geq \bar{k}\} \to \mathbb{R}^n$) that assigns to every non-negative integer $k \in \mathbb{N}$ a vector*

$$f(k) = x(k) = x_k \in \mathbb{R}^n$$

A sequence in $\mathbb{R}^n$ is denoted $\{x_k\}_{k\in\mathbb{N}}$, $\{x_k\}_{\mathbb{N}}$, or $\{x_k\}$

Example

$$\left\{\begin{pmatrix} k \\ k+3 \end{pmatrix}\right\}, \quad \left\{\begin{pmatrix} \frac{1}{k} \\ 1+\sin k \end{pmatrix}\right\}$$

Sequences in $\mathbb{R}^n$

1. $\{x_k\}$ is **convergent** if there exists $\bar{x} \in \mathbb{R}^n$ such that

$$\lim_{k \to \infty} x_k = \bar{x}$$

2. $\{x_k\}$ is **divergent** if

$$\lim_{k \to \infty} x_k = \infty \in \mathbb{R}^n_\star$$

3. $\{x_k\}$ is **irregular** if it is neither convergent nor divergent
4. A sequence in $\mathbb{R}^n$, $\{x_k\}$, is bounded if there exists a constant $M > 0$ such that

$$\|x_k\| \leq M, \quad \forall \ k$$

Proposition

A sequence $\{x_k\}$ is bounded if and only if the sequences defined by each individual component are bounded

Subsequences in $\mathbb{R}^n$

Let $\{x_k\}$ be a sequence. The sequence obtained from $\{x_k\}$ by selecting the elements x_k corresponding to an infinite set of indices K is a subsequence

Example

Let $\{x_k\}$ be defined as

$$x_k = \begin{pmatrix} 1/k \\ (-1)^k \end{pmatrix}$$

The sequence obtained by selecting elements corresponding to even values of the index k is a subsequence

Subsequences in $\mathbb{R}^n$

Formally

Definition

Given a sequence $\{x_k\}$ and a strictly increasing sequence of integers $i_k \in \mathbb{N}$, the sequence $\{y_k\}$ defined as

$$y_k = x_{i_k}$$

is a subsequence of $\{x_k\}$

The set $K = \{i_k, k \in \mathbb{N}\}$ is an infinite subset of $\mathbb{N}$. A subsequence is very often denoted as

$$\{x_k, k \in K \subseteq \mathbb{N}\}$$

$$\{x_k\}_{k \in K}$$

$$\{x_k\}_K$$

Sequences in $\mathbb{R}^n$

1. $\{x_k\}$ is **convergent** if there exists $\bar{x} \in \mathbb{R}^n$ such that
$$\lim_{k \to \infty} x_k = \bar{x}$$
2. $\{x_k\}$ is **divergent** if
$$\lim_{k \to \infty} x_k = \infty \in \mathbb{R}^n_\star$$
3. $\{x_k\}$ is **irregular** if it is neither convergent nor divergent
4. A sequence in $\mathbb{R}^n$, $\{x_k\}$, is bounded if there exists a constant $M > 0$ such that
$$\|x_k\| \leq M, \quad \forall\ k$$
5. $\{x_k\}$ is **unbounded** if there exists a set of indices $K \subseteq \mathbb{N}$ such that the subsequence $\{x_k\}_{k \in K}$ is divergent.

Recursive sequences

Example in $\mathbb{R}$: Given $f\colon \mathbb{R} \to \mathbb{R}$, we want to find a point x^* such that $f(x^*) = 0$

Suppose

1. that f is differentiable on $\mathbb{R}$;
2. that we have x_0 such that $f(x_0) \neq 0$, $f'(x_0) \neq 0$

equation of the tangent line to the graph of f at $(x_0, f(x_0))$

$$y - f(x_0) = f'(x_0)(x - x_0)$$

we find the point where the line crosses the horizontal axis

$$x = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Recursive sequences

In general we can define a recursive sequence as follows. Given m vectors $a_i \in \mathbb{R}^n$, $i = 0, \dots, m-1$, and a function $f \colon \underbrace{\mathbb{R}}_{\to k} \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{m \text{ times}} \to \mathbb{R}^n$,

we define

$$x_k = \begin{cases} a_k, & \text{if } k < m, \\ f\big(k, x_{k-1}, \dots, x_{k-m}\big), & \text{if } k \geq m. \end{cases}$$

The first elements of the sequence are:

$$a_0,\ a_1,\ \dots,\ a_{m-1},\ \underbrace{x_m}_{f\left(m, a_{m-1}, \dots, a_0\right)},\ \underbrace{x_{m+1}}_{f\left(m+1, x_m, a_{m-1}, \dots, a_1\right)},\ \dots$$

Classification of sequences

1. $\{x_k\}$ **convergent** when (by def.) there exists $\bar{x} \in \mathbb{R}^n$: $\lim_{k\to\infty} x_k = \bar{x}$
2. $\{x_k\}$ **divergent** when (by def.) $\lim_{k\to\infty} x_k = \infty$, i.e. $\lim_{k\to\infty} \|x_k\| = +\infty$
3. $\{x_k\}$ **irregular** when (by def.) it is neither convergent nor divergent
4. $\{x_k\}$ **bounded** when (by def.) there exists $M > 0$: $\|x_k\| < M$ for every k
5. $\{x_k\}$ **unbounded** when (by def.) there exists $K \subseteq \mathbb{N}$: $\{x_k\}_K$ is divergent

Properties of sequences

1. $\{x_k\}$ **convergent** $\Rightarrow$ $\{x_k\}$ **bounded**
2. $\{x_k\}$ **bounded** $\Rightarrow$ there exists K: $\{x_k\}_K$ **convergent**
3. $\{x_k\}$ **convergent** $\Leftrightarrow$ for every K, $\{x_k\}_K$ **convergent**

$$\lim_{k\to\infty} x_k = \bar{x} \quad \Leftrightarrow \quad \lim_{k\to\infty, k\in K} x_k = \bar{x}, \qquad \forall\ K \subseteq \mathbb{N}$$

4. $\{x_k\} \subseteq C$, C **compact** $\Rightarrow$ there exists K: $\{x_k\}_K$ **convergent**
5. (Bridge theorem) $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$, x_0 accumulation point of X

$$\lim_{x\to x_0} f(x) = \ell \quad \Leftrightarrow \quad \lim_{k\to\infty} f(x_k) = \ell$$

for every $\{x_k\} \subseteq X \setminus \{x_0\}$: $\lim_{k\to\infty} x_k = x_0$

Definition

Definition

Let $f\colon E \to \mathbb{R}$, $E \subseteq \mathbb{R}^n$. f is continuous at $x_0 \in E$ when

- *x_0 is an* **isolated point** *of E, or*
- *x_0 is an* **accumulation point** *of E and*

$$\lim_{x \to x_0} f(x) = f(x_0).$$

that is

1. *the limit exists*
2. *the limit equals $f(x_0)$*

Definition

f is continuous on $X \subseteq E$ when it is continuous at every point of X

What does it mean?

Being able to say that a function $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ is continuous at an accumulation point $x_0 \in D$ of D means that

> *f takes values* **arbitrarily** *close to $f(x_0)$ (i.e., as close as we wish) for points x that are* **sufficiently** *close to x_0*

Operations that preserve continuity

Let $f\colon X \subseteq \mathbb{R}^n$, $g\colon Y \subseteq \mathbb{R}^n \to \mathbb{R}$ be continuous at $x_0 \in X \cap Y$. Then the following are also continuous at x_0

1. the function $f \pm g$
2. the function αf, $\alpha \in \mathbb{R}$
3. the function $f \cdot g$
4. the function f/g if $g(x_0) \neq 0$
5. the function $|f|$

Let $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ be continuous at $x_0 \in X$ and $g\colon Y \subseteq \mathbb{R} \to \mathbb{R}$ be continuous at $y = f(x_0) \in Y$. Then

6. the composite function $f(g(x))$ is continuous at x_0

Linear function

Definition

A function $f\colon \mathbb{R}^n \to \mathbb{R}$ *is called* **linear** *when*

1. *for every* $x, y \in \mathbb{R}^n$, $f(x+y) = f(x) + f(y)$
2. *for every* $x \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$, $f(\alpha x) = \alpha f(x)$

A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is linear $\Leftrightarrow$ there exist $c_1, c_2, \ldots, c_n \in \mathbb{R}$ such that

$$f(x) = c_1x_1 + c_2x_2 + \cdots + c_nx_n = c^\top x$$

Definition

A function $f\colon \mathbb{R}^n \to \mathbb{R}$ *is called* **affine** *when there exist* $c \in \mathbb{R}^n$ *and* $d \in \mathbb{R}$ *such that*

$$f(x) = c^\top x + d$$

Quadratic forms

- A quadratic form in one variable is a homogeneous polynomial of degree two: $q(x) = ax^2$
- A quadratic form in two variables is a homogeneous polynomial of degree two in two variables: $q(x, y) = ax^2 + bxy + cy^2$

$$\begin{aligned} q(x, y) &= (ax + by)x + (0x + cy)y = (ax + by,\ 0x + cy)\begin{pmatrix} x \\ y \end{pmatrix} \\ &= (x,\ y)\begin{pmatrix} a & 0 \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (x,\ y)\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= (x,\ y)\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Quadratic functions

Definition

A quadratic function is (by def.) the sum of a quadratic form and an affine function

$$q(x) = \underbrace{x^\top Q x}_{\textit{quadratic form}} + \underbrace{c^\top x + d}_{\textit{affine function}}$$

with Q not identically zero

N.B. one can always assume Q to be a symmetric matrix. If it were not, we can replace Q with $(Q + Q^\top)/2$, since

$$x^\top Q x = x^\top \left(\frac{1}{2}(Q + Q^\top) \right) x$$

CAUTION in general $Q \neq (Q + Q^\top)/2$

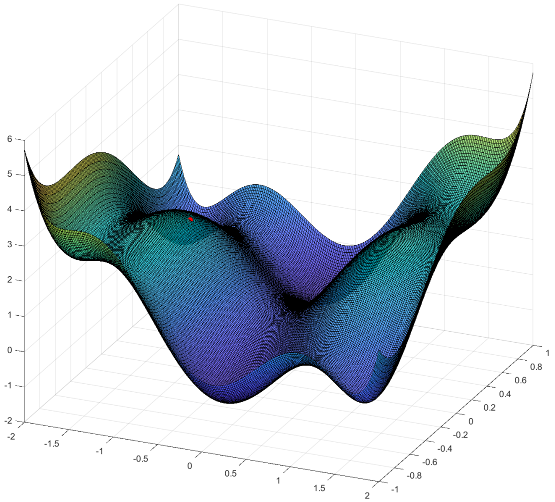
Restrictions of f

Given the function

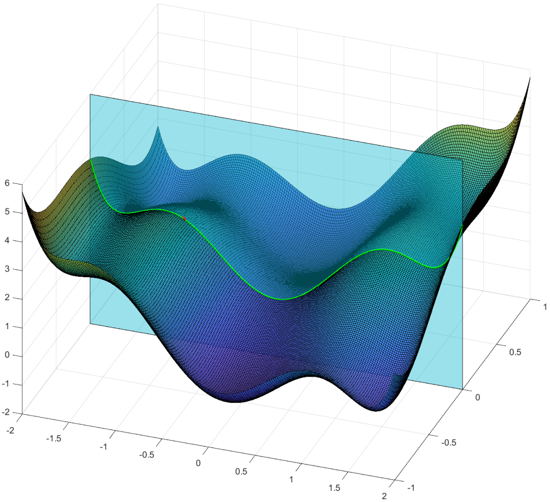
$$f(x,y) = (4 - 2.1x^2 + \frac{1}{3}x^4)x^2 + xy + (-4 + 4y^2)y^2$$

- Determine whether and why it is continuous on $\mathbb{R}^2$
- write the expression of the restriction of f to the set $A = \{(x,y) \in \mathbb{R}^2 : \; y = 0\}$
- write the expression of the restriction of f to the set $A = \{(x,y) \in \mathbb{R}^2 : \; x = -1\}$
- write the expression of the restriction of f to the set $A = \{(x,y) \in \mathbb{R}^2 : \; x = -1 + 2t, y = t\}$

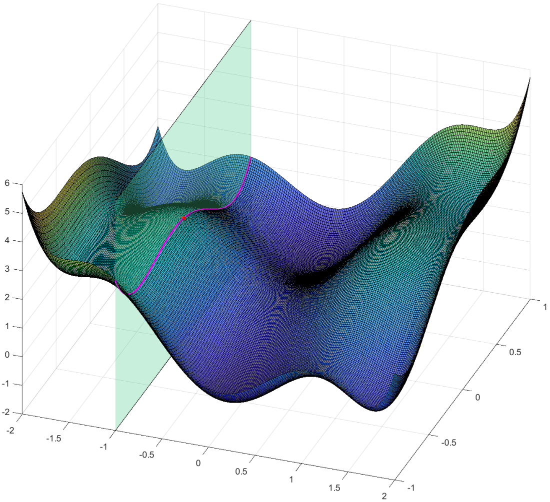
Graph of f



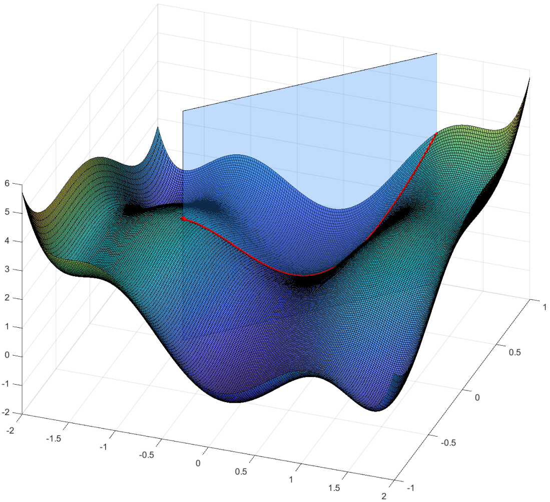
Restriction of f to $y = 0$



Restriction of f to $x = -1$



Restriction of f to $x = -1 + 2t, y = t$



Differentiable function and derivative

Given $f\colon \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}$

Definition

f is differentiable at x_0 if the following limit exists and is finite

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \ell$$

the value of the limit (if it exists and is finite) is called the derivative of f at x_0, $f'(x_0)$

Differentiable function

Given $f\colon \mathbb{R} \to \mathbb{R}$ and $x_0 \in \mathbb{R}$

Definition

f is differentiable at x_0 when there exists a constant $a \in \mathbb{R}$ such that, as $x \to x_0$

$$f(x) = f(x_0) + a(x - x_0) + o(x - x_0)$$

where

$$\lim_{x \to x_0} \frac{o(x - x_0)}{x - x_0} = 0$$

N.B. we can approximate f with the linear function $f(x_0) + a(x - x_0)$ to arbitrary precision, provided that x is sufficiently close to x_0

Equivalence in $\mathbb{R}$

Theorem

Let $f : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}$. f is differentiable at x_0 **if and only if** *f is (Fréchet) differentiable at x_0*

Differentiable function

$f : \mathbb{R} \to \mathbb{R}$ differentiable at x_0.
The graph of the linear function $f(x_0) + f'(x_0)(x - x_0)$ is the set

$$G_t = \{(x, y) : \; y = f(x_0) + f'(x_0)(x - x_0)\}$$

and is called the **tangent line** to the graph of f at the point $(x_0, f(x_0))$

The vector $(-f'(x_0), 1)$ is the **normal vector to the tangent** (and to the graph of f) at x_0

What does it mean?

Suppose we have a function $f\colon \mathbb{R} \to \mathbb{R}$ differentiable at a point x_0. What do we know about f?

1. f is continuous at x_0
2. f can be approximated "near" x_0 by a linear function
3. the tangent to the graph of f at $(x_0, f(x_0))$ exists, i.e., the graph is "smooth" at x_0
    - it certainly has no jumps (because f is continuous)
    - it has no "corners" where the tangent would not be defined

Example

In $\mathbb{R}^2$ consider the function

$$f(x_1, x_2) = x_1^2 + x_2^2$$

and

$$P_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad d_1 = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \quad d_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

- $A = \{x \in \mathbb{R}^2 : x = x(t) = P_1 + td_1, t \geq 0\}$
- $B = \{x \in \mathbb{R}^2 : x = x(t) = P_2 + td_2, t \geq 0\}$

$$f|_A(x) = \phi_1(t), \qquad f|_B(x) = \phi_2(t)$$

Compute:

$$\lim_{t \to 0^+} \frac{\phi_1(t) - \phi_1(0)}{t}, \qquad \lim_{t \to 0^+} \frac{\phi_2(t) - \phi_2(0)}{t}$$

Directional derivative

Congratulations — you have just computed your first directional derivative of a function of two variables!!!

Definition

Given $f : \mathbb{R}^n \to \mathbb{R}$, we say that f admits a **directional derivative** *at $x \in \mathbb{R}^n$ along $d \in \mathbb{R}^n$ when the following difference quotient limit exists and is finite*

$$\lim_{t \to 0^+} \frac{f(x+td) - f(x)}{t} = Df(x, d)$$

Example

Consider the function $f\colon \mathbb{R}^3 \to \mathbb{R}$

$$f(x_1, x_2, x_3) = x_1^2 + 3x_1x_3 + x_2x_3^2$$

Given $x \in \mathbb{R}^3$, consider the set

$$A = \{(y_1, y_2, y_3) : y_1 = x_1, y_2 = x_2, y_3 = x_3 + h\}$$

Compute

$$f|_A(x) = \phi(h), \qquad \lim_{h \to 0} \frac{\phi(h) - \phi(0)}{h}$$

Partial derivative

Definition

Given $f\colon \mathbb{R}^n \to \mathbb{R}$, we say that f admits a **partial derivative** *at $x \in \mathbb{R}^n$ with respect to the variable x_i when the following difference quotient limit exists and is finite*

$$\lim_{h \to 0} \frac{f(x_1, x_2, \dots, x_i + h, x_{i+1}, \dots, x_n) - f(x)}{h} = \frac{\partial f(x)}{\partial x_i}$$

The vector $\nabla f(x)$

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$. f is **differentiable** *at x when it admits partial derivatives with respect to all variables.*

If f is differentiable at x, the column vector of partial derivatives is denoted by $\nabla f(x)$

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

Connection between directional derivative and partial derivative

Note that, given $x \in \mathbb{R}^n$ and $e_i \in \mathbb{R}^n$, we have

$$x + he_i = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} + h \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i + h \\ \vdots \\ x_n \end{pmatrix}$$

therefore, since $f(x_1, \dots, x_i + h, \dots, x_n) = f(x + he_i)$, if $\partial f(x)/\partial x_i$ exists

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h} = \lim_{t \to 0^+} \frac{f(x + te_i) - f(x)}{t} = Df(x, e_i)$$

Example 1

$$f(x, y) = \begin{cases} 1 & \text{if } xy \neq 0 \\ 0 & \text{if } xy = 0 \end{cases}$$

1. f is **differentiable** at $(0, 0)$, however
2. f is NOT continuous at $(0, 0)$, the limit

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

   does NOT exist
3. f cannot be approximated by a linear function “near” the origin

Time to bring out the heavy artillery

$$f(x,y) = \begin{cases} 1 & \text{if } xy \neq 0 \\ 0 & \text{if } xy = 0 \end{cases}$$

Does f admit directional derivatives at $(0,0)$? **NO**

Try choosing $d \neq \pm e_i$ and computing

$$\lim_{t \to 0^+} \frac{f(0+td) - f(0)}{t}$$

Example 2

$$f(x,y) = \begin{cases} 1 & \text{if } x^2 < y < 2x^2 \\ 0 & \text{otherwise} \end{cases}$$

Let $d \in \mathbb{R}^2$, $d_2 > 0$, $d_1 \neq 0$. Consider the half-line

$$r^+ = \{(x,y) \in \mathbb{R}^2 : x = td_1,\ \ y = td_2,\ \ t \geq 0\}$$

$$r^+ \cap G_{2x^2} = \{(x,y) \in \mathbb{R}^2 : x = td_1, y = td_2, td_2 = 2t^2 d_1^2\} = \left\{(0,0),\ \left(\frac{d_2}{2d_1}, \frac{d_2^2}{2d_1^2}\right)\right\}$$

$td_2 > 2t^2 d_1^2$ for $t > 0$, when $t \in (0, d_2/2d_1^2)$. Therefore, if we consider the set $A = \{(x,y) \in \mathbb{R}^2 : x = td_1,\ \ y = td_2,\ \ t \in (0, d_2/2d_1^2)\}$

$$f|_A = 0$$

Then $Df(0,d) = 0$.

In conclusion...

- the existence of partial derivatives at a point, and
- the existence of directional derivatives at a point along every direction

are NOT sufficient to say that f is "linearizable" in a neighborhood of that point

So what then?

Definition

$f : \mathbb{R}^n \to \mathbb{R}$ is **differentiable** *(in the strong sense) at x_0 when there exists a vector $g(x_0) \in \mathbb{R}^n$ such that*

$$\lim_{\|d\| \to 0} \frac{|f(x_0 + d) - f(x_0) - g^\top d|}{\|d\|} = 0$$

the vector $g(x_0)$ is called the **gradient** *of f at x_0*

Properties

$f : \mathbb{R}^n \to \mathbb{R}$ **differentiable** at x_0

1. if and only if for every $d \in \mathbb{R}^n$, $d \neq 0$

$$f(x_0 + d) = f(x_0) + g(x_0)^\top d + o(\|d\|) \quad \text{when } \|d\| \to 0$$

2. f is continuous at x_0, that is

$$\lim_{x \to x_0} f(x) = f(x_0)$$

3. $f(x_0 + td) - f(x_0) = tg^\top d + o(|t|\|d\|)$, therefore

$$Df(x_0, d) = g^\top d$$

4. $Df(x_0, e_i) = g_i$, $Df(x_0, -e_i) = -g_i$, hence f is differentiable at x_0 and $\nabla f(x_0) = g$

Tangent hyperplane

Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is **differentiable** at x_0

Definition

The **tangent hyperplane** *to the graph of f at the point* $(x_0, f(x_0))$ *exists and is*

$$T(x_0) = \{(x, y):\ \ y = f(x_0) + \nabla f(x_0)^\top (x - x_0)\}$$

The vector $v = (-\nabla f(x_0)^\top, 1)^\top$ *is the* **normal vector** *to the tangent hyperplane and to the graph of f*

A function that is continuous and differentiable but not (strongly) differentiable

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be defined as:

$$f(x,y) = \begin{cases} \dfrac{x^2 y}{x^2+y^2} & \text{if } x^2+y^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Show that f is NOT differentiable at $(0,0)^\top$

How to determine whether f is differentiable at x_0

Do we always have to resort to the definition to determine whether a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable at x_0?

The answer is **NO**, not always

We have two tools

1. operations that **preserve** differentiability
2. the **total differential** theorem

Sum and product of functions

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be differentiable at x_0

1. $\boxed{h(x) = f(x) + g(x)}$ is differentiable at x_0 and

$$\nabla h(x_0) = \nabla f(x_0) + \nabla g(x_0)$$

2. $\boxed{h(x) = f(x) \cdot g(x)}$ is differentiable at x_0 and

$$\nabla h(x_0) = g(x_0)\nabla f(x_0) + f(x_0)\nabla g(x_0)$$

Composite function

Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. Let g be differentiable at x_0 and f be differentiable at $g(x_0)$

3. $\boxed{h(x) = f(g(x))}$ is differentiable at x_0 and

$$\nabla h(x_0) = f'(g(x_0))\nabla g(x_0)$$

Reciprocal and quotient

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ and $g\colon \mathbb{R}^n \to \mathbb{R}$ be differentiable at x_0 and such that $g(x_0) \neq 0$

4. $\boxed{h(x) = 1/g(x)}$ is differentiable at x_0 and

$$\nabla h(x_0) = -\frac{1}{g(x_0)^2}\nabla g(x_0)$$

5. $\boxed{h(x) = f(x)/g(x)}$ is differentiable at x_0 and

$$\nabla h(x_0) = \frac{g(x_0)\nabla f(x_0) - f(x_0)\nabla g(x_0)}{g(x_0)^2}$$

Total differential

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$. *Then:*

1. *If* f *is differentiable at* x **then** f *is continuous and differentiable at* x *(the vector* $\nabla f(x)$ *exists and coincides with the gradient of* f *at* x*)*
2. *if* f *is differentiable at* x *and* $\nabla f(x)$ *is continuous* **then** f *is (strongly) differentiable at* x

Jacobian matrix

Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be a vector of functions

Suppose all functions $g_i(x)$ are differentiable, i.e., the following exist

$$\nabla g_1(x), \dots, \nabla g_m(x)$$

it is possible to organize all these partial derivatives into a single **Jacobian** matrix

Definition (Jacobian matrix)

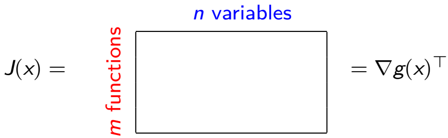
Let $g : \mathbb{R}^n \to \mathbb{R}^m$ and $x \in \mathbb{R}^n$. If the first-order partial derivatives $\partial g_i(x)/\partial x_j$ exist for $i = 1 \dots, m$ and $j = 1 \dots n$ at x, we define the Jacobian matrix *of g at x as the $m \times n$ matrix*

$$J(x) := \begin{pmatrix} \dfrac{\partial g_1(x)}{\partial x_1} & \cdots & \dfrac{\partial g_1(x)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial g_m(x)}{\partial x_1} & \cdots & \dfrac{\partial g_m(x)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \nabla g_1(x)^\top \\ \nabla g_2(x)^\top \\ \vdots \\ \nabla g_m(x)^\top \end{pmatrix} = \nabla g(x)^\top$$

Jacobian matrix

$$g : \mathbb{R}^n \to \mathbb{R}^m$$

n **variables** for *m* **functions**



Recall

For functions $f\colon \mathbb{R}^n \to \mathbb{R}$

1. if f is differentiable: $\nabla f(x)$ is the vector of partial derivatives
2. if f is (strongly) differentiable: gradient $g(x)$ such that

$$\lim_{\|d\| \to 0} \frac{|f(x+d) - f(x) - g(x)^\top d|}{\|d\|} = 0$$

Differentiability of a vector of functions

Definition (First derivative of a vector of functions)

Let $g : \mathbb{R}^n \to \mathbb{R}^m$*. We say that* g *is differentiable at the point* $x \in \mathbb{R}^n$ *if there exists a matrix* $J(x)$ *such that, for every* d*,*

$$\lim_{\|d\| \to 0} \frac{\|g(x+d) - g(x) - J(x)d\|}{\|d\|} = 0.$$

The linear operator $J(x) : \mathbb{R}^n \to \mathbb{R}^m$ *is called the derivative of* g *at* x*.*

N.B. the mere existence of the Jacobian matrix at x does not imply differentiability

Differentiability of a vector of functions

The following holds (analogously to the case $m = 1$)

Proposition

Let $g : \mathbb{R}^n \to \mathbb{R}^m$ *and* $x \in \mathbb{R}^n$.

(i) *if* g *is differentiable at* x*, then* g *is continuous at* x*, the Jacobian matrix* $J(x)$ *exists, and* $J(x)$ *coincides with the derivative of* g *at* x

(ii) *if the Jacobian matrix* $J(x)$ *of* g *at* x *exists and* J *is continuous with respect to* x*, then* g *is differentiable at* x*, and the derivative of* g *at* x *coincides with* $J(x)$.

Chain rule

Let

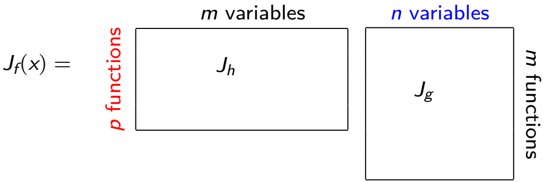
- $g : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at x
- $h : \mathbb{R}^m \to \mathbb{R}^p$ be differentiable at $g(x)$

Then $f : \mathbb{R}^n \to \mathbb{R}^p$ defined by $f(x) = h(g(x))$ is differentiable at x and

$$\nabla f(x) = \nabla g(x) \nabla h(y)|_{y=g(x)}$$

N.B. recalling that $\nabla f(x) = J_f(x)^\top$, we have

$$J_f(x) = J_h(y)|_{y=g(x)} J_g(x)$$



Special case: $n = 1$, $p = 1$

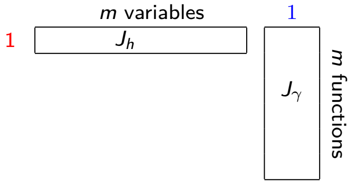
Let

- $\gamma : (a, b) \to \mathbb{R}^m$ be differentiable at $t_0 \in (a, b)$
- $h : \mathbb{R}^m \to \mathbb{R}$ be differentiable at $\gamma(t_0) \in \mathbb{R}^m$

Then $F : (a, b) \to \mathbb{R}$ defined by $F(t) = h(\gamma(t))$ is differentiable at t_0 and

$$F'(t) = \nabla\gamma(t)\nabla h(y)|_{y=\gamma(t)} = J_h(\gamma(t))J_\gamma(t)$$

$F'(t) =$



Differentiability criterion

A sufficient condition for $f\colon \mathbb{R}^n \to \mathbb{R}^m$ to be differentiable at x is that

- f is **differentiable** (in the weak sense) at x
- $\nabla f(x)$ is **continuous** at x

1. a function f **continuous** at every $x \in X$ is said to be of **class 0** on X, $\boxed{f \in C^0(X)}$
2. a function f that satisfies the **sufficient condition for differentiability** at every $x \in X$ is said to be of **class 1** on X, $\boxed{f \in C^1(X)}$