

Calculo Diferencial

Contents

1	Topología en el espacio euclídeo	2
2	Aplicaciones diferenciables	18
2.1	Derivadas direccionales y gradiente	18
2.2	Aplicaciones diferenciables	20
2.3	Regla de la cadena	24
2.4	Teoremas del valor medio	25
2.5	Derivadas de orden superior. Teorema de Taylor	27
2.6	Extremos locales	31
2.7	Teoremas de la función inversa, implícita y de los multiplicadores de Lagrange	34

1 Topología en el espacio euclídeo

Definición 1.0.1 [Longitud o Norma euclídea]

Se denomina **longitud o norma euclídea** de un vector $\vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ al número real mayor o igual que cero definido por

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Definición 1.0.2 [Distancia euclídea]

Se llama **distancia euclídea** entre dos vectores $\vec{x} = (x_1, x_2, \dots, x_n)$ y $\vec{y} = (y_1, y_2, \dots, y_n)$ al número real mayor o igual que 0 definido por:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{\frac{1}{2}}$$

Definición 1.0.3 [Producto escalar euclídeo]

Se llama **producto escalar euclídeo** entre dos vectores $\vec{x} = (x_1, x_2, \dots, x_n)$ y $\vec{y} = (y_1, y_2, \dots, y_n)$ al número real, no necesariamente positivo, definido por:

$$\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i.$$

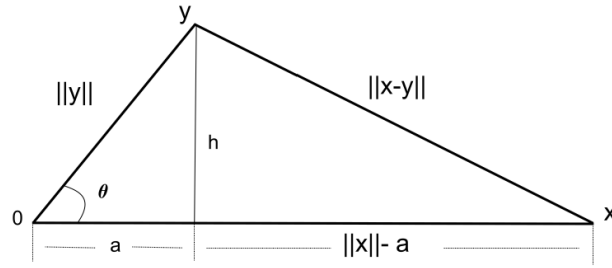
Teorema 1.0.1

1. $\langle \vec{x}, \vec{y} \rangle \geq 0 \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n.$
2. $\langle \vec{x}, \vec{y} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0} \text{ o } \vec{y} = \vec{0}.$
3. $\forall \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ se cumple que $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle.$
4. $\langle \alpha \vec{x}, \vec{y} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle$ para todo $\alpha \in \mathbb{R}.$
5. $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle.$

Teorema 1.0.2

Para cualesquiera $x, y \in \mathbb{R}^2$ se verifica que $\langle x, y \rangle = \|\vec{x}\| \|\vec{y}\| \cos(\theta)$, donde θ es el ángulo entre los vectores \vec{x} y \vec{y} .

Demostración. Dados dos vectores x e y de \mathbb{R}^2 , que supondremos distintos de 0 (pues si uno de ellos es 0 el resultado es inmediato), consideremos el triángulo de vértices 0, x , y :



Utilizando trigonometría elemental, tenemos que:

$$\cos \theta = \frac{a}{\|y\|}$$

Además, usando el teorema de Pitágoras, tenemos que:

$$\|y\|^2 = a^2 + h^2 \implies \|y\|^2 - a^2 = h^2 = \|x - y\|^2 - (\|x\| - a)^2$$

Con lo que:

$$\|x - y\|^2 = \|y\|^2 - a^2 + \|x\|^2 - 2a\|x\| + a^2 = \|y\|^2 + \|x\|^2 - 2a\|x\|$$

Usando que $a = \|y\| \cos \theta$, obtenemos:

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \cos \theta$$

Si ahora usamos las propiedades del producto interior, obtenemos que:

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$$

De donde se deduce, teniendo en cuenta el valor previamente obtenido de $\|x - y\|^2$, que:

$$\langle x, y \rangle = \|x\|\|y\| \cos \theta$$

□

Definición 1.0.4 [Vectores ortogonales]

Se dice que dos vectores \vec{x} y \vec{y} son **ortogonales** si $\langle \vec{x}, \vec{y} \rangle = 0$.

Proposición 1.0.1 [Propiedades de la norma euclídea]

1. $\|\vec{x}\| \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n$.
2. $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$.
3. $\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|$ para todo $\alpha \in \mathbb{R}$.
4. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n$ (desigualdad triangular).

Teorema 1.0.3 [Desigualdad de Cauchy-Schwarz]

Sea $\vec{x}, \vec{y} \in \mathbb{R}^n$. Entonces se cumple que:

$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|$$

Equivalentemente

$$\left\| \sum_{i=1}^n x_i y_i \right\| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}$$

Demostración. Fijemos \vec{x} y $\vec{y} \in \mathbb{R}^n$. Para cada $\alpha \in \mathbb{R}$ se tiene que

$$\langle \alpha \vec{x} + \vec{y}, \alpha \vec{x} + \vec{y} \rangle = \alpha^2 \langle \vec{x}, \vec{x} \rangle + 2\alpha \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle \geq 0$$

Si tomamos $A = \langle \vec{x}, \vec{x} \rangle$, $B = 2\langle \vec{x}, \vec{y} \rangle$ y $C = \langle \vec{y}, \vec{y} \rangle$, tenemos que:

$$A\alpha^2 + B\alpha + C \geq 0 \quad \forall \alpha \in \mathbb{R}$$

Entonces podemos distinguir dos casos:

1. Si $A = 0$, entonces $\vec{x} = \vec{0}$ y la desigualdad es trivial.
2. Si $A > 0$, entonces la desigualdad anterior es una ecuación cuadrática en α , y por las propiedades del producto escalar es necesario que su discriminante sea no positivo, pues de lo contrario tendría dos raíces reales distintas y entonces la ecuación tomaría algún valor negativo

$$\implies D = B^2 - 4AC \leq 0 \iff B^2 \leq 4AC \iff 4\langle \vec{x}, \vec{y} \rangle^2 \leq 4\langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle = 4\|\vec{x}\|^2 \|\vec{y}\|^2$$

□

Proposición 1.0.2 [Propiedades de la distancia euclídea]

1. $d(\vec{x}, \vec{y}) \geq 0 \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n$.
2. $d(\vec{x}, \vec{y}) = 0 \iff \vec{x} = \vec{y}$.
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$.
4. $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ (desigualdad triangular).

Definición 1.0.5 [Métrica]

Se llama **métrica** sobre un conjunto arbitrario M a cualquier aplicación $d : M \times M \rightarrow \mathbb{R}$ que cumple las siguientes propiedades:

1. $d(x, y) \geq 0 \quad \forall x, y \in M$.
2. $d(x, y) = 0 \iff x = y$.
3. $d(x, y) = d(y, x)$.

4. $d(x, z) \leq d(x, y) + d(y, z)$ (desigualdad triangular).

Definición 1.0.6 [Espacio métrico]

Se llama **espacio métrico** a un par (M, d) donde M es un conjunto no vacío y d es una métrica sobre M .

Ejemplo

Vemos algunos ejemplos de métricas:

1. La métrica euclídea en \mathbb{R}^n
2. $d_1(x, y) = \sum_{i=1}^n |x_i - y_i|$
3. $d_\infty(x, y) = \max_{i=1, \dots, n} |x_i - y_i|$
4. $d(f, g) = \int_a^b |f(x) - g(x)| dx$ para funciones $f, g : [a, b] \rightarrow \mathbb{R}$.
5. $d_\infty(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$ para funciones $f, g : [a, b] \rightarrow \mathbb{R}$.
6. $d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y \end{cases}$, que se conoce como la **métrica discreta**.

Definición 1.0.7 [Diámetro]

Se llama **diámetro** de un subconjunto S de un espacio métrico (M, d) a

$$\text{diam}(S) = \sup\{d(x, y) \mid x, y \in S\}$$

si el conjunto de números reales $\{d(x, y) : x, y \in S\}$ es acotado superiormente y se define $\text{diam}(S) = +\infty$ en caso contrario. Cuando el diámetro es infinito se dice que el conjunto no es **acotado**.

Definición 1.0.8 [Norma]

Sea V un espacio vectorial sobre \mathbb{R} . Se llama **norma** en V a toda aplicación $\|\cdot\| : V \rightarrow \mathbb{R}$ que cumple las siguientes propiedades:

1. $\|\vec{x}\| \geq 0 \quad \forall \vec{x} \in V$.
2. $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$.
3. $\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|$ para todo $\alpha \in \mathbb{R}$.
4. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ (desigualdad triangular).

Ejemplo

1. $\|\vec{x}\| = |x|$
2. $\|\vec{x}\|_2 = \left(\sum_{j=1}^n x_j^2\right)^{\frac{1}{2}}$ (norma euclídea).

3. $\|\vec{x}\|_1 = \sum_{j=1}^n |x_j|$ (norma l^1).
4. $\|\vec{x}\|_\infty = \max_{j=1,\dots,n} |x_j|$ (norma l^∞).
5. $\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$ para funciones $f : [a, b] \rightarrow \mathbb{R}$.

Definición 1.0.9 [Producto escalar o interior]

Llamaremos *producto escalar o producto interior* en V a toda aplicación $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ que cumple las siguientes propiedades:

1. $\langle \vec{x}, \vec{y} \rangle \geq 0 \quad \forall \vec{x}, \vec{y} \in V$.
2. $\langle \vec{x}, \vec{x} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0}$.
3. $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$.
4. $\langle \alpha \vec{x}, \vec{y} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle$ para todo $\alpha \in \mathbb{R}$.
5. $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$ para todo $\vec{x}, \vec{y}, \vec{z} \in V$.

Definición 1.0.10 [Igualdad del paralelogramo]

Sea una norma $\|\cdot\|$ en un espacio vectorial V . Se dice que la norma cumple la **igualdad del paralelogramo** si la norma procede de un producto escalar

$$\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2$$

Demostración.

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle + \langle \vec{x} - \vec{y}, \vec{x} - \vec{y} \rangle = \\ &= \|\vec{x}\|^2 + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \|\vec{y}\|^2 + \|\vec{x}\|^2 - \langle \vec{x}, \vec{y} \rangle - \langle \vec{y}, \vec{x} \rangle + \|\vec{y}\|^2 = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2 \end{aligned}$$

□

Definición 1.0.11 [Bola abierta]

Dados $x_0 \in \mathbb{R}^n$ y un número real $r > 0$, llamamos **bola abierta** de centro x_0 y radio r al conjunto

$$B(x_0, r) = \{x \in \mathbb{R}^n \mid d(x, x_0) < r\}$$

donde d es la métrica que se está considerando en \mathbb{R}^n .

Definición 1.0.12 [Conjunto abierto]

Se dice que un conjunto $A \subset \mathbb{R}^n$ es **abierto** si para todo punto $x_0 \in A$ existe un número real $r > 0$ tal que $B(x_0, r) \subset A$.

Proposición 1.0.3 [Propiedades de los conjuntos abiertos]

1. El conjunto vacío y el espacio euclídeo \mathbb{R}^n son abiertos.
2. La unión de abiertos es un abierto
3. La intersección finita de abiertos es un abierto.

Definición 1.0.13 [Punto abierto]

Se dice que un punto $x \in S \subset \mathbb{R}^n$ es un **punto abierto** de S si existe una bola abierta $B(x, r)$ tal que $B(x, r) \subset S$. Denotamos por S° al conjunto de los puntos abiertos de S .

Observación 1.0.1

S° puede ser vacío, por ejemplo si S es un subconjunto con un solo punto

Proposición 1.0.4 [Propiedades de los puntos abiertos]

1. S° es el mayor abierto contenido en S
2. S° es la unión de todos los abiertos contenidos en S .
3. S es abierto si y solo si $S = S^\circ$.

Demostración. 1. S° es abierto, pues dado $x \in S^\circ$, existe $r > 0$ tal que $B(x, r) \subset S$. Entonces sucede que $B(x, r) \subset S^\circ$, pues al ser $B(x, r)$ un abierto, entonces $B(x, r) = [B(x, r)]^\circ \subset S^\circ$. Por otra parte, si A es un abierto de \mathbb{R}^n contenido en S , entonces para todo punto de A hay una bola centrada en él contenida en A y por lo tanto en S , luego todos los puntos de A están en S°

2. Es claro que el mayor abierto contenido en S es la unión de todos los abiertos contenidos en S

3. Si S es abierto, entonces él es el mayor abierto contenido en S , luego $S = S^\circ$. Por otra parte, si $S = S^\circ$, entonces S es abierto, pues para todo punto $x \in S$, existe una bola abierta $B(x, r)$ tal que $B(x, r) \subset S$, luego S es abierto.

□

Definición 1.0.14 [Bola cerrada]

Dados $x_0 \in \mathbb{R}^n$ y un número real $r > 0$, llamamos **bola cerrada** de centro x_0 y radio r al conjunto

$$\overline{B}(x_0, r) = \{x \in \mathbb{R}^n \mid d(x, x_0) \leq r\}$$

Proposición 1.0.5

Sea $x_0 \in \mathbb{R}^n$ y $r > 0$. Entonces se cumple que:

$$[\overline{B}(x_0, r)]^\circ = B(x_0, r)$$

Demostración. “ \subset ”

Sea $x \in [\overline{B}(x_0, r)]^\circ \implies \exists r_x > 0 : B(x, r_x) \subset \overline{B}(x_0, r)$ tal que $\|x - x_0\| = r$

Sea $y = x + \frac{x - x_0}{\|x - x_0\|} \cdot \frac{1}{2}r_x \implies \|y - x\| = \frac{1}{2}r_x < r_x \implies y \in B(x, r_x)$

No obstante,

$$\|y - x_0\| = \|x - x_0\| \left(1 + \frac{1}{\|x - x_0\|} \frac{1}{2}r_x\right) > \|x - x_0\| = r$$

Por tanto llegamos a que ningún punto de la frontera de la bola cerrada puede estar en su interior. “ \supset ”

Esta inclusión es inmediata pues $B(x_0, r) \subset \overline{B}(x_0, r)$ y al ser $B(x_0, r) = [B(x_0, r)]^\circ$ resulta que $B(x_0, r) \subset [\overline{B}(x_0, r)]^\circ$.

□

Definición 1.0.15 [Conjunto cerrado]

Se dice que un subconjunto $C \subset \mathbb{R}^n$ es **cerrado** si su complementario (respecto a \mathbb{R}^n) es abierto.

Proposición 1.0.6 [Propiedades de los conjuntos cerrados]

1. El conjunto vacío y el espacio euclídeo \mathbb{R}^n son cerrados.
2. La intersección de cerrados es un cerrado.
3. La unión finita de cerrados es un cerrado.

Definición 1.0.16 [Punto de acumulación]

Se dice que un punto $x \in \mathbb{R}^n$ es un **punto de acumulación** de un subconjunto $S \subset \mathbb{R}^n$ si toda bola abierta centrada en x contiene algún punto de S distinto de x . Equivalentemente,

$$x \text{ es un punto de acumulación de } S \Leftrightarrow \forall r > 0, B(x, r) \cap (S \setminus \{x\}) \neq \emptyset$$

Al conjunto de puntos de acumulación de un conjunto se le suele denominar S'

Ejemplo

1. Todo punto del intervalo $[0, 1]$ es un punto de acumulación de $(0, 1)$.
2. El punto 0 es un punto de acumulación de $S = \{1/n \mid n \in \mathbb{N}\}$
3. $S = \mathbb{Q} \cap [0, 1] \implies S' = [0, 1]$.

Teorema 1.0.4

S es cerrado si y solo si $S' \subset S$.

Demostración. “ \Rightarrow ” Supongamos que S es cerrado. Entonces sea $x \in S' : x \notin S \Rightarrow$ cualquier bola de centro en x contiene puntos de $S \Rightarrow \mathbb{R}^n \setminus S$ no es abierto, luego llegamos a contradicción y que $x \in S$.

“ \Leftarrow ” Dado $x \in \mathbb{R}^n \setminus S \Rightarrow \exists r > 0 : B(x, r) \subset \mathbb{R}^n \setminus S$ (pues x no es punto de acumulación de S). Por tanto, $B(x, r) \subset \mathbb{R}^n \setminus S$, luego $\mathbb{R}^n \setminus S$ es abierto y por tanto S es cerrado. \square

Definición 1.0.17 [Punto adherente]

Se dice que un punto $x \in \mathbb{R}^n$ es un **punto adherente** de un subconjunto $S \subset \mathbb{R}^n$ si toda bola abierta centrada en x contiene algún punto de S . Equivalentemente,

$$x \text{ es un punto adherente de } S \Leftrightarrow \forall r > 0, B(x, r) \cap S \neq \emptyset$$

Definición 1.0.18 [Adherencia o clausura]

Se llama **adherencia** o **clausura** de un conjunto $S \subset \mathbb{R}^n$ al conjunto de sus puntos adherentes, denotado por \overline{S} .

Proposición 1.0.7 [Propiedades de la adherencia]

1. \overline{S} es cerrado y es el menor cerrado que contiene a S ($S \subset \overline{S}$)
2. \overline{S} es la intersección de todos los cerrados que contienen a S .
3. S es cerrado si y solo si $S = \overline{S}$.

Demostración. 1. Sea $x \in \mathbb{R}^n \setminus \overline{S} \Rightarrow \exists r > 0 : B(x, r) \cap S = \emptyset \Rightarrow \forall y \in B(x, r) : B(y, r - d(y, x)) \cap S = \emptyset \Rightarrow y \notin \overline{S} \Rightarrow B(x, r) \subset \mathbb{R}^n \setminus \overline{S}$.

Por otro lado, si C es un cerrado que contiene a S , entonces C debe contener a \overline{S} , pues C contiene a S y a S' debido a que $S' \subset C' \subset C$ por ser C cerrado.

2. Es claro que el menor cerrado que contiene a un conjunto es la intersección de todos los cerrados que lo contienen. Además, la intersección finita de cerrados es cerrada.

3. Si S es cerrado es claro que él mismo es el menor cerrado que lo contiene, luego $S = \overline{S}$. Recíprocamente, por lo visto en el primer apartado, si \overline{S} es cerrado, entonces si $S = \overline{S}$, entonces S es cerrado. \square

Definición 1.0.19 [Distancia]

Dados un conjunto $S \subset \mathbb{R}^n$ y un punto $x \in \mathbb{R}^n$, se define la **distancia** de x a S como

$$d(x, S) = \inf\{d(x, y) : y \in S\}$$

Teorema 1.0.5

Un punto $x \in \mathbb{R}^n$ es un punto adherente de un conjunto $S \subset \mathbb{R}^n$ si y solo si $d(x, S) = 0$.

Demostración. “ \rightarrow ” $x \in \bar{S} \implies \forall r > 0, \exists y \in S : d(x, y) < r \implies \inf\{d(x, y) : y \in S\} = 0$. Obsérvese que como $x \in S$ este ínfimo se alcanza y es mínimo.

“ \leftarrow ” El $\inf\{d(x, y) : y \in S\} = 0$ implica que $\forall \epsilon > 0, \exists y \in S : d(x, y) < \epsilon$. Entonces $B(x, \epsilon) \cap S \neq \emptyset$, luego x es un punto adherente de S . \square

Definición 1.0.20 [Punto frontera]

Se llama **punto frontera** de un conjunto $S \subset \mathbb{R}^n$ a todo punto $x \in \mathbb{R}^n$ tal que $\forall r > 0$, se cumple que:

$$B(x, r) \cap S \neq \emptyset \text{ y } B(x, r) \cap (\mathbb{R}^n \setminus S) \neq \emptyset$$

Equivalentemente, un punto $x \in \mathbb{R}^n$ es un punto frontera de S si y solo si $x \in \bar{S} \cap \overline{\mathbb{R}^n \setminus S}$.

Definición 1.0.21 [Punto aislado]

Se dice que un punto $x \in S$ es un **punto aislado** de S si existe un número real $r > 0$ tal que $B(x, r) \cap S = \{x\}$.

Definición 1.0.22 [Sucesión]

Se llama **sucesión** en \mathbb{R}^n a toda aplicación $x : \mathbb{N} \rightarrow \mathbb{R}^n$. Como se hace en el caso de sucesiones de números reales, se identificará x con la “tira” de sus valores $(x(1), x(2), \dots)$. Normalmente se escribirá x_k en vez de $x(k)$ y las sucesiones se representarán por $\{x_k\}_{k=1}^\infty$, $\{x_k\}$, $(x_k)_{k=0}^\infty$ o (x_k) . Obsérvese que cada x_k es un punto de \mathbb{R}^n y por lo tanto es de la forma $x_k = (x_{k1}, \dots, x_{kn})$.

Definición 1.0.23 [Convergencia de sucesiones]

Se dice que una sucesión $\{x_k\}$ **converge** a un punto $x_0 \in \mathbb{R}^n$ si la sucesión de números reales $\{d(x_k, x_0)\}$ converge a 0. Esto es,

$$\forall \epsilon > 0 \exists \nu \in \mathbb{N} \text{ tal que } k \geq \nu \implies d(x_k, x_0) < \epsilon.$$

Definición 1.0.24 [Subsucesión]

Se llama **subsucesión** de una sucesión $\{x_k\}$ a toda aplicación $m \in \mathbb{N} \rightarrow x_{k_m}$ es uno de los términos de la sucesión $\{x_k\}$ que verifique la condición de que la aplicación $m \in \mathbb{N} \rightarrow k_m$ sea estrictamente creciente.

Proposición 1.0.8

$$x \in \bar{S} \Leftrightarrow \exists \{x_k\} \subset S : x_k \rightarrow x$$

Demostración. “ \Rightarrow ” Sea $x \in \bar{S} \implies \forall r > 0, B(x, r) \cap S \neq \emptyset \implies \exists y : d(x, y) < r$. Si tomamos $r = \frac{1}{k} : k \in \mathbb{N}$, entonces $\exists y_k \in S : d(x, y_k) < \frac{1}{k}$. Por tanto, la sucesión $\{y_k\}$ converge a x .

“ \Leftarrow ” Sea $\{x_k\} \subset S$ tal que $x_k \rightarrow x$. Entonces, por definición de convergencia, para todo $\epsilon > 0 \exists k_0 : \forall k \geq k_0 \implies d(x_k, x) < \epsilon \iff x_k \in B(x, \epsilon)$ y como $x_k \in S$, entonces $B(x, \epsilon) \cap S \neq \emptyset$. Por tanto, $x \in \bar{S}$. \square

Teorema 1.0.6

Si una sucesión $\{x_k\}$ convergen a un punto $x_0 \in \mathbb{R}^n$, entonces toda subsucesión de $\{x_k\}$ también converge a x_0 .

Demostración. $\{x_k\} \rightarrow x_0 \implies \forall \epsilon \exists n \in \mathbb{N} : \forall k \geq n \implies d(x_k, x_0) < \epsilon \implies$ Sea $\{x_{k_m}\}$ una subsucesión de $\{x_k\}$, dado $\epsilon > 0$, existe $n \in \mathbb{N}$ tal que $k_m \geq n \implies d(x_{k_m}, x_0) < \epsilon$, entonces $\forall m \geq n$ se tiene que $k_m \geq m \geq n$ y por tanto $d(x_{k_m}, x_0) < \epsilon$. \square

Definición 1.0.25 [Sucesión de Cauchy]

Se dice que una sucesión es de Cauchy si

$$\forall \epsilon > 0 \exists N \in \mathbb{N} : \forall m, n \geq N \implies d(x_m, x_n) < \epsilon$$

Teorema 1.0.7

Una sucesión de \mathbb{R}^n es convergente si y solo si es de Cauchy.

Demostración. “ \implies ” Sea $\{x_k\}$ una sucesión convergente a x_0 , entonces dado $\epsilon > 0 \exists n \in \mathbb{N} : \forall k \in \mathbb{N} k \geq n \implies d(x_k, x_0) < \frac{\epsilon}{2}$. Por si tomamos $m, n \geq N$, entonces tenemos por la desigualdad triangular que:

$$d(x_m, x_n) \leq d(x_m, x_0) + d(x_n, x_0) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Lo que nos proporciona la condición de Cauchy. “ \Leftarrow ” Sea $\{x_k\}$ una sucesión de Cauchy, entonces dado $\epsilon > 0 \exists N \in \mathbb{N} : \forall m, n \geq N \implies d(x_m, x_n) < \epsilon$. En particular, la sucesión $\{x_{k,i}\}_{i \in \mathbb{N}}$ es una sucesión de Cauchy de número reales y por tanto por el curso de Análisis en Variable Real sabemos que existe $x_{0,i} \in \mathbb{R}$ tal que $x_{k,i} \rightarrow x_{0,i}$. Considerando que $x_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,n})$ \square

Definición 1.0.26

1. Llamaremos **recubrimiento** de un subconjunto S de \mathbb{R}^n a cualquier colección de subconjuntos de \mathbb{R}^n cuya unión contenga a S
2. Se dice que un recubrimiento es **abierto** si los conjuntos que lo forman son abiertos.
3. Llamaremos **subrecubrimiento** de un recubrimiento de un conjunto S a toda colección de elementos del recubrimiento que sea un recubrimiento de S
4. Finalmente llamaremos **recubrimiento finito** de un conjunto $S \subset \mathbb{R}^n$ cuando se trate de un recubrimiento de S formado por una cantidad finita de subconjuntos de \mathbb{R}^n .

Definición 1.0.27 [Compacidad]

Se dice que un subconjunto $K \subset \mathbb{R}^n$ es **compacto** si todo recubrimiento abierto de K admite un subrecubrimiento finito.

Ejemplo

1. $\{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$ es conjunto compacto de \mathbb{R} . En efecto, si tomamos un recubrimiento abierto $\{A_\alpha : \alpha \in \Lambda\}$ (Λ es un conjunto arbitrario) de nuestro conjunto, entonces el 0 debe estar en un cierto A_{α_0} . Al ser A_{α_0} abierto existe un intervalo centrado en 0 contenido en A_{α_0} . Como la sucesión $\{1/k\}$ converge a 0 existe ν tal que los x_k , con $k \geq \nu$, pertenecen a A_{α_0} . Luego si consideramos una cantidad finita de abiertos A_α que contengan a los puntos $1, \frac{1}{2}, \dots, \frac{1}{\nu-1}$ y el A_{α_0} tendremos un subrecubrimiento finito del conjunto $\{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$. El mismo argumento prueba que el conjunto formado por los puntos de una sucesión convergente y su límite es un conjunto compacto.
2. El conjunto $\{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ no es compacto pues del recubrimiento de él formado por los abiertos $A_i : i \in \mathbb{N}$, donde $A_i = \left(\frac{1}{i} - \frac{1}{i+1}, \frac{1}{i} + \frac{1}{i+1}\right)$, no se puede extraer ningún subrecubrimiento finito.
3. El conjunto $(0, 2)$ no es compacto, pues del recubrimiento abierto

$$\left\{ (1, 3), \left(\frac{1}{2}, 3\right), \left(\frac{1}{3}, 3\right), \dots \right\}$$

de él no se puede extraer ningún subrecubrimiento finito.

Proposición 1.0.9

Los subconjuntos cerrados contenidos en un compacto son compactos.

Demostración. Sea $C \subset \mathbb{R}^n$ cerrado contenido en K compacto. Consideremos un recubrimiento $\{A_\alpha : \alpha \in \Delta\}$ de C formado por abiertos. Llamaremos A al complementario de C en \mathbb{R}^n (i.e. $A = \mathbb{R}^n \setminus C$). Entonces $\{A_\alpha \cup A : \alpha \in \Delta\}$ es un recubrimiento abierto de K por lo que existirán $\alpha_1, \dots, \alpha_k \in \Delta$ tales que $K \subset A_{\alpha_1} \cup \dots \cup A_{\alpha_k} \cup A$. Evidentemente, $C \subset A_{\alpha_1} \cup \dots \cup A_{\alpha_k}$, luego $\{A_{\alpha_1}, \dots, A_{\alpha_k}\}$ es un subrecubrimiento finito de C . \square

Teorema 1.0.8 [Teorema de Bolzano-Weierstrass de compacidad por sucesiones]

Un subconjunto $K \subset \mathbb{R}^n$ es compacto si y sólo si toda sucesión de elementos de K tiene una subsucesión convergente a un punto de K .

Demostración. “ \Rightarrow ” Sea K un conjunto compacto y sea $\{x_k\} \subset K$ una sucesión de elementos de K tal que no tiene ninguna subsucesión convergente a un punto de K .

Entonces $\{x_k\}$ tiene que tener infinitos elementos distintos, pues en caso contrario tendría subsucesiones convergente a un punto de K . Denotemos por $\{x_{k_m}\}$ subsucesión de $\{x_k\}$ y S al conjunto formado por los puntos de la subsucesión. Entonces, $S' = \emptyset$ pues si existiese $x_0 \in S'$ entonces habría una subsucesión de puntos de S convergente a x_0 .

Ésto nos permite afirmar que $\forall m \in \mathbb{N} \exists r_m > 0 : B(x_{k_m}, r_m) \cap S = \{x_{k_m}\}$ (ya que el conjunto de puntos de acumulación de S es nulo $S' = \emptyset$) y que S es cerrado (pues contiene a $S' = \emptyset$) ya que como S está contenido en el compacto K es también compacto, pero ésto no es posible ya que del recubrimiento de S formado por las bolas abiertas $B(x_{k_m}, r_m)$ no es posible extraer ninguno finito.

“ \Leftarrow ” Supongamos que K tiene una sucesión de elementos de K que no tiene una subsucesión convergente a un punto de K . Entonces,

1. Veamos que $\forall r > 0$ existe un recubrimiento de abiertos definido por $\bigcup_{k \in \mathbb{N}} B(x_k, r)$, entonces existe una cantidad finita x_1, \dots, x_m de puntos de K tales que

$$K \subset \bigcup_{i=1}^m B(x_i, r)$$

Si ésto no fuera así, existiría un $r_0 > 0$ tal que K no se podría recubrir por ninguna cantidad finita de bolas de radio r_0 centradas en puntos de K :

Si fijamos $x_1 \in K$ como $B(x_1, r_0)$ no recubre a K existe $x_2 \in K \setminus B(x_1, r_0)$, como tampoco $B(x_1, r_0) \cup B(x_2, r_0)$ recubre a K existe $x_3 \in K \setminus (B(x_1, r_0) \cup B(x_2, r_0))$ y así sucesivamente, obtenemos una sucesión de puntos de K tales que no tienen ninguna subsucesión convergente, pues para cada $p, q \in \mathbb{N}$ $d(x_p, x_q) \geq r_0$, pero ésto llega a contradicción con la hipótesis.

2. Sea $\{A_\alpha : \alpha \in \Delta\}$ un recubrimiento abierto de K veremos como podemos extraer un subrecubrimiento finito:

Sabemos que existe $r^* > 0$ que $\forall x \in K \exists \alpha_x \in \Delta : B(x, r^*) \subset A_{\alpha_x}$, ya que en caso contrario existiría para todo $k \in \mathbb{N}$ un $x_k \in K$ tal que $B(x_k, \frac{1}{k})$ no estaría contenida en A_β cualquiera que sea $\beta \in \Delta$; es decir, las bolas pequeñas alrededor de x_k no caben enteramente en ningún A_β . La sucesión $\{x_k\}$ así formada deberá tener una subsucesión $\{x_{k_m}\}_{m=1}^{+\infty}$ convergente a un punto $x_0 \in K$ (por hipótesis). Ese punto x_0 deberá estar en algún A_{α_0} abierto, por lo que existe un $r_0 > 0$ tal que $B(x_0, r_0) \subset A_{\alpha_0}$.

Si tomamos un m lo suficientemente grande para que $d(x_{k_m}, x_0) < \frac{r_0}{2}$ y $\frac{1}{k_m} < \frac{r_0}{2}$. Entonces $B(x_{k_m}, \frac{1}{k_m}) \subset B(x_0, r_0) \subset A_{\alpha_0}$, lo que contradice que el que $B(x_{k_m}, \frac{1}{k_m})$ no estuviera contenido en A_β para ningún $\beta \in \Delta$, por la suposición inicial.

3. Finalmente, fijemos un $r^* > 0$ tal que para todo $x \in K \exists \alpha_x \in \Delta : B(x, r^*) \subset A_{\alpha_x}$, por lo que, por lo visto anteriormente existen puntos $x_1, \dots, x_k \in K$ tales que $K \subset \bigcup_{i=1}^k B(x_i, r^*)$ y dado que cada $B(x_i, r^*)$ está contenido en algún $A_{\alpha_{x_i}}$, vemos que una cantidad finita de A_α ya recubren a K .

□

Teorema 1.0.9 [Teorema de Bolzano-Weierstrass de compacidad por conjuntos]

Un conjunto K es compacto si y sólo si todo subconjunto de él con infinitos elementos tiene un punto de acumulación en K .

Demostración. “ \Rightarrow ” Sea K un conjunto compacto y sea $S \subset K$ un subconjunto de K con infinitos elementos, entonces podemos formar una sucesión de elementos distintos y por el Teorema de Bolzano-Weierstrass de compacidad por sucesiones, tenemos que existe una subsucesión convergente a un punto $x_0 \in K$. Por tanto, x_0 es un punto de acumulación de S y por tanto de K .

“ \Leftarrow ” Dada una sucesión de elementos de K pueden suceder dos cosas:

1. Tengan infinitos elementos distintos, en cuyo caso, por hipótesis, es un punto de acumulación de K y en consecuencia una subsucesión convergente a un punto de K .
2. No tenga infinitos términos distintos, en cuyo caso, necesariamente hay algún término que se repite infinitas veces, con lo que ya tiene una subsucesión convergente a un elemento de K .

□

Teorema 1.0.10 [Teorema de Heine-Borel]

Un conjunto es compacto \iff es cerrado y acotado.

Demostración. “ \Leftarrow ”

Sea K un conjunto cerrado y acotado. Consideremos una sucesión $\{x_k\} \subset K$ de puntos de K , y tomemos las sucesiones de sus coordenadas, dadas por $\{x_{k,i}\}_{k \in \mathbb{N}}$.

Tomemos $i = 1$. Se sabe, por el curso de Análisis en Variable Real, que al ser la sucesión $\{x_{k,1}\}$ acotada en \mathbb{R} , existe una subsucesión $\{x_{k_{\ell},1}\}$ que converge a un punto $x_{0,1}$ (teorema de Bolzano–Weierstrass para sucesiones de números reales).

La sucesión $\{x_{k_{\ell},2}\}$, por la misma razón, tiene una subsucesión convergente $\{x_{k_{\ell_r},2}\}$ a un punto $x_{0,2}$.

Reiterando este proceso para cada coordenada $i = 1, \dots, n$, obtenemos finalmente una sucesión

$$\{x_{k_{\ell_1 \dots \ell_n}}\}$$

convergente a un punto $x_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,n}) \in \mathbb{R}^n$.

Este punto es adherente a la sucesión $\{x_k\}$, y por tanto, adherente a K . Como K es cerrado, se concluye que $x_0 \in K$.

” \Rightarrow ” Demostraremos primero que si un conjunto es compacto, entonces es cerrado:

Sea K un compacto, probaremos que $\mathbb{R}^n \setminus K$ es abierto. Para ello sea $x \in \mathbb{R}^n \setminus K$. Como $x \notin K$ entonces $\forall y \in K \exists r_y > 0 : B(x, r_y) \cap B(y, r_y) = \emptyset$. Es claro que K está contenido en $\bigcup_{y \in K} B(y, r_y)$ y por tanto existen puntos $y_1, \dots, y_k \in K$ tales que $K \subset \bigcup_{i=1}^k B(y_i, r_{y_i})$. Entonces la bola de centro x y radio igual al mínimo de los r_{y_i} está contenida en $\mathbb{R}^n \setminus K$.

Ahora veamos que si es compacto, está acotado:

Sea K un conjunto compacto. Del recubrimiento por bolas abiertas $\{B(0, k) : k = 1, 2, \dots\}$ se tiene que poder extraer no finito. Esto es $k_0 \in \mathbb{N}$ tal que $K \subset B(0, k_0)$ con lo que

$$\sup\{d(x, y) : x, y \in K\} \leq 2k_0$$

y entonces K es acotado. □

Definición 1.0.28 [Limite]

Se dice que f tiene por **límite** l en el punto x_0 si

$$\forall \epsilon > 0 \exists \delta > 0 : \forall x \in B(x_0, \delta) \setminus \{x_0\} \implies d(f(x), l) < \epsilon$$

Teorema 1.0.11

Si la función f toma valores reales $\lim_{x \rightarrow x_0} f(x) = l$ y $l \neq 0$, entonces $\lim_{x \rightarrow x_0} \frac{1}{f(x)} = \frac{1}{l}$.

Demostración. Como $l \neq 0$ existe una bola cenrada en x_0 tal que la función es no nula en ella, o equivalentemente, $l \neq 0 \exists \delta_1 > 0 : f(x) \neq 0 \forall x \in B(x_0, \delta_1) \cap S$ tal que $x \neq x_0$.

Asimismo, también tenemos que para cierto $\delta_1 > 0$ se cumple que para $\epsilon = \frac{|l|}{2}$, $|f(x) - l| < \frac{|l|}{2}$ para todo $x \in B(x_0, \delta_1) \setminus \{x_0\}$. Con lo que obtenemos que $|f(x)| > \frac{|l|}{2}$. Para los $x \in B(x_0, \delta_1) \cap S$ se verifica que:

$$\left| \frac{1}{f(x)} - \frac{1}{l} \right| = \left| \frac{l - f(x)}{lf(x)} \right| = \frac{|f(x) - l|}{|l||f(x)|} \leq \frac{|f(x) - l|}{\frac{|l|^2}{2}}$$

Dado un $\epsilon > 0 \exists \delta_2 > 0 : \forall x \in B(x_0, \delta_2) \setminus \{x_0\} \implies |f(x) - l| < \epsilon \cdot \frac{|l|^2}{2}$. Luego para $\delta = \min\{\delta_1, \delta_2\}$ se cumple que para todo $x \in B(x_0, \delta) \setminus \{x_0\}$:

$$\left| \frac{1}{f(x)} - \frac{1}{l} \right| < \epsilon$$

□

Teorema 1.0.12 [Criterio del límite por sucesiones]

Sea $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función que tiene límite l en x_0 si y sólo si para toda sucesión $\{x_k\} \subset S$ tal que $x_k \rightarrow x_0$, se cumple que $f(x_k) \rightarrow l$.

Demostración. “ \implies ” Sean $x_0 \in S$ y $\{x_k\} \subset S : x_k \rightarrow x_0$ y $x_k \neq x_0 \forall k$. Dado $\epsilon > 0 \exists \delta > 0$ tal que si $x \in S$ y $0 < d(x, x_0) < \delta$, entonces $d(f(x), l) < \epsilon$. Como $x_k \rightarrow x_0$ entonces existe $N \in \mathbb{N}$ tal que para todo $k \geq N$, $d(x_k, x_0) < \delta$ y por tanto $d(f(x_k), l) < \epsilon$. Luego $f(x_k) \rightarrow l$.

“ \impliedby ”

Supongamos que f no tiene límite $\implies \exists \epsilon_0 > 0 : \forall \delta > 0 \exists x \in S : 0 < d(x, x_0) < \delta$ tal que $d(f(x), l) \geq \epsilon_0$. En particular $\forall k (\delta = \frac{1}{k}) \exists x_k \in S : 0 < d(x_k, x_0) < \frac{1}{k}$ y $d(f(x_k), l) \geq \epsilon_0$. Por tanto, la sucesión $\{x_k\}$ vemos que $d(x_k, x_0) \rightarrow 0$ pero como $d(f(x_k), l) \geq \epsilon_0$ entonces $f(x_k) \not\rightarrow l$. Luego llegamos a contradicción y si que tiene que haber límite. □

Definición 1.0.29 [Continuidad]

Sean S un subconjunto de \mathbb{R}^n , x_0 un punto de $S \cap S'$ y $f : S \rightarrow \mathbb{R}^m$ se dice que f es **continua** en x_0 si existe $\lim_{x \rightarrow x_0} f(x) = f(x_0)$, equivalentemente,

$$\forall \epsilon > 0 \exists \delta > 0 : \forall x \in S, 0 < d(x, x_0) < \delta \implies d(f(x), f(x_0)) < \epsilon$$

Teorema 1.0.13 [Criterio de continuidad por sucesiones]

Sean S un subconjunto de \mathbb{R}^n , $x_0 \in S \cap S'$ y $f : S \rightarrow \mathbb{R}^m$. Entonces, f es continua en x_0 si y solo si para toda sucesión $\{x_k\} \subset S$ tal que $x_k \rightarrow x_0$, se cumple que $f(x_k) \rightarrow f(x_0)$.

Definición 1.0.30 [Aplicación continua]

Diremos que una aplicación $f : S \rightarrow \mathbb{R}^m$ es **continua** si es continua en cada punto de S .

Observación 1.0.2

Sean $f, g : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $x_0 \in S$, y $\alpha \in \mathbb{R}$. Se verifican las siguientes propiedades:

- Si f y g son continuas en x_0 , entonces $f + g$ también es continua en x_0 .
- Si f es continua en x_0 , entonces αf también es continua en x_0 .
- Si $f, g : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ son continuas en x_0 , entonces fg es continua en x_0 .

- Si $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es continua en x_0 y $f(x) \neq 0$ para todo x en un entorno de x_0 , entonces $\frac{1}{f}$ es continua en x_0 .

Teorema 1.0.14

Para una aplicación $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, las siguientes afirmaciones son equivalentes:

1. f es continua en S
2. Para todo abierto $A \subset \mathbb{R}^m$, existe un abierto $B \subset S$ tal que $f^{-1}(A) = B \cap S$ (es decir, $f^{-1}(A) \subset S$)

Demostración. “ \Rightarrow ”

Hagamos una distinción de casos:

1. $f^{-1}(A) = \emptyset \implies B = \emptyset$ que cómo es abierto, lo cumple.
2. Si $f^{-1}(A) \neq \emptyset$: $\forall x \in f^{-1}(A) \exists \epsilon_x > 0 : B(f(x), \epsilon_x) \subset A$. La continuidad de f en x nos da que existe $\delta_x > 0$ tal que $f(B(x, \delta_x) \cap S) \subset B(f(x), \epsilon_x)$. Sea $B = \bigcup_{x \in f^{-1}(A)} B(x, \delta_x)$ abierto. Veamos que se cumple que $f^{-1}(A) = B \cap S$:
 - “ \subset ”: $x_0 \in f^{-1}(A) \implies x \in B \cap S$
 - “ \supset ”: Si $x_0 \in B \cap S$, entonces existe $x \in f^{-1}(A)$ tal que $x_0 \in B(x, \delta_x)$ como $f(B(x, \delta_x) \cap S) \subset B(f(x), \epsilon_x) \subset A \implies f(x_0) \in A \implies x_0 \in f^{-1}(A)$.

“ \Leftarrow ”

Dado $x_0 \in S$ y $\epsilon > 0$ consideremos la bola abierta $A = B(f(x_0), \epsilon)$ y el abierto B tal que $f^{-1}(A) = B \cap S$. Como $x_0 \in f^{-1}(A)$, entonces $x_0 \in B \cap S$, y dado que B es un abierto, existe $\delta > 0$ tal que $B(x_0, \delta) \subset B \implies f(B(x_0, \delta) \cap S) \subset B(f(x_0), \epsilon)$. \square

Definición 1.0.31 [Homeomorfismo]

Sea $f : S \subset \mathbb{R}^n \rightarrow T \subset \mathbb{R}^m$ una aplicación continua y biyectiva. Se dice que f es un **homeomorfismo** si su inversa $f^{-1} : T \rightarrow S$ también es continua.

Teorema 1.0.15

Si f es una aplicación continua en K compacto $\implies f(K)$ es un compacto en \mathbb{R}^m .

Demostración. Sea $\{A_\alpha : \alpha \in \Delta\}$ un recubrimiento abierto de $f(K)$, entonces $K \subset \bigcup_{\alpha \in \Delta} f^{-1}(A_\alpha)$. Para cada $\alpha \in \Delta$ sea G_α un abierto de \mathbb{R}^n tal que $f^{-1}(A_\alpha) = G_\alpha \cap K$. Entonces, $K \subset \bigcup_{\alpha \in \Delta} G_\alpha$. La compacidad de K nos da la existencia de un conjunto finito de índices $\alpha_1, \dots, \alpha_k \in \Delta$ tal que $K \subset G_{\alpha_1} \cup \dots \cup G_{\alpha_k}$. Por tanto, $f(K) \subset A_{\alpha_1} \cup \dots \cup A_{\alpha_k}$, lo que implica que $f(K)$ es compacto. \square

Teorema 1.0.16 [Teorema del máximo y el mínimo]

Sea $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función continua, siendo K compacto

$$\implies \exists a, b \in K : f(a) = \inf f(x) : x \in K \text{ y } f(b) = \sup f(x) : x \in K$$

Demostración. $f(K)$ es compacto (por teorema anterior) $\implies f(K)$ es acotado \implies tiene ínfimo y supremo. El ínfimo es un punto adherente pero como $f(K)$ es cerrado ese punto pertenece al conjunto luego es $f(a)$ para algún $a \in K$ y lo mismo con el supremo. \square

Definición 1.0.32 [Continuidad uniforme]

Sea $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ se dice que es **uniformemente continua** en S si

$$\forall \epsilon > 0 \exists \delta > 0 : \forall x, y \in S, d(x, y) < \delta \implies d(f(x), f(y)) < \epsilon$$

Observación 1.0.3

La diferencia esencial con respecto a la continuidad en cada punto es que aquí dado un ϵ existe un único δ que sirva para todo par de puntos de S que estén a una distancia menor que δ .

Teorema 1.0.17

Si $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es uniformemente continua en S y $\{x_k\}$ es una sucesión de Cauchy en S , entonces $\{f(x_k)\}$ es una sucesión de Cauchy en \mathbb{R}^m .

Demostración. Sea $\epsilon > 0$ Por cada $x \in K \exists \delta_x > 0$ tal que si $y \in K$ y $d(x, y) < \delta_x$ entonces $d(f(x), f(y)) < \frac{\epsilon}{2}$. La colección de bolas $B(x, \frac{\delta_x}{2})$ cuando x varia en K es un recubrimiento abierto de K , luego existen $x_1, \dots, x_k \in K$ tales que $K \subset B(x_1, \frac{\delta_{x_1}}{2}) \cup \dots \cup B(x_k, \frac{\delta_{x_k}}{2})$. Sea δ el mínimo de los δ_{x_i} , entonces dados x e $y \in K$ tales que $d(x, y) < \delta$ consideramos un $j = 1, \dots, k$ tal que $x \in B(x_j, \frac{\delta_{x_j}}{2})$ resulta que $y \in B(x_j, \delta_{x_j})$ pues:

$$d(y, x_j) \leq d(y, x) + d(x, x_j) < \delta_{x_j}$$

En consecuencia

$$d(f(x), f(y)) \leq d(f(x), f(x_j)) + d(f(x_j), f(y)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

\square

2 Aplicaciones diferenciables

2.1 Derivadas direccionales y gradiente

Definición 2.1.1 [Dirección]

Llamaremos **dirección** a un vector $v \in \mathbb{R}^n$. Normalmente de norma 1.

Si por ejemplo tenemos $n = 1$ tenemos sólo dos direcciones, $v = 1$ y $v = -1$. En cambio para $n > 1$ tenemos infinitas direcciones. En el caso de \mathbb{R}^2 las direcciones de norma 1 pueden escribirse de la forma $v = (\cos \theta, \sin \theta)$, con $\theta \in [0, 2\pi)$.

Definición 2.1.2 [Recta]

Llamaremos **recta** pasando por x_0 y de dirección v a la recta $x(t) = x_0 + tv$, donde $t \in \mathbb{R}$.

Definición 2.1.3 [Derivada direccional]

Si f es una función definida en un subconjunto abierto A de \mathbb{R}^n , x_0 es un punto de A y v es una dirección de \mathbb{R}^n , se define la derivada de f en x_0 en la dirección v como

$$D_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

Observación 2.1.1

Para cualquier dirección v tanto ella como su opuesta $-v$ definen la misma recta pasando por x_0 (el vector de dirección también determina una orientación). Sin embargo, las derivadas en las direcciones v y $-v$ son de signo opuesto:

$$\begin{aligned} D_{-v} f(x_0) &= \lim_{t \rightarrow 0} \frac{f(x_0 + t(-v)) - f(x_0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(x_0 + (-t)v) - f(x_0)}{t} \\ &= \lim_{s \rightarrow 0} \frac{f(x_0 + sv) - f(x_0)}{-s} \\ &= - \lim_{s \rightarrow 0} \frac{f(x_0 + sv) - f(x_0)}{s} \\ &= -D_v f(x_0). \end{aligned}$$

Definición 2.1.4 [Derivadas parciales]

Considémos en \mathbb{R}^m las direcciones dadas por los vectores $e_i = (0, \dots, 0, 1, 0, \dots, 0)$, donde 1 está en la i -ésima posición. Las derivadas en un punto x_0 de una función f en estas direcciones (si es que existen) se llaman **derivadas parciales** de f en x_0 y se denotan por $D_i f(x_0) = D_{e_i} f(x_0)$, o también por $\frac{\partial f}{\partial x_i}(x_0)$ o $f_{x_i}(x_0)$.

Definición 2.1.5 [Gradiente]

Dada una función $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, que tenga todas las derivadas parciales en un punto $x_0 \in A$, se llama **gradiente** de f en x_0 al vector

$$\nabla f(x_0) = (D_1 f(x_0), D_2 f(x_0), \dots, D_n f(x_0)) = \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right) \in \mathbb{R}^n$$

Definición 2.1.6 [Diferenciable]

Dada una función $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, es **diferenciable** en un punto $x_0 \in A$ si existe una aplicación lineal $L : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = 0$$

Lo cual equivale:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{\|h\|} = 0$$

Observación 2.1.2

La existencia del gradiente no garantiza la diferenciabilidad de la función.

Proposición 2.1.1

Toda aplicación lineal es diferenciable $\forall x \in \mathbb{R}^n$ y su gradiente es la aplicación lineal misma.

Teorema 2.1.1

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable en $x_0 \in A$, entonces es derivable en x_0 en todas las direcciones $v \in \mathbb{R}^n$. Además sea $v \in \mathbb{R}^n$ una dirección, que supondremos de norma 1, y una aplicación lineal $L : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{\|h\|} = 0$$

entonces

$$L(v) = D_v f(x_0) = \nabla f(x_0) \cdot v$$

Demostración. Tomando sólo vectores h de la forma $h = tv$ con $t \in \mathbb{R}$, tales que $\|v\| = 1$, tenemos que, por la diferenciabilidad de f en x_0 , se cumple que

$$\begin{aligned} & \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0) - L_{x_0}(tv)}{\|tv\|} = \\ &= \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0) - tL_{x_0}(v)}{|t|} = 0 \iff \\ &\iff \lim_{t \rightarrow 0} \left| \frac{f(x_0 + tv) - f(x_0) - tL_{x_0}(v)}{|t|} \right| = 0 \\ &\iff \lim_{t \rightarrow 0} \frac{|f(x_0 + tv) - f(x_0) - tL_{x_0}(v)|}{|t|} = 0 \iff \end{aligned}$$

$$\begin{aligned}
&\Longleftrightarrow \lim_{t \rightarrow 0} \left| \frac{f(x_0 + tv) - f(x_0) - tL_{x_0}(v)}{|t|} \right| = 0 \Longleftrightarrow \\
&\Longleftrightarrow \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{|t|} - L_{x_0}(v) = 0 \\
&\Longleftrightarrow \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{|t|} = L_{x_0}(v)
\end{aligned}$$

luego f es derivable en x_0 en la dirección v y se tiene que $D_v f(x_0) = L(v)$. □

2.2 Aplicaciones diferenciables

Definición 2.2.1 [Diferencial]

Para representar la función L usaremos la notación df_{x_0} , que se llama **diferencial** de f en x_0 .

Observación 2.2.1

Cuando f es diferenciable en x_0 :

$$\frac{\partial f}{\partial x_i}(x_0) = D_i f(x_0) = df(x_0)(e_i) = \nabla f(x_0) \cdot e_i$$

donde e_i es el vector de dirección en la i -ésima coordenada.

Corolario 2.2.1

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una aplicación diferenciable en un punto $x_0 \in A$ tal que $\nabla f(x_0) \neq 0$. Entonces el valor máximo de $|D_v f(x_0)|$ se alcanza para $v = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$ y ese valor máximo es $\|\nabla f(x_0)\|_2$.

Demostración. Sabemos que la derivada direccional de f en el punto x_0 y en la dirección del vector unitario v viene dada por

$$D_v f(x_0) = \langle \nabla f(x_0), v \rangle.$$

Por la desigualdad de Cauchy-Schwarz, se tiene que

$$|\langle \nabla f(x_0), v \rangle| \leq \|\nabla f(x_0)\|_2 \|v\|_2 = \|\nabla f(x_0)\|_2,$$

ya que v es unitario ($\|v\|_2 = 1$).

El valor máximo se alcanza cuando v tiene la misma dirección que $\nabla f(x_0)$, es decir,

$$v = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|_2}.$$

En ese caso,

$$D_v f(x_0) = \left\langle \nabla f(x_0), \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|_2} \right\rangle = \frac{\|\nabla f(x_0)\|_2^2}{\|\nabla f(x_0)\|_2} = \|\nabla f(x_0)\|_2.$$

Por lo tanto, el valor máximo de $|D_v f(x_0)|$ es $\|\nabla f(x_0)\|_2$ y se alcanza precisamente en la dirección del gradiente. □

Definición 2.2.2 [Espacio afín tangente]

Cuando f es diferenciable en x_0 llamaremos **espacio afín tangente** a la gráfica $\{(x, f(x)) : x \in A\} \subset \mathbb{R}^{n+1}$ en el punto $(x_0, f(x_0))$ a

$$T = \{(x, f(x_0)) : \langle \nabla f(x_0), x - x_0 \rangle : x \in \mathbb{R}^n\} = \{(x, f(x_0)) + df(x_0)(x - x_0)\}$$

Observación 2.2.2

Los espacios afines tangentes son hiperplanos

Observación 2.2.3

El concepto intuitivo de tangencia que tenemos para el caso de curvas en \mathbb{R}^2 mantiene también para superficies en \mathbb{R}^n pues si T es el espacio afín tangente a la superficie $\{(x, f(x)) : x \in A\}$ en el punto $(x_0, f(x_0))$, entonces para cada punto $(x, f(x)) : x \in A$ existe un punto $(x, y) \in T$ tal que

$$\lim_{x \rightarrow x_0} \frac{f(x) - y}{\|x - x_0\|} = 0$$

basta tomar $y = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$.

Teorema 2.2.1

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable en un punto x_0 entonces f es continua en x_0 .

Demostración. Escribamos

$$f(x) - f(x_0) = f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle + \langle \nabla f(x_0), x - x_0 \rangle$$

y consideremos el límite

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle}{\|x - x_0\|} = 0$$

resulta que

$$\exists \|x - x_0\| < \delta_1 \implies |f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle| \leq \|x - x_0\|$$

Dado que por Cauchy-Schwarz $\|\langle \nabla f(x_0), x - x_0 \rangle\| \leq \|\nabla f(x_0)\| \|x - x_0\|$, podemos escribir

$$|f(x) - f(x_0)| < (1 + \|\nabla f(x_0)\|) \|x - x_0\|$$

si $\|x - x_0\| < \delta_1$. Luego si dado $\epsilon > 0$ tomamos $\delta = \min\{\delta_1, \frac{\epsilon}{1 + \|\nabla f(x_0)\|}\}$, tenemos que si $\|x - x_0\| < \delta$ entonces $|f(x) - f(x_0)| < \epsilon$. \square

Teorema 2.2.2

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una aplicación que tiene derivadas parciales en cada punto de A . Si para cada $i = 1, \dots, n$ la función

$$df : x \in A \mapsto \frac{\partial f}{\partial x_i}(x)$$

es continua en un punto $x_0 \in A$, entonces f es diferenciable en x_0

Demostración. Queremos ver si

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle}{\|x - x_0\|} = 0$$

Sea una función de la forma $\varphi_i(x) = f(x_{0_1}, \dots, x_{0_{i-1}}, x_i, x_{0_{i+1}}, \dots, x_{0_n})$ para $i = 1, \dots, n$ y fijemos el punto $x_0 \in A$ en el que todas las derivadas de f son continua y consideremos un $r > 0$ tal que $B(x_0, r) \subset A$. Entonces, para cada punto $x \in B(x_0, r)$, tenemos que

$$\begin{aligned} f(x) - f(x_0) &= f(x_1, \dots, x_n) - f(x_{0_1}, \dots, x_{0_n}) = \\ &= f(x_1, x_2, \dots, x_n) - f(x_{0_1}, x_2, \dots, x_n) + \\ &+ f(x_{0_1}, x_2, \dots, x_n) - f(x_{0_1}, x_{0_2}, \dots, x_n) + \\ &+ \dots + f(x_{0_1}, x_{0_2}, \dots, x_n) - f(x_{0_1}, x_{0_2}, \dots, x_{0_n}) \\ &= f(x) - \varphi_1(x) + \varphi_1(x) - \varphi_2(x) + \varphi_2(x) + \dots + \varphi_n(x) - \varphi_n(x) - f(x_0) \end{aligned}$$

Entonces, apliquemos el Teorema de Valor Medio a $\varphi_1(x)$: $\varphi_1(s) = f(s, x_{0_2}, \dots, x_{0_n})$ para $s \in [x_{0_1}, x_1]$ es continua y derivable por lo que debe existir un punto $u_1 \in (x_{0_1}, x_1)$ tal que

$$\varphi_1(x_1) - \varphi_1(x_{0_1}) = \varphi_1'(u_1)(x_1 - x_{0_1})$$

Si $x_1 = x_{0_1}$ pasamos a la siguiente coordenada, pues en esta primera la diferencia es nula. Pero además tenemos que:

$$\varphi_1'(u_1) = \frac{\partial f}{\partial x_1}(u_1, x_{0_2}, \dots, x_{0_n})$$

Por lo que,

$$f(x_1, x_2, \dots, x_n) - f(x_{0_1}, x_2, \dots, x_n) = D_1 f(u_1, x_{0_2}, \dots, x_{0_n})(x_1 - x_{0_1})$$

Repitiendo el proceso, conseguimos la existencia de un vector $u = (u_1, u_2, \dots, u_n)$ tales que

$$f(x_{0_1}, \dots, x_{0_{i-1}}, x_i, x_{i+1}, \dots, x_n) - f(x_{0_1}, \dots, x_{0_{i-1}}, x_{0_i}, x_{i+1}, \dots, x_n) = D_i f(x_{0_1}, \dots, x_{0_{i-1}}, u_i, x_{i+1}, \dots, x_n)(x_i - x_{0_i})$$

Tenemos entonces que,

$$f(x) - f(x_0) = \sum_{i=1}^n D_i f(x_{0_1}, \dots, x_{0_{i-1}}, u_i, x_{i+1}, \dots, x_n)(x_i - x_{0_i})$$

Ahora volvamos al límite que queríamos calcular:

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle}{\|x - x_0\|}$$

Sustituamos la expresión que hemos obtenido para $f(x) - f(x_0)$:

$$\begin{aligned} &\frac{f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle}{\|x - x_0\|} = \\ &= \frac{D_1(u_1, \dots, x_n)(x_1 - x_{0_1}) + \dots + D_n f(x_{0_1}, \dots, u_n)(x_n - x_{0_n}) - [D_1 f(x_0)(x_1 - x_{0_1}) + \dots + D_n f(x_0)(x_n - x_{0_n})]}{\|x - x_0\|} \\ &\leq \frac{|D_1 f(x_{0_1}, \dots, u_n)(x_1 - x_{0_1}) - D_1 f(x_0)(x_1 - x_{0_1})| + \dots + |D_n f(x_{0_1}, \dots, u_n)(x_n - x_{0_n}) - D_n f(x_0)(x_n - x_{0_n})|}{\|x - x_0\|} \\ &\leq \frac{(D_1 f(u_1, \dots, x_n) - D_1 f(x_0))(x - x_0)}{\|x - x_0\|} + \dots + \frac{(D_n f(x_{0_1}, \dots, u_n) - D_n f(x_0))(x - x_0)}{\|x - x_0\|} = \\ &= (D_1 f(u_1, \dots, x_n) - D_1 f(x_0)) + \dots + (D_n f(x_{0_1}, \dots, u_n) - D_n f(x_0)) \xrightarrow{x \rightarrow x_0} 0 \end{aligned}$$

□

Observación 2.2.4

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es tal que $\exists df(x) \forall x \in A$, entonces se puede hablar de la función diferencial $df : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$, que es una aplicación lineal en cada punto $x \in A$

Definición 2.2.3 [Matriz Jacobiana]

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una aplicación diferenciable, entonces la matriz $J_f(x)$ de la función diferencial $df(x)$ se llama **matriz jacobiana** de f en el punto x . Si f tiene derivadas parciales en x , entonces la matriz jacobiana es la matriz cuyas entradas son las derivadas parciales de f en x :

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

donde $f = (f_1, f_2, \dots, f_m)$ y $x = (x_1, x_2, \dots, x_n)$.

Observación 2.2.5

Para aplicaciones lineales $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$, se verifica que existe $C > 0$ tal que

$$\|L(x)\| \leq C\|x\| \quad \forall x \in \mathbb{R}^n$$

Proposición 2.2.1 [Propiedades de la matriz jacobiana]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable. Entonces:

1. $d(f+g)(x_0) = df(x_0) + dg(x_0)$ para $f, g : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciables en x_0 .
2. $d(\alpha f)(x_0) = \alpha df(x_0)$ para $\alpha \in \mathbb{R}$ y $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en x_0 .
3. $d(f \cdot g) = g(x_0)df(x_0) + f(x_0)dg(x_0)$ para $f, g : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciables en x_0 .

Teorema 2.2.3

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable en un punto de $x_0 \in A$. Entonces existen constantes $M > 0$ y $\delta > 0$ tales que

$$\|x - x_0\| < \delta \implies \|f(x) - f(x_0)\| \leq M\|x - x_0\|$$

En particular, esto implica la continuidad de f en x_0 .

Demostración. Dado que f es diferenciable se cumple que:

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - df(x_0)(x - x_0)}{\|x - x_0\|} = 0$$

Sea $\epsilon = 1 \implies \exists \delta_1 > 0$ tal que si $\|x - x_0\| < \delta_1$ entonces

$$0 < \|x - x_0\| < \delta_1 \implies |f(x) - f(x_0) - df(x_0)(x - x_0)| < \|x - x_0\|$$

Con lo que

$$f(x) - f(x_0) = df(x_0)(x - x_0) + [f(x) - f(x_0) - df(x_0)(x - x_0)] \implies$$

$$|f(x) - f(x_0)| = |df(x_0)(x - x_0) + [f(x) - f(x_0) - df(x_0)(x - x_0)]|$$

$$\|f(x) - f(x_0)\| \leq \|df(x_0)(x - x_0)\| + \|x - x_0\| \leq |df(x_0)(x - x_0)| + |f(x) - f(x_0) - df(x_0)(x - x_0)|$$

Pero por la hipótesis tenemos que

$$|f(x) - f(x_0) - df(x_0)(x - x_0)| < \|x - x_0\|$$

Entonces

$$\|f(x) - f(x_0)\| \leq |df(x_0)(x - x_0)| + \|x - x_0\|$$

Pero como $df(x_0)$ es una aplicación lineal, existe una constante $C > 0$ tal que

$$\|df(x_0)(x - x_0)\| \leq C\|x - x_0\|$$

Entonces tenemos que $\forall x \in \mathbb{R}^n$ se cumple que

$$\|f(x) - f(x_0)\| \leq M\|x - x_0\|$$

para los x tales que $\|x - x_0\| < \delta_1$, donde $M = C + 1$. □

2.3 Regla de la cadena

Teorema 2.3.1 [Regla de la cadena]

Sea $A \subset \mathbb{R}^n$ y $B \subset \mathbb{R}^m$ abiertos y $f : A \rightarrow \mathbb{R}^m$ y $g : B \rightarrow \mathbb{R}^p$ tales que $f(A) \subset B$, f es diferenciable en $x_0 \in A$ y g es diferenciable en $f(x_0)$. Entonces $g \circ f$ es diferenciable en x_0 y

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)$$

A nivel de matrices, ésto es equivalente a que

$$J_{g \circ f}(x_0) = J_g(f(x_0)) \cdot J_f(x_0)$$

Demostración. Tenemos que

$$\frac{\|g(f(x)) - g(f(x_0)) - dg(f(x_0)) \circ df(x_0)(x - x_0)\|}{\|x - x_0\|} \leq$$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - dg(f(x_0))(f(x) - f(x_0))\|}{\|x - x_0\|} + \frac{\|dg(f(x_0))(f(x) - f(x_0) - df(x_0)(x - x_0))\|}{\|x - x_0\|}$$

Al ser f diferenciable en x_0 por el teorema anterior, $M > 0$ y $\delta_1 > 0$ tales que

$$\|x - x_0\| < \delta_1 \implies \|f(x) - f(x_0)\| \leq M\|x - x_0\|$$

Por otra parte tenemos que al ser g diferenciable en $f(x_0)$, dado $\frac{\epsilon}{2M} > 0$ existe $\delta_2 > 0$ tal que

$$0 < \|y - f(x_0)\| < \delta_2 \implies \frac{\|g(y) - g(f(x_0)) - dg(f(x_0))(y - f(x_0))\|}{\|y - f(x_0)\|} < \frac{\epsilon}{2M}$$

Y esto se cumple $\forall y \in B$, en particular para $y = f(x)$.

Tomando $\delta_3 = \min\{\delta_1, \delta_2\}$ tenemos que

$$0 < \|x - x_0\| < \delta_3 \implies \frac{\|g(f(x)) - g(f(x_0)) - dg(f(x_0))(f(x) - f(x_0))\|}{\|x - x_0\|} =$$

$$= \frac{\|g(f(x)) - g(f(x_0)) - dg(f(x_0))(f(x) - f(x_0))\| \|f(x) - f(x_0)\|}{\|x - x_0\| \|f(x) - f(x_0)\|} \leq \frac{\epsilon}{2M} \cdot \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \leq \frac{\epsilon}{2M} \cdot M = \frac{\epsilon}{2}$$

Además, por ser $dg(f(x_0))$ una aplicación lineal, existe $C^* > 0$ tal que $\|dg(f(x_0))(y)\| \leq C^* \|y\|$ para todo $y \in \mathbb{R}^m$.

Además, por ser f diferenciable en x_0 , existe $\delta_4 > 0$ tal que

$$0 < \|x - x_0\| < \delta_4 \implies \frac{\|f(x) - f(x_0) - df(x_0)(x - x_0)\|}{\|x - x_0\|} < \frac{\epsilon}{2C^*} \implies$$

$$\frac{\|df(x_0)[f(x) - f(x_0) - df(x_0)(x - x_0)]\|}{\|x - x_0\|} \leq \frac{C^* \cdot \|f(x) - f(x_0) - df(x_0)(x - x_0)\|}{\|x - x_0\|} \leq \frac{C^*}{2} \cdot \frac{\epsilon}{C^*} = \frac{\epsilon}{2}$$

Tomando $\delta = \min\{\delta_3, \delta_4\}$ y tomando x tales que $0 < \|x - x_0\| < \delta$ resulta que

$$\begin{aligned} & \frac{\|g(f(x)) - g(f(x_0)) - dg(f(x_0)) \circ df(x_0)(x - x_0)\|}{\|x - x_0\|} \leq \\ & \leq \frac{\|g(f(x)) - g(f(x_0)) - dg(f(x_0))(f(x) - f(x_0))\|}{\|x - x_0\|} + \frac{\|dg(f(x_0))(f(x) - f(x_0) - df(x_0)(x - x_0))\|}{\|x - x_0\|} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

□

Definición 2.3.1 [Convexidad]

Se dice que un subconjunto $S \subset \mathbb{R}^n$ es convexo si para todo par de puntos $x, y \in S : x \neq y$ se verifica que el segmento $L[x, y] = \{x + t(y - x) : t \in [0, 1]\} = \{xt + y(1 - t) : t \in [0, 1]\}$ de extremos x e y está contenido en S

Observación 2.3.1

Las bolas son conjuntos convexos con lo que $L[x, y] \subset B(x_0, r)$

2.4 Teoremas del valor medio

Teorema 2.4.1 [Teorema del Valor Medio]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable en cada punto de A . Entonces para todo $x, y \in A$ tales que $L[x, y] \subset A$ y para todo $z \in \mathbb{R}^m$ se verifica que existe $c \in L[x, y]$ tales que:

$$\langle z, f(y) - f(x) \rangle = \langle z, df(c)(y - x) \rangle$$

En principio, c depende de x, y, z

Demostración. Como A es abierto $\implies \exists \delta > 0 : (1 - t)x + ty \in A \forall t \in [-\delta, \delta + 1]$.

Fijemos $z \in \mathbb{R}^m$ un vector cualquiera, definamos la función

$$\varphi(t) = \langle z, f((1 - t)x + ty) \rangle : t \in (-\delta, \delta + 1)$$

Esta función es derivable en $(-\delta, \delta + 1)$ y

$$\varphi'(t) = \langle z, df((1 - t)x + ty)((y - x)) \rangle$$

pues

$$\varphi(t) = z_1 f_1((1-t)x + ty) + z_2 f_2((1-t)x + ty) + \dots + z_m f_m((1-t)x + ty)$$

téngase en cuenta que se trata de la diferencial de la composición de f con la función lineal $g(t) = (1-t)x + ty$. Luego por el Teorema del Valor Medio para funciones en \mathbb{R} existe $s \in (0, 1)$ tal que

$$\varphi(1) - \varphi(0) = \varphi'(s)(1 - 0)$$

de donde se sigue que:

$$\langle z, f(y) - f(x) \rangle = \langle z, f(y) \rangle - \langle z, f(x) \rangle = \varphi(1) - \varphi(0) = \varphi'(s) = \langle z, df((1-s)x + sy)(y-x) \rangle = \langle z, df(c)(y-x) \rangle$$

Basta tomar $c = (1-s)x + sy$ para concluir la demostración. \square

Teorema 2.4.2 [Teorema del valor medio para funciones $\mathbb{R}^n \rightarrow \mathbb{R}$]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable en cada punto de A . Entonces para todo $x, y \in A$ tales que $L[x, y] \subset A$ existe $c \in L[x, y]$ tales que

$$f(y) - f(x) = df(c)(y - x)$$

Demostración. Como en este caso $m = 1$ si tomamos $z = 1$ el resultado se sigue directamente del Teorema del Valor Medio anterior, pues

$$f(y) - f(x) = \langle 1, f(y) - f(x) \rangle = \langle 1, df(c)(y - x) \rangle = df(c)(y - x)$$

\square

Teorema 2.4.3 [Desigualdad del valor medio para aplicaciones de $\mathbb{R}^n \rightarrow \mathbb{R}^m$ o Teorema de los incrementos finitos]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable en A . Entonces para todo $x, y \in A$ tales que $L[x, y] \subset A$ se verifica que

$$\|f(y) - f(x)\| \leq \sup_{c \in L[x, y]} \|df(c)\| \cdot \|y - x\|$$

Demostración. Sea $x, y \in A : L[x, y] \subset A$ y sea $z = \frac{f(y) - f(x)}{\|f(y) - f(x)\|}$, vector unitario. Por el Teorema del Valor Medio para funciones de $\mathbb{R}^n \rightarrow \mathbb{R}$, existe $c \in L[x, y]$ tal que

$$\langle z, f(y) - f(x) \rangle = \langle z, df(c)(y - x) \rangle$$

De donde se sigue que

$$\begin{aligned} \|f(y) - f(x)\| &= \frac{\|f(y) - f(x)\|^2}{\|f(y) - f(x)\|} = \left\langle \frac{f(y) - f(x)}{\|f(y) - f(x)\|}, f(y) - f(x) \right\rangle = \left\langle \frac{f(y) - f(x)}{\|f(y) - f(x)\|}, df(c)(y - x) \right\rangle \leq \\ &\leq \left\| \frac{f(y) - f(x)}{\|f(y) - f(x)\|} \right\| \|df(c)(y - x)\| = \|df(c)(y - x)\| \leq \|df(c)\| \cdot \|y - x\| \leq \sup_{c \in L[x, y]} \|df(c)\| \cdot \|y - x\| \end{aligned}$$

\square

Teorema 2.4.4

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación diferenciable en A . Entonces para todo $x, y \in A$ tales que $L[x, y] \subset A$ se verifica que

$$f_j(y) - f_j(x) = df_j(c_j)(y - x)$$

para algún $c_j \in L[x, y]$ y para cada $j = 1, \dots, m$.

Corolario 2.4.1

Sea $A \subset \mathbb{R}^n$ abierto y convexo y $f : A \subset \mathbb{R}^m$ es diferenciable en A y $df(x) = 0 \forall x \in A$. Entonces f es constante en A .

2.5 Derivadas de orden superior. Teorema de Taylor

Definición 2.5.1 [Derivada parciales segundo]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una aplicación diferenciable en A . Entonces la **derivada parcial segunda** de f respecto a la variable x_i es

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

y la **derivada parcial mixta** de f respecto a las variables x_i y x_j es

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

Definición 2.5.2 [Clase C^1]

Sea $A \subset \mathbb{R}^n$ un conjunto abierto. Una aplicación $f : A \rightarrow \mathbb{R}$ se dice que es de clase C^1 si tiene derivadas parciales en cada punto de A y estas son continuas en A . Si $f : A \rightarrow \mathbb{R}^m$ es una aplicación de clase C^1 , entonces se dice que es de clase C^1 si cada componente f_i es de clase C^1 .

Definición 2.5.3 [Clase C^2]

Sea $A \subset \mathbb{R}^n$ un conjunto abierto. Una aplicación $f : A \rightarrow \mathbb{R}$ se dice que es de clase C^2 si tiene derivadas parciales de orden 2 en cada punto de A y estas son continuas en A . Si $f : A \rightarrow \mathbb{R}^m$ es una aplicación de clase C^2 , entonces se dice que es de clase C^2 si cada componente f_i es de clase C^2 .

Definición 2.5.4 [Clase C^p]

Sea $A \subset \mathbb{R}^n$ un conjunto abierto. Una aplicación $f : A \rightarrow \mathbb{R}$ se dice que es de clase C^p si tiene derivadas parciales de orden k en cada punto de A y estas son continuas en A para todo $k = 1, \dots, p$. Si $f : A \rightarrow \mathbb{R}^m$ es una aplicación de clase C^p , entonces se dice que es de clase C^p si cada componente f_i es de clase C^p .

Lema 2.5.1

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^2 . Fijado $x_0 \in A$ sea $\delta > 0$ tal que $B(x_0, \delta) \subset A$. Entonces para cada u y $v \in B(0, \frac{\delta}{2}) \setminus \{0\}$ existen $\alpha, \beta \in (0, 1)$ tales que

$$f(x_0 + u + v) - f(x_0 + u) - (f(x_0 + v) - f(x_0)) = D_v(D_u f)(x_0 + \alpha u + \beta v).$$

Demostración. Consideremos la función

$$g : B(0, \frac{\delta}{2}) \rightarrow \mathbb{R}$$

$$g(x) = f(x + v) - f(x)$$

Esta función g es diferenciable en $B(x_0, \frac{\delta}{2})$ y $dg(x) = df(x + v) - df(x)$. Apliquemos el teorema del valor medio a g en el segmento $L[x_0, x_0 + u]$ y obtenemos la existencia de un punto c entre x_0 y $x_0 + u$ tal que

$$g(x_0 + u) - g(x_0) = dg(c)(u)$$

Esto es,

$$f(x_0 + u + v) - f(x_0 + u) - (f(x_0 + v) - f(x_0)) = (df(c + v) - df(c))(u)$$

Si escribimos $c = x_0 + \alpha u$ para algún $\alpha \in (0, 1)$, tenemos que

$$f(x_0 + u + v) - f(x_0 + u) - (f(x_0 + v) - f(x_0)) = (df(x_0 + \alpha u + v) - df(x_0 + \alpha u))(u) = D_u f(x_0 + \alpha u + v) - D_u f(x_0 + \alpha u)$$

Ahora como $D_u f$ es diferenciable en A y $L[x_0 + \alpha u, x_0 + \alpha u + v] \subset A$, aplicamos el teorema del valor medio a $D_u f$ y nos da que existe un punto $e \in L[x_0 + \alpha u, x_0 + \alpha u + v]$ tal que

$$D_u f(x_0 + \alpha u + v) - D_u f(x_0 + \alpha u) = d(D_u f)(e)(v) = D_v(D_u f)(e)$$

Si escribimos $e = x_0 + \alpha u + \beta v$ para algún $\beta \in (0, 1)$, tenemos el resultado □

Teorema 2.5.1 [Teorema de Schwarz]

Sea $A \subset \mathbb{R}^n$ un abierto y $f \in C^2(A)$, entonces $D_{ij}f(x) = D_{ji}f(x)$ para todo $x \in A$ y para todo $i, j = 1, \dots, n$.

Demostración. Sea $x_0 \in A$. Probaremos que $\forall \epsilon > 0 |D_{ij}f(x_0) - D_{ji}f(x_0)| < \epsilon$. Como $D_{ij}f$ y $D_{ji}f$ son continuas en x_0 , existe $\delta > 0$ tal que $B(x_0, \delta) \subset A$ y

$$\|x - x_0\| < \delta \implies \begin{cases} |D_{ij}f(x) - D_{ij}f(x_0)| < \epsilon/2 \\ |D_{ji}f(x) - D_{ji}f(x_0)| < \epsilon/2 \end{cases}$$

Tomemos $u = te_i$ y $v = se_j$ con $t, s \in (0, \frac{\delta}{2})$. Tenemos así puntos que están en las condiciones del lema anterior, aplicando ese lema obtenemos que existen $\alpha_1, \beta_1, \alpha_2, \beta_2 \in (0, 1)$ tales que

$$\begin{cases} f(x_0 + te_i + se_j) - f(x_0 + te_i) - (f(x_0 + se_j) - f(x_0)) = D_{se_j}(D_{te_i}f)(x_0 + \alpha_1 te_i + \beta_1 se_j) \\ f(x_0 + se_j + te_i) - f(x_0 + se_j) - (f(x_0 + te_i) - f(x_0)) = D_{te_i}(D_{se_j}f)(x_0 + \alpha_2 se_j + \beta_2 te_i) \end{cases}$$

Como en general $D_{\lambda u}f(x) = \lambda D_u f(x)$ y $A = B$, tenemos que

$$stD_{ij}f(x_0 + \alpha_1 te_i + \beta_1 se_j) = tsD_{ji}f(x_0 + \alpha_2 se_j + \beta_2 te_i)$$

Entonces por la desigualdad triangular tenemos que

$$|D_{ij}f(x_0) - D_{ji}f(x_0)| \leq |D_{ij}f(x_0) - D_{ij}f(x_0 + \alpha_2 se_j + \beta_2 te_i)| + |D_{ji}f(x_0 + \alpha_2 se_j + \beta_2 te_i) - D_{ji}f(x_0)|$$

y tal como habíamos tomado s y t esta suma es menor que ϵ pues $\|\alpha_2 se_j + \beta_2 te_i\| < s + t < \delta$ y $\|\alpha_1 te_i + \beta_1 se_j\| < s + t < \delta$. Por tanto, la arbitrariedad de ϵ nos da que $D_{ij}f(x_0) = D_{ji}f(x_0)$. □

Observación 2.5.1

1. Si f es de clase C^1 y existe $\frac{\partial^2 f}{\partial x \partial y}$ y es continua entonces también existe $\frac{\partial^2 f}{\partial y \partial x}$ y coinciden.
2. También se puede obtener la igualdad entre D_{12} y D_{21} a partir de la existencia de $D_1 f$ y $D_2 f$ en un entorno de (x_0, y_0) y su diferenciabilidad.

Corolario 2.5.1

Si $f \in C^3(A)$ entonces $D_{ijk}f(x) = D_{\sigma(i),\sigma(j),\sigma(k)}f(x)$ para todo $x \in A$ y toda permutación σ de $\{i, j, k\}$.

Demostración. Supongamos que σ está dada por $\sigma(i) = k, \sigma(j) = i, \sigma(k) = j$.

$$\begin{aligned} D_{ijk}f(x) &= D_i(D_j(D_k f))(x) \\ &= D_i(D_{jk}f)(x) = D_i(D_{kj}f)(x) \\ &= D_i(D_k(D_j f))(x) = D_{ik}(D_j f)(x) = D_{ki}(D_j f)(x) = D_{kij}f(x). \end{aligned}$$

Las $D_p f$ son de clase C^2 y se les puede aplicar el teorema de Schwarz. □

Teorema 2.5.2 [Teorema de Taylor]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^{k+1} en A y sean $x_0 \in A$ y $h \neq 0$ tales que $L[x_0, x_0 + h] \subset A$. Entonces existe un punto $c \in L[x_0, x_0 + h]$ tal que

$$\begin{aligned} f(x_0 + h) &= f(x_0) + \sum_{i_1=1}^n D_{i_1}f(x_0)h_{i_1} \\ &\quad + \frac{1}{2!} \sum_{i_1=1}^n \sum_{i_2=1}^n D_{i_1 i_2}f(x_0)h_{i_1}h_{i_2} + \cdots \\ &\quad + \frac{1}{k!} \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n D_{i_1 \cdots i_k}f(x_0)h_{i_1} \cdots h_{i_k} \\ &\quad + \frac{1}{(k+1)!} \sum_{i_1=1}^n \cdots \sum_{i_{k+1}=1}^n D_{i_1 \cdots i_{k+1}}f(c)h_{i_1} \cdots h_{i_k}h_{i_{k+1}}, \end{aligned}$$

o equivalentemente usando $h = (x - x_0)$

$$\begin{aligned} f(x) &= f(x_0) + \sum_{i_1=1}^n D_{i_1} f(x_0)(x_{i_1} - x_{0i_1}) \\ &+ \frac{1}{2!} \sum_{i_1=1}^n \sum_{i_2=1}^n D_{i_1 i_2} f(x_0)(x_{i_1} - x_{0i_1})(x_{i_2} - x_{0i_2}) + \cdots \\ &+ \frac{1}{k!} \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n D_{i_1 \dots i_k} f(x_0)(x_{i_1} - x_{0i_1}) \cdots (x_{i_k} - x_{0i_k}) \\ &+ \frac{1}{(k+1)!} \sum_{i_1=1}^n \cdots \sum_{i_{k+1}=1}^n D_{i_1 \dots i_{k+1}} f(c)(x_{i_1} - x_{0i_1}) \cdots (x_{i_k} - x_{0i_k})(x_{i_{k+1}} - x_{0i_{k+1}}). \end{aligned}$$

Demostración. Sea la aplicación $\varphi : \mathbb{R} \rightarrow \mathbb{R}^n$ definida por $\varphi(t) = x_0 + th$. Dado que φ es continua entonces $B = \varphi^{-1}(A) \subset \mathbb{R}$ es abierto.

Consideremos la composicion $g = f \circ \varphi : B \subset \mathbb{R} \rightarrow \mathbb{R}$, definida por $g(t) = f(x_0 + th)$, que es de clase C^{k+1} en B (pues es una función polinómica y por tanto es f la que restringe la clase). Además se tienq eue $\varphi'(t) = h$. Entonces aplicando la regla de la cadena a φ obtenemos que:

$$\begin{aligned} g'(t) &= df(\varphi(t)) \circ d\varphi(t) = \sum_{i_1=1}^n \frac{\partial f}{\partial x_{i_1}}(\varphi(t)) \cdot \frac{\partial \varphi}{\partial t}(t) = \sum_{i_1=1}^n D_{i_1} f(\varphi(t)) h_{i_1} = \langle \nabla f(\varphi(t)), h \rangle \\ g^{(j)} &= \sum_{i_1=1, \dots, i_j=1}^n D_{i_1 \dots i_j} f(\varphi(t)) \cdot h_{i_1} \cdots h_{i_j} \end{aligned}$$

Si aplicamos nuevamente la regla de la cadena a φ obtenemos que:

$$\frac{\partial}{\partial t} \nabla f(\varphi(t)) = H_f(\varphi(t)) \cdot \frac{\partial \varphi}{\partial t}(t) = H_f(\varphi(t)) \cdot h$$

Entonces tendríamos que

$$g''(t) = \langle \langle H_f(\varphi(t)), h \rangle, h \rangle = H_f(\varphi(t)) \cdot h \cdot h = \sum_{i_1=1, i_2=1}^n \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(\varphi(t)) h_{i_1} h_{i_2}$$

Entonces por inducción podemos llegar a la derivada $\varphi^{(j)}$:

$$g^{(j)}(t) = \sum_{i_1=1, \dots, i_n=1}^n D_{i_1 \dots i_j} f(\varphi(t)) h_{i_1} \cdots h_{i_j}$$

Podemos aplicar el Teorema de Taylor para una variable una variable real φ centrado en a :

$$f(x) = \sum_{i=0}^k \frac{f^{(i)}(a)(x-a)^i}{i!}$$

Tomando $a = 0$ como centro y $x = 1$ y además tomaremos la fórmula del resto del valor medio del resto $R_k(x) = \frac{f^{(k+1)}(\alpha)}{(k+1)!} (x-a)^{k+1}$, donde $\alpha \in (0, 1)$, entonces tenemos que:

$$\begin{aligned} g(1) &= g(0) + \frac{1}{2!} g''(0) + \cdots + \frac{1}{k!} g^{(k)}(0) + R_k(1) = \\ &= f(x_0) + \sum_{i_1=1}^n D_{i_1} f(x_0) h_{i_1} + \frac{1}{2!} \sum_{i_1=1}^n \sum_{i_2=1}^n D_{i_1 i_2} f(x_0) h_{i_1} h_{i_2} + \cdots + \frac{1}{k!} \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n D_{i_1 \dots i_k} f(x_0) h_{i_1} \cdots h_{i_k} + R_k(1) \end{aligned}$$

□

2.6 Extremos locales

Definición 2.6.1 [Extremos relativos y absolutos]

Diremos $f : A \subset \mathbb{R}^n$ y $x_0 \in A$. Diremos que f tiene un **máximo relativo** en x_0 si existe $r > 0$ tal que $B(x_0, r) \subset A$ y $f(x) \leq f(x_0)$ para todo $x \in B(x_0, r)$. Si existe una bola centrada en x_0 en la $f(x) \leq f(x_0)$ diremos que f tiene un **mínimo relativo** en x_0 . Cuando se da alguna de esas desigualdades para todo $x \in A$ diremos que f tiene un **máximo absoluto** en x_0 y un **mínimo absoluto** en x_0 .

Definición 2.6.2 [Punto crítico]

Si f tiene todas las derivadas parciales de primer orden en un punto x_0 diremos que x_0 es un **punto crítico** de f si $\nabla f(x_0) = 0$.

Teorema 2.6.1

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ tiene un extremo relativo en $x_0 \in A$ y f tiene todas las derivadas parciales en ese punto entonces $\nabla f(x_0) = 0$.

Demostración. f tiene un máximo relativo en x_0 , entonces

$$\frac{f(x_0 + te_i) - f(x_0)}{t} \leq 0 \quad \forall t > 0$$

Dado que existe el límite de ese cociente, necesariamente tiene que ser 0. □

Definición 2.6.3 [Punto de ensilladura]

Los puntos críticos de f que no son ni máximos ni mínimos relativos se denominan **puntos de ensilladura** o **puntos de silla**.

Definición 2.6.4 [Matriz hesiana]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^2 en A . Entonces la **matriz hesiana** de f en un punto $x_0 \in A$ es la matriz cuadrada de orden n cuyas entradas son las derivadas parciales segundas de f :

$$H_f(x_0) = \begin{pmatrix} D_{11}f(x_0) & D_{12}f(x_0) & \cdots & D_{1n}f(x_0) \\ D_{21}f(x_0) & D_{22}f(x_0) & \cdots & D_{2n}f(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ D_{n1}f(x_0) & D_{n2}f(x_0) & \cdots & D_{nn}f(x_0) \end{pmatrix}$$

Definición 2.6.5 [Definición de signo]

Se dice que una forma cuadrática $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ es **definida positiva** si $Q(x) > 0$ para todo $x \neq 0$. Se dice que es **definida negativa** si $Q(x) < 0$ para todo $x \neq 0$. Se dice que es **indefinida** si no es ni positiva ni negativa. Se dice que una forma cuadrática es **semidefinida positiva** si $Q(x) \geq 0$ para todo $x \in \mathbb{R}^n$. Se dice que es **semidefinida negativa** si $Q(x) \leq 0$ para todo $x \in \mathbb{R}^n$.

Teorema 2.6.2

Sea f una función de clase C^2 en un abierto de \mathbb{R}^n y supongamos que un punto x_0 de él es un punto crítico para f .

- (a) Si la forma cuadrática Q asociada a $H_f(x_0)$ es definida negativa, entonces f tiene en x_0 un máximo relativo.
- (b) Si f tiene un máximo relativo en x_0 , entonces Q es semidefinida negativa.
- (c) Si Q es definida positiva, entonces f tiene en x_0 un mínimo relativo.
- (d) Si f tiene un mínimo relativo en x_0 , entonces Q es semidefinida positiva.
- (e) Si Q es indefinida entonces la función tiene un punto de ensilladura en x_0 .

Demostración. (a) Como el conjunto $K = \{x \in \mathbb{R}^n : \|x\| = 1\}$ es un conjunto compacto y por tanto acotado, existe $\tilde{x} \in K$ tal que

$$Q(\tilde{x}) \geq Q(x) \quad \forall x \in K$$

Ya que al ser Q una aplicación continua y K un compacto, en él debe alcanzar un máximo. Por otro lado al ser Q definida negativa, tenemos que $Q(\tilde{x}) < 0$, por tanto, sea $\epsilon_0 = -Q(\tilde{x})$. Tenemos que por la continuidad de las segundas derivadas de f en x_0 , existe $\delta > 0$ tal que $B(x_0, \delta) \subset A$ y

$$\|x - x_0\| < \delta \implies |D_{ij}f(x) - D_{ij}f(x_0)| < \frac{\epsilon_0}{2n^2} \quad \forall i, j = 1, \dots, n$$

Por otra parte, para $h \in \mathbb{R}^n : h \neq 0$ se verifica que:

$$Q(h) = \|h\|^2 Q\left(\frac{h}{\|h\|}\right) \leq \|h\|^2 Q(\tilde{x}) = -\epsilon_0 \|h\|^2$$

Apliquemos ahora el Teorema de Taylor, hasta el grado 1, y obtengamos un $c \in L[x_0, x_0 + h]$ para el resto tal que

$$f(x_0 + h) - f(x_0) = \frac{1}{2} \sum_{i,j=1}^n D_{ij}f(c) h_i h_j =$$

Nota: Las primeras derivadas son nulas pues es un punto crítico.

$$\begin{aligned} &= \sum_{i,j=1}^n D_{ij}f(c) h_i h_j + D_{ij}f(x_0) h_i h_j - D_{ij}f(x_0) h_i h_j = \sum (D_{ij}f(c) - D_{ij}f(x_0)) h_i h_j + D_{ij}f(x_0) h_i h_j = \\ &= \sum_{i,j=1}^n (D_{ij}f(c) - D_{ij}f(x_0)) h_i h_j + Q(h) \end{aligned}$$

Si tomamos la cota de la continuidad anterior y lo aplicamos a c , tenemos que:

$$\|c - x_0\| \leq \|x_0 + h - x_0\| = \|h\| < \delta \implies |D_{ij}f(c) - D_{ij}f(x_0)| < \frac{\epsilon_0}{2n^2}$$

Con lo que obtenemos que:

$$\sum_{i,j=1}^n (D_{ij}f(c) - D_{ij}f(x_0)) h_i h_j < \frac{\epsilon_0}{2} \|h\|^2$$

Además, como por otra parte tenemos que

$$Q(h) \leq -\epsilon_0 \|h\|^2$$

Obtenemos finalmente que

$$f(x_0 + h) - f(x_0) = \frac{1}{2} \sum_{i,j=1}^n D_{ij} f(x) h_i h_j \leq \frac{1}{2} (-\epsilon_0 + \frac{\epsilon_0}{2}) \|h\|^2 < 0$$

Entonces obtenemos que $f(x) < f(x_0) \forall x \in B(x_0, \delta)$ y finalmente que f tiene un máximo local en x_0 .

(b) Demostremos esto por medio del absurdo:

Si existe h_0 tal que $Q(h_0) > 0$ consideremos la función $g(t) = -f(x_0 + th_0) = -f(\varphi(t))$ siendo $\varphi(t) = x_0 + th_0$. Se verifique que g es de clase C^2 en un intervalo centrado en 0

$$g'(t) = -(f \circ \varphi)'(t) = -\sum_{i=1}^n D_i f(x_0 + th_0) h_{0_i} = -\sum_{i=1}^n (D_i f \circ \varphi)(t) h_{0_i} = \langle \nabla f(x_0 + th_0), h_0 \rangle$$

$$g''(t) = -(f \circ \varphi)''(t) = -\sum_{i=1}^n (D_i f \circ \varphi)'(t) h_{0_i} = -\sum_{i=1}^n \left(\sum_{j=1}^n D_j (D_i f)(\varphi(t)) h_{0_j} \right) h_{0_i}$$

Entonces, cómo $g'(0) = 0$ (por ser un punto crítico) y $g''(0) = -Q(h_0) < 0$, tenemos que g es cóncava hacia abajo, entonces g decrece cerca de $t = 0 \implies$

$$-f(x_0 + th_0) < -f(x_0) \quad \forall t \in (-\delta, \delta)$$

por lo que f no tendría un máximo local en x_0 .

□

Teorema 2.6.3

Sea $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ un función de clase C^2 y $(x_0, y_0) \in A$ un punto crítico de f y sea $\Delta = \det(H_f(x_0, y_0)) = ac - b^2$, tenemos que:

1. Si $\Delta > 0$ y $a > 0$ entonces f tiene un mínimo relativo en (x_0, y_0)
2. Si $\Delta > 0$ y $a < 0$ entonces f tiene un máximo relativo en (x_0, y_0)
3. Si $\Delta < 0$ entonces f tiene un punto de silla en (x_0, y_0)
4. Si $\Delta = 0$ entonces no se sabe nada

Demostración. Para la forma cuadrática Q asociada a $H_f(x_0, y_0)$ se verifica que:

$$Q(h_1, h_2) = ah_1^2 + 2bh_1h_2 + ch_2^2$$

Si $\Delta > 0$ entonces $a \neq 0$ pues $\Delta = ac - b^2$ por lo que:

$$Q(h_1, h_2) = ah_1^2 + 2bh_1h_2 + ch_2^2 = a \left(h_1 + \frac{b}{a} h_2 \right)^2 + \frac{\Delta}{a} h_2^2$$

1. Si $a > 0 \implies Q(h_1, h_2) > 0 \quad \forall h \in \mathbb{R}^2 \implies f$ tiene un mínimo relativo en (x_0, y_0) .
2. Si $a < 0 \implies Q(h_1, h_2) < 0 \quad \forall h \in \mathbb{R}^2 \implies f$ tiene un máximo relativo en (x_0, y_0) .

3. Si $\Delta < 0$, haremos una distinción de casos:

(a) $a = 0 \implies Q(h_1, h_2) = 2bh_1h_2 + ch_2^2$, y tendríamos que $b \neq 0$ pues si $b = 0$ y $a = 0$ tendríamos que $\Delta = 0$ y hacemos nuevamente distinción de casos:

$$\text{i. } c = 0 \implies Q(h_1, h_2) = 2bh_1h_2 \implies \begin{cases} Q(1, 1) = 2b \\ Q(1, -1) = -2b \end{cases}$$

$$\text{ii. } c \neq 0 \implies Q(h_1, h_2) = 2bh_1h_2 + ch_2^2 \implies \begin{cases} Q(\frac{-c}{b}, 1) = -c \\ Q(\frac{-c}{4b}, 1) = \frac{c}{2} \end{cases}$$

Estos dos casos nos dan que Q toma valores opuestos en diferentes puntos, es decir, Q es indefinida

4. $a \neq 0 \implies \begin{cases} Q(1, 0) = a \\ Q(\frac{-b}{a}, 1) = \frac{\Delta}{a} < 0 \end{cases} \implies Q$ toma valores opuestos en diferentes puntos, es decir, Q es indefinida.

5. Ni podemos decir nada pues depende de cada caso particular

□

Teorema 2.6.4 [Criterio de Sylvester]

1. Si $\Delta_k > 0 \quad \forall k = 1, \dots, n$ y f tiene un mínimo relativo en x_0
2. Si $(-1)^k \Delta_k > 0 \quad \forall k = 1, \dots, n$ y f tiene un máximo relativo en x_0
3. Si $\Delta \neq 0$ entonces no hay extremo

Observación 2.6.1

Las funciones de más de una variable pueden tener un único punto crítico y éste puede ser un extremo relativo y no absoluto. Por ejemplo, la función

$$f(x, y) = y^2 + x^2(1 + y)^3$$

tiene un único punto crítico en $(0, 0)$ y es un mínimo relativo pero no absoluto, ya que $f(0, 0) = 0$ y $f(1, -4) = -9$

2.7 Teoremas de la función inversa, implícita y de los multiplicadores de Lagrange

Proposición 2.7.1

Sea $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ k -lipschitz con $k \in (0, 1) \iff \forall x, y \in U \implies \|f(x) - f(y)\| \leq k\|x - y\|$. Entonces, la función $g = \text{id} + f$ es un homeomorfismo de U en V -abierto de \mathbb{R}^n

$$g : U \rightarrow V$$

$$g^{-1} : V \rightarrow U$$

Teorema 2.7.1

Sea $f : \overline{B(x_0, r)} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una aplicación continua en $\overline{B(x_0, r)}$ y y diferenciable en $B(x_0, r)$. Si $\det(J_f(x)) \neq 0$ para todo $x \in B(x_0, r)$ y $f(x) \neq f(x_0) \forall x : \|x - x_0\| = r$, entonces $f(x_0)$ es un punto interior de $f(B(x_0, r))$

Demostración. Consideremos la función

$$\varphi : x \in \overline{B(x_0, r)} \subset \mathbb{R}^n \mapsto \varphi(x) = \|f(x) - f(x_0)\|$$

Como φ es continua (ya que f es continua en la bola) y como la frontera de la bola es compacta, entonces $\exists x^* \in \{x : \|x - x_0\| = r\}$ tal que:

$$\varphi(x^*) = \min_{x \in \partial B(x_0, r)} \{\|f(x) - f(x_0)\|\}$$

Entonces, como $f(x) \neq f(x_0) \forall x \in \partial B(x_0, r)$, tenemos que $\varphi(x^*) > 0$. Ahora beamos que $B(f(x_0), \frac{m}{2}) \subset f(B(x_0, r))$:

Fijado un $y \in B(f(x_0), \frac{m}{2})$, tomemos la función auxiliar:

$$\psi : x \in \overline{B(x_0, r)} \subset \mathbb{R}^n \mapsto \psi(x) = \|f(x) - y\|$$

ψ es continua en el compacto $\overline{B(x_0, r)}$ por lo que existe $x^{**} \in \overline{B(x_0, r)}$ tal que $\psi(x^{**}) = \min_{x \in \overline{B(x_0, r)}} \{\|f(x) - y\|\}$. Ahora veremos que:

•

□

Corolario 2.7.1

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ una aplicación de clase $C^1(A)$ y supongamos que f es inyectiva y que $\det(J_f(x)) \neq 0$ para todo $x \in A$. Entonces f es una aplicación abierta.

Proposición 2.7.2

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ una aplicación de clase C^1 en el abierto de A y supongamos que para un punto $x_0 \in A$ se verifica que $\det(J_f(x_0)) \neq 0$. Entonces, $\exists r > 0$ tal que $B(x_0, r) \subset A$, f es inyectiva en $B(x_0, r)$ y $\det(J_f(x)) \neq 0 \forall x \in B(x_0, r)$

Teorema 2.7.2 [Teorema de la función inversa]

Sea $f : A \rightarrow \mathbb{R}^n \in C^1(A, \mathbb{R}^n)$ y supongamos que $x_0 \in A$ se verifica que $\det(J_f(x_0)) \neq 0$. Entonces existen entornos abiertos U de x_0 y V de $f(x_0)$ y una única aplicación $g : V \subset \mathbb{R}^n \rightarrow U \subset \mathbb{R}^n$ tal que:

1. $f(U) = V$
2. f es inyectiva en U
3. $g \circ f = Id$ en U
4. $g \in C^1(V, \mathbb{R}^n)$

Teorema 2.7.3 [Teorema de la función implícita]

Sea $F : A \subset \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n$ una aplicación de clase C^1 en A -abierto y supongamos que para todo punto $(x_0, y_0) \in A$ se verifica que $F(x_0, y_0) = 0$ y que $J_y F(x_0, y_0) \neq 0$. Entonces existen entornos abiertos Y_0 de y_0 y una única aplicación $G : Y_0 \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$ tal que:

1. $G \in C^1(Y_0, \mathbb{R}^n)$
2. $G(y_0) = x_0$
3. $F(G(y), y) = 0 \quad \forall y \in Y_0$
4. Existe un entorno $U \subset A$ de (x_0, y_0) tal que:

$$\{(x, y) \in U : F(x, y) = 0\} = \{(G(y), y) : y \in Y_0\}$$

Definición 2.7.1 [Puntos regulares]

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$ siendo A un abierto, un punto x_0 es un punto regular si f es de clase C^1 en un entorno abierto de x_0 la matriz jacobiana de f en x_0 es de rango máximo

Definición 2.7.2 [Variedad regular]

Se llama **variedad regular** de dimensión m en \mathbb{R}^n ($m < n$) a todo subconjunto M de \mathbb{R}^n tal que $\forall x_0 \in M$ existe $A \subset \mathbb{R}^n$ abierto que contiene a x_0 y existe una aplicación $F : A \rightarrow \mathbb{R}^{n-m}$, regular en cada punto de A , por lo que el rango de $J_F(x_0)$ es $n - m$, tal que:

$$M \cup A = \{x \in A : F(x) = 0 \in \mathbb{R}^{n-m}\}$$

Definición 2.7.3 [Extremos relativos condicionados]

Sean B es un conjunto abierto de \mathbb{R}^n y f una aplicación definida en B en \mathbb{R} y M un subconjunto de B , se dice que un punto $x_0 \in M$ es un **máximo relativo condicionado** de f en M si existe un entorno abierto U de x_0 tal que $U \cap M \subset B$ y $f(x) \leq f(x_0)$ para todo $x \in U \cap M$. Se dice que es un **mínimo relativo condicionado** si se cumple la desigualdad inversa.

Lema 2.7.1

Sea $F : A \subset \mathbb{R}^{n-m} \times \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$ de clase C^1 en el abierto de A y sea

$$M = \{x = (x_1, \dots, x_{n-m}, \dots, x_n) \in A : F(x) = 0\}$$

Entonces para todo $x_0 = (x_{01}, \dots, x_{0n-m}, \dots, x_{0n}) \in M$ tal que $\det(D_i F_j(x_0))_{i,j=1,\dots,n-m} \neq 0$ existe un entorno abierto $Y_0 \subset \mathbb{R}^m$ y existe $\psi : Y_0 \rightarrow A$ regular en cada punto de Y_0 tal que $\psi(y_0) = x_0$ y $F(\psi(y)) = 0 \forall y \in Y_0$.

Teorema 2.7.4 [Teorema de los multiplicadores de Lagrange]

Sea $f : B \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^1 en el abierto de B , definamos una variedad regular M contenida en B de dimensión $m < n$ tal que contenga a un extremo relativo $x_0 = (x_{0_1}, \dots, x_{0_{n-m}}, \dots, x_{0_n}) \in M$ de la función $f|_M$, entonces podemos definir una función $F : \mathbb{R}^{n-m} \times \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$ tal que:

$$M \cup A = \{x \in A : F(x) = 0\}$$

y podemos definir $\lambda_1, \dots, \lambda_{n-m} \in \mathbb{R}$ tales que:

$$\nabla f(x) = \lambda_1 \nabla F_1(x) + \dots + \lambda_{n-m} \nabla F_{n-m}(x)$$

Observación 2.7.1

1. A los λ_i se les llama **multiplicadores de Lagrange** y a la función $L = f + \sum_{i=1}^m \lambda_i g_i$ se le llama **función de Lagrange** o **función langrangiana**. El teorema nos dice que debemos buscar x_0 tal que:

$$\begin{cases} \nabla L(x_0) = \nabla f(x_0) + \sum_{i=1}^m \lambda_i \nabla g_i(x_0) = 0 \\ g_i(x_0) = 0 \quad i = 1, \dots, m \end{cases}$$

2. El teorema se puede enunciar en mayor generalidad que sobre variedades diferenciales M
3. En todo caso, el conjunto M debe ser muy regular (sin picos)