

CÁLCULO DIFERENCIAL

Juan Diego Barrado Daganzo e Iker Muñoz Martínez
2º de Carrera

3 de noviembre de 2022¹

¹Este documento se actualiza, para consultar las últimas versiones entrar en el enlace <https://github.com/JuanDiegoBarrado/CalculoDiferencial>

QUIÉNES SOMOS

Somos un grupo de estudiantes de la Universidad Complutense de Madrid, concretamente del Doble Grado de Informática y Matemáticas que queremos compartir unos apuntes de calidad y, como mínimo, ordenados para que os sea más fácil llevar la asignatura al día (sobre todo a estudiantes de Doble Grado).

Estos apuntes son posibles gracias a la colaboración de más alumnos como tú que deciden aportar un granito de arena al proyecto. Puedes contribuir de la siguiente manera:

- Notificando erratas
- Modificando erratas
- Proponiendo mejoras
- Aportando ejemplos nuevos
- Aportando nuevas versiones

Para contribuir no tienes más que ponerte en contacto con juandbar@ucm.es o dejarnos un *Pull Request* en <https://github.com/JuanDiegoBarrado/CalculoDiferencial>. Los detalles para que la contribución de todos sea lo más homogénea posible estarán en el fichero *Contribute.md* de dicho repositorio o, en caso de no aparecer correctamente, podéis poneros en contacto con el correo anteriormente mencionado.

Muchas gracias, esperamos que este documento te sea útil.

AGRADECIMIENTOS

Queremos dar gracias especialmente al Profesor Javier Soria, por ser el profesor que impartió la asignatura de *Cálculo Diferencial* durante la elaboración de estos apuntes y por darnos *feedback* sobre la calidad y las posibles mejoras de los mismos.

También queremos dar las gracias a los Profesores Víctor Manuel Sánchez, Jose María Martínez Ansemil y Socorro Ponte Miramontes, por elaborar otros manuales más formales sobre la asignatura que nos han permitido contrastar adecuadamente los nuestros y entender adecuadamente los contenidos aquí expuestos.

Cálculo Diferencial © 2021 by Juan Diego Barrado & Iker Muñoz is licensed under Attribution-NonCommercial 4.0 International. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc/4.0/>

Índice general

1. Espacios Métricos	3
1.1. Concepto de Métrica y Espacios Normados	3
1.2. Topología en Espacios Métricos	7
1.3. Sucesiones completitud y compacidad	13
1.3.1. Convergencia y sucesiones sobre espacios métricos	13
1.3.2. Concepto de compacidad y propiedades	15
2. Continuidad y Diferenciabilidad	23
2.1. Definición de Continuidad	23
2.1.1. Cálculo de límites en \mathbb{R}^n	26
2.2. Definición de Diferenciabilidad	30
2.2.1. Diferencial y propiedades	30
2.2.1.1. Interpretación geométrica de la diferencial	34
2.2.2. Taylor, extremos y puntos críticos	39
3. Teoremas Notables	47
3.1. Teorema de la función inversa	47
3.2. Teorema de la función implícita	51
3.3. Multiplicadores de Lagrange	54

ESPACIOS MÉTRICOS

Vamos a introducir un elemento fundamental en esta asignatura, y muy útil para las posteriores, que no es otra cosa que la capacidad de definir qué es una medida y cómo podemos medir las cosas según un criterio general.

CONCEPTO DE MÉTRICA Y ESPACIOS NORMADOS

En esta sección, se definen los elementos básicos para el estudio de medidas, distancias y se estudian las características de las estructuras que se generan a partir de dichas definiciones, con sus consecuentes resultados para otras áreas como la geometría.

Definición (métrica)

Sea $E \neq \emptyset$ un conjunto, decimos que $d : E \times E \rightarrow \mathbb{R}$ es una **métrica o distancia** siempre que se satisfaga las siguientes propiedades¹

- *Positiva:* $d(x, y) \geq 0 : \forall x, y \in E$
- *No degenerada:* $d(x, y) = 0 \Leftrightarrow x = y$
- *Simetría:* $d(x, y) = d(y, x) : \forall x, y \in E$
- *Desigualdad triangular:* $d(x, y) \leq d(x, z) + d(z, y) : \forall x, y, z \in E$

Al par (E, d) lo denotamos como **espacio métrico**.

Ejemplo:

- Un ejemplo sencillo de comprobar es considerar \mathbb{R} con la métrica $d(x, y) = |x - y|$ tradicional.
- Si escogemos un $E \neq \emptyset$ cualquiera, definimos la métrica discreta como:

$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

Y el par formado por ambos elementos constituyen una peculiar definición de espacio métrico.

- Definimos como **Espacio Euclídeo Usual** al conjunto:

$$\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, \cdots, x_n) : x_i \in \mathbb{R}\} \text{ donde } d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

Este, junto con la suma, el producto por escalares usual y la métrica definida (que más adelante veremos que es la métrica euclídea), tiene estructura de espacio vectorial y también de espacio métrico.

¹Nótese que no está permitido valores que tiendan a infinito (comprendidos en $\bar{\mathbb{R}}$)

Definición (Norma)

Sea E un espacio vectorial real, se dice que $\|\cdot\| : E \rightarrow \mathbb{R}$ es una **norma** si se cumplen las siguientes propiedades:

- *Positiva:* $\|x\| \geq 0 : \forall x \in E$
- *No degenerada:* $\|x\| = 0 \Leftrightarrow x = 0$
- *Homogénea:* $\|\lambda x\| = |\lambda| \cdot \|x\| : \forall \lambda \in \mathbb{R} \wedge \forall x \in E$
- *Desigualdad triangular:* $\|x + y\| \leq \|x\| + \|y\|$

Al par $(E, \|\cdot\|)$ se le denota como **espacio normado**.

Ejemplo:

- Considerando \mathbb{R}^n denominamos como la **clásica norma euclídea** a:

$$\|x\| = \|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2} : \forall x \in \mathbb{R}^n$$

La demostración de que constituye una norma es trivial salvo el último apartado, que no se puede demostrar hasta más adelante puesto que se necesita de la *Desigualdad de Cauchy-Schwarz*.

- Consideramos en \mathbb{R}^n la **norma p** donde $1 < p < \infty$:

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

- Denominamos como la **norma infinito** al extremo de la norma p:

$$\|x\|_\infty = \lim_{p \rightarrow \infty} \|x\|_p = \max\{|x_j| : j = 1, \dots, n\}$$

Proposición (Métrica asociada a una norma)

Si consideramos $(E, \|\cdot\|)$ un espacio normado, podemos definir la **métrica asociada a dicha norma** como:

$$d_{\|\cdot\|}(x, y) = \|x - y\| : \forall x, y \in E$$

Cuya definición, por ser en base a una norma ya dada, hace fácilmente verificable las condiciones de métrica.

Definición (Métrica Euclídea)

De este modo, se define la **métrica o distancia euclídea** en \mathbb{R}^n como la métrica asociada a la norma euclídea:

$$d(x, y) = d_2(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

Ejemplo:

- Definimos el conjunto de las funciones continuas $E = \{f : [0, 1] \rightarrow \mathbb{R} : f \text{ continua}\} = \mathcal{C}[0, 1]$ y la función:

$$d(f, g) = \sup\{|f(x) - g(x)| : \forall x \in [0, 1]\}$$

Lo primero de todo, comprobamos que está bien definida: como son funciones continuas la diferencia es continua y, por ser el valor absoluto una función continua, la composición con él también es continua. Consecuentemente, al tratarse de una función continua y acotada en $[0, 1]$ el máximo se alcanza, así que, de hecho, no solo el supremo es un número finito, sino que es máximo de la función en ese intervalo.

Para ver que se trata de una norma, comprobamos la última propiedad (las anteriores son triviales):

$$\begin{aligned} d(f, g) &= \sup |f(x) - g(x)| = \sup |f(x) - h(x) + h(x) - g(x)| \leq \\ &\leq \sup |f(x) - h(x)| + \sup |h(x) - g(x)| \leq d(f, h) + d(h, g) \end{aligned}$$

- Junto al ejemplo anterior, podemos definir:

$$\|f\|_{\infty} = \sup\{|f(x)| : x \in [0, 1]\}$$

y entonces el par $(E, \|\cdot\|_{\infty})$ es un espacio normado. Para comprobar que esto es una norma, se puede tratar dicha norma como $\|f\|_{\infty} = d(f, 0)$ para demostrar sus propiedades.

- Sea $E = \{f : (0, 1] \rightarrow \mathbb{R} \text{ continua}\}$ definimos:

$$d(f, g) = \sup\{|f(x) - g(x)| : \forall x \in (0, 1]\}$$

Sin embargo, no es una métrica porque no está bien definida; por ejemplo:

$$d\left(\frac{1}{x}, 0\right) = \infty$$

Observación:

Sin embargo, no toda métrica tiene asociada una norma en un espacio vectorial. Para verlo, tomamos \mathbb{R}^n con la métrica discreta definida: observamos que no existe una norma que haga que el par (\mathbb{R}^n, d) sea un espacio normado puesto que en caso de existir, por ejemplo, no se verifica la 3 propiedad:

$$\text{Sea } x \in \mathbb{R} : x \neq 0, \exists \|\cdot\| \Rightarrow \underbrace{d(\lambda x, 0)}_{=1} = \|\lambda x\| = |\lambda| \cdot \|x\| = |\lambda| \cdot d(x, 0) = |\lambda|$$

Y como debe ocurrir para todo lambda, es absurdo.

Definición (Producto escalar)

Sea E un espacio vectorial, se dice que $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$ es un **producto escalar** si es una forma bilineal definida positiva, es decir, cumple las propiedades:

- Definida positiva: $\langle x, x \rangle \geq 0 : \forall x \in E$
- No degenerada: $\langle x, x \rangle = 0 \Leftrightarrow x = 0$
- Homogeneidad: $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle : \forall x, y \in E$
- Bilinealidad: $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle : \forall x, y, z \in E$
- Simétrica: $\langle x, y \rangle = \langle y, x \rangle : \forall x, y \in E$

Al par $(E, \langle \cdot, \cdot \rangle)$ se le denomina espacio vectorial con producto escalar o espacio **pre-Hilbert**.

Ejemplo:

- En \mathbb{R}^n , definimos

$$\langle x, y \rangle = \sum_{j=1}^n x_j \cdot y_j : \forall x, y \in \mathbb{R}^n$$

- En $\mathcal{C}[0, 1]$, definimos

$$\langle f, g \rangle = \int_0^1 f(x) \cdot g(x) dx$$

Y es muy sencillo demostrar que ambas definiciones suponen un producto escalar en el espacio en el que están definidas.

Teorema (Desigualdad de Cauchy-Schwarz)

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio pre-Hilbert y sean $x, y \in E$ dos vectores cualesquiera. Entonces ocurre que:

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \cdot \langle y, y \rangle}$$

Demostración:

- Si $x = 0$ o $y = 0$, la desigualdad es trivial.
- Tomemos en primer lugar un $\lambda \in \mathbb{R}$ arbitrario y escojamos $x, y \in E \setminus \{0\}$:

$$0 \leq \langle \lambda x + y, \lambda x + y \rangle = \lambda^2 \langle x, x \rangle + 2\lambda \langle x, y \rangle + \langle y, y \rangle$$

Esta ecuación² describe una parábola que, a lo sumo, es tangente al eje X pero nunca lo llega a cruzar porque es siempre ≥ 0 . Consecuentemente, el discriminante de esta ecuación nunca será estrictamente positivo ya que esto implicaría tener dos raíces, es decir:

$$\Delta = 4\langle x, y \rangle^2 - 4\langle x, x \rangle \cdot \langle y, y \rangle \leq 0 \Leftrightarrow \langle x, y \rangle^2 - \langle x, x \rangle \cdot \langle y, y \rangle \leq 0 \Leftrightarrow |\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \cdot \langle y, y \rangle}$$

Observación:

Cabe destacar que hay igualdad si y solo si los vectores son proporcionales, es decir:

$$|\langle x, y \rangle| = \sqrt{\langle x, x \rangle \cdot \langle y, y \rangle} \Leftrightarrow x = \alpha y : \alpha \in \mathbb{R}$$

Si son proporcionales es trivial demostrar la igualdad. En cambio, si tenemos la igualdad, entonces ello implica que la parábola de la que hablábamos antes corta en un único punto al eje de abscisas, luego:

$$\langle \lambda \cdot x + y, \lambda \cdot x + y \rangle = 0 \Leftrightarrow \lambda^2 \langle x, x \rangle + 2\lambda \langle x, y \rangle + \langle y, y \rangle = 0 \Leftrightarrow \lambda = \frac{-2 \cdot \langle x, y \rangle}{2 \cdot \langle x, x \rangle} \Leftrightarrow x = -\lambda y$$

Proposición

Dado un espacio pre-Hilbert $(E, \langle \cdot, \cdot \rangle)$ y definimos $\|x\| = \sqrt{\langle x, x \rangle}$, entonces $(E, \|\cdot\|)$ es normado.

Demostración:

La única propiedad no trivial es la desigualdad triangular. Sea

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

Aplicando la desigualdad de Cauchy-Schwarz

$$\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

Ejemplo:

- En \mathbb{R}^n , si $x, y \in \mathbb{R}^n$, entonces:

$$\sum_{j=1}^n |x_j \cdot y_j| \leq \sqrt{\sum_{j=1}^n x_j^2} \cdot \sqrt{\sum_{j=1}^n y_j^2}$$

- Sea $E = \ell^2(\mathbb{N}) = \{x = \{x_j\} \in \mathbb{R} : \|x\|_{\ell^2}^2 = \sum_{n=1}^{\infty} x_j^2 < \infty\}$. Definimos

$$\langle x, y \rangle = \sum_{j=1}^{\infty} x_j y_j$$

Comprobemos que se encuentra bien definida:

$$\sum_{j=1}^{\infty} |x_j y_j| \leq \sqrt{\sum_{j=1}^n x_j^2} \cdot \sqrt{\sum_{j=1}^n y_j^2} \leq \sqrt{\sum_{j=1}^{\infty} x_j^2} \cdot \sqrt{\sum_{j=1}^{\infty} y_j^2} < \infty$$

$$\Rightarrow \sum_{j=1}^{\infty} |x_j| |y_j| \leq \|x\|_{\ell^2} \cdot \|y\|_{\ell^2}$$

²Esto es mayor o igual que 0 por la propiedad de definida positiva

- Sea $\mathcal{C}[0, 1]$

$$\int_0^1 |f(x) \cdot g(x)| \, dx \leq \left(\int_0^1 |f(x)|^2 \, dx \right)^{\frac{1}{2}} \cdot \left(\int_0^1 |g(x)|^2 \, dx \right)^{\frac{1}{2}}$$

Definición (Equivalencia de normas)

Se dice que dos normas son equivalentes en un espacio vectorial E , $\|\cdot\|_1 \approx \|\cdot\|_2$, si

$$\exists c_1, c_2 > 0 : c_1 \|x\|_2 \leq \|x\|_1 \leq c_2 \|x\|_2 : \forall x \in E$$

Observación:

Dado un espacio métrico (E, d) , una pregunta razonable es cuándo existe una norma que genera d , tal que $d(x, y) = \|x - y\|$. Esto ocurre cuando la métrica cumple dos condiciones:

- $d(x + z, y + z) = d(x, y)$ (Invariante por traslaciones)
- $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ (Invariante por dilataciones)

Teorema (Ley del Paralelogramo)

Sea $(E, \langle \cdot, \cdot \rangle)$ un espacio vectorial normado, cuya norma es $\|x\|^2 = \langle x, x \rangle$, se cumple que:

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

Si se diese que $x \perp y \Rightarrow \|x + y\| = \|x - y\|$, es decir, obtenemos el Teorema de Pitágoras:

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

Es interesante observar que el recíproco también es cierto, es decir, dado un espacio normado $(E, \|\cdot\|)$, si $\|\cdot\|$ satisface la Ley del Paralelogramo, entonces existe un producto escalar $\langle \cdot, \cdot \rangle$ tal que $\|x\|^2 = \langle x, x \rangle$. Concretamente es:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2)$$

TOPOLOGÍA EN ESPACIOS MÉTRICOS

Igual que \mathbb{R} necesitábamos definir las nociones de entorno, intervalo, supremo... es necesario hacer extensible dichos conceptos a espacios métricos para ser capaces de referirnos a conceptos como las proximidades de un punto o decir que un conjunto es cerrado.

Definición (Bola)

Sea el espacio métrico (E, d) . Se define la **bola abierta** de centro $x \in E$ y radio $r > 0$ al conjunto:

$$B(x, r) = \{y \in E : d(x, y) < r\}$$

Asimismo, definimos como **bola cerrada** al conjunto:

$$\bar{B}(x, r) = \{y \in E : d(x, y) \leq r\}$$

Ejemplo:

- En \mathbb{R} , tomando $x = \frac{a+b}{2}$ donde $a < b$ y $a, b \in \mathbb{R}$, elegimos $r = \frac{b-a}{2} > 0$. Luego ocurre que:

$$B(x, r) = \left\{ y \in \mathbb{R} : \left| y - \frac{a+b}{2} \right| < \frac{b-a}{2} \right\} = (a, b)$$

- Sea (X, d) , siendo d la métrica discreta $d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$, tenemos entonces que:

$$B(x, 2) = X$$

$$B(x, 1) = \{x\}$$

$$\bar{B}(x, 1) = X$$

Definición (Conjunto abierto)

Sea (X, d) un espacio métrico, se dice que un conjunto $A \subset X$ es **abierto** si:

$$\forall x \in A : \exists \varepsilon > 0 : B(x, \varepsilon) \subset A$$

Observación:

- Si $\varepsilon_1 < \varepsilon_2 \Rightarrow B(x, \varepsilon_1) \subset B(x, \varepsilon_2)$
- El conjunto vacío \emptyset y X son abiertos.
- El intervalo $(0, 1]$ no es abierto en \mathbb{R} , ya que para $x = 1, \nexists \varepsilon > 0 : B(1, \varepsilon) \subset (0, 1]$
- Consideramos el espacio métrico $((0, 1], d_2)$. El conjunto $(0, 1]$ es abierto en este espacio métrico determinado. Es decir, que el hecho de ser abierto o no depende del espacio métrico en el que nos encontramos.
- $\{x\}$ no es abierto en \mathbb{R} . Sin embargo, $\{x\}$ sí es abierto en $(\mathbb{R}, d_{\text{discreta}})$

Proposición

En un espacio métrico (E, d) , toda bola abierta es un abierto.

Demostración:

Sean $x \in E$ y $r > 0$ tomamos $y \in B(x, r)$. Queremos probar que $\exists \varepsilon > 0 : B(y, \varepsilon) \subset B(x, r)$, luego sea $\varepsilon = r - d(x, y) > 0$ vamos a ver que $B(y, \varepsilon) \subset B(x, r)$. Para ello sea $z \in B(y, \varepsilon)$ comprobemos que $d(x, z) < r$:

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \varepsilon = r$$

Proposición

Sea el espacio métrico (E, d) , entonces³ la apertura se conserva por intersecciones y uniones:

1. $\forall A_1, \dots, A_n$ familia de conjuntos abiertos en E , $\bigcap_{j=1}^n A_j$ es abierto en E .
2. $\forall \{A_\alpha\}_{\alpha \in I}$ familia arbitraria de conjuntos abiertos en E , $\bigcup_{\alpha \in I} A_\alpha$ es abierto en E .

Demostración:

- Tomamos $x \in \bigcap_{j=1}^n A_j$ para intentar probar que:

$$\exists \varepsilon > 0 : B(x, \varepsilon) \subset \bigcap_{j=1}^n A_j$$

Para ello, la idea es que fijado un j tendremos un ε_j que valdrá, luego como es un conjunto finito podemos quedarnos con el más pequeño y valdrá para todos los demás:

$$j \in \{1, \dots, n\} \Rightarrow \exists \varepsilon_j > 0 : B(x, \varepsilon_j) \subset A_j \Rightarrow \varepsilon = \min_{j \in \{1, \dots, n\}} \{\varepsilon_j\} > 0$$

³Cabe destacar que la 1ª afirmación es solo para familias finitas

De esta forma ocurre que:

$$B(x, \varepsilon) \subset \bigcap_{j=1}^N B(x, \varepsilon_j) \subset \bigcap_{j=1}^N A_j$$

- Sea $x \in \bigcup_{\alpha \in I} A_\alpha \Rightarrow \exists \beta \in I : x \in A_\beta$, luego se tiene que:

$$\exists \varepsilon > 0 : B(x, \varepsilon) \subset A_\beta \subset \bigcup_{\alpha \in I} A_\alpha$$

Cabe destacar que la condición de intersección finita es indispensable, puesto que en (\mathbb{R}^n, d_2) , $A_j = \left(-\frac{1}{j}, \frac{1}{j}\right)$ es abierto, pero $\bigcap_{j=1}^{\infty} A_j = \{0\}$ **no** es abierto en \mathbb{R} .

Definición (Topología)

Toda familia de subconjuntos T de un conjunto X , $T \subset \mathcal{P}(X)$, que satisface:

- $\emptyset, X \in T$
- T es invariante por intersecciones finitas.
- T es invariante por uniones arbitrarias.

Se denomina como **topología de X** .

Así, la familia de abiertos de un espacio métrico (X, d) es una topología.

Definición (Punto interior)

Dado un espacio métrico (E, d) , se dice que x es un **punto interior** de $A \subset E$ si cumple que:

$$\exists \varepsilon > 0 : B(x, \varepsilon) \subset A$$

Denotamos como $\mathring{A} = \text{int}(A) = \{x \in A : x \text{ es punto interior}\}$

Proposición (Caracterización de abierto)

Dado un espacio métrico (E, d) , $A \subset E$ es abierto si y sólo si:

$$A = \mathring{A}$$

O lo que es lo mismo, todos los puntos son interiores.

Observación:

- $\mathring{A} \subset A$
- $x \in \mathring{A} \Leftrightarrow \exists U \text{ abierto} : x \in U \subset A$
- $A = \{0\} \Rightarrow \mathring{A} = \emptyset$
- $A = [0, 1] \Rightarrow \mathring{A} = (0, 1)$
- $\mathring{A} = \bigcup_{U \subset A} U$ donde U es abierto.

Definición

En un espacio métrico (E, d) , decimos que un conjunto $A \subset E$ es cerrado si $E \setminus A$ es abierto.

Ejemplo:

- \emptyset, E son cerrados.
- En \mathbb{R} , $(0, 1]$ no es abierto ni cerrado.
- En $(0, 2]$ el conjunto $(0, 1]$, es cerrado puesto que su complementario es $(0, 2] \setminus (0, 1] = (1, 2]$ es abierto en este espacio ambiente.
- En \mathbb{R}^2 los puntos son cerrados puesto que si tomamos $A = \{x\}$, entonces $\mathbb{R} \setminus A$ es abierto. Tomamos $y \in \mathbb{R}$ y $\varepsilon = \frac{\|x-y\|_2}{2}$ entonces $x \notin B(y, \varepsilon)$, luego es abierto.

Proposición

En cualquier espacio métrico (E, d) , las bolas cerradas $\bar{B}(x, \varepsilon)$ son conjuntos cerrados.

Demostración:

Se reduce todo a ver que el complementario es abierto. Sea $z \in E \setminus \bar{B}(x, \varepsilon)$, entonces $d(z, x) > \varepsilon$. Si escogemos $0 < \delta < d(z, x) - \varepsilon$, entonces ¿se podrá verificar que $B(z, \delta) \subset E \setminus \bar{B}(x, \varepsilon)$? o lo que es lo mismo, que la intersección de esta bola de centro z con la bola cerrada inicial es el vacío:

$$y \in B(z, \delta) \Rightarrow d(y, x) \geq d(x, z) - d(z, y) > d(x, z) - \delta > \varepsilon \Rightarrow y \notin \bar{B}(x, \varepsilon)$$

Luego ya tenemos lo que queríamos probar

Proposición

Sea un espacio métrico (E, d) se cumple que:

1. Si $\{A_\alpha\}_{\alpha \in I}$ son cerrados, entonces: $\bigcap_{\alpha \in I} A_\alpha$ cerrado
2. Si $\{A_\alpha\}_{j=1, \dots, n}$ son cerrados, entonces: $\bigcup_{j=1}^n A_n$ cerrado

Demostración:

Se demuestra por las Leyes de De Morgan

Definición (Punto de acumulación)

Sea (E, d) un espacio métrico, y sea $A \subset E$, se dice que $x \in E$ es un **punto de acumulación** de A si:

$$\forall \varepsilon > 0 : (B(x, \varepsilon) \setminus \{x\}) \cap A \neq \emptyset$$

Al **conjunto de puntos de acumulación** se le conoce como A' .

Observación:

- Sea $A = (0, 1)$ en \mathbb{R} , entonces $A' = [0, 1]$, de lo que se desprende que no tiene por qué darse $A' \subset A$
- $A = \{1\}$ en \mathbb{R} , entonces $A' = \emptyset$, de lo que se desprende que no tiene por qué darse $A \subset A'$
- $x \in A' \Leftrightarrow \forall U$ abierto $\subset E$, $x \in U \Rightarrow (U \setminus \{x\}) \cap A \neq \emptyset$
- La acumulación en un espacio métrico con la métrica discreta es siempre el vacío. En efecto, tomamos (E, d) siendo d la métrica discreta y un conjunto $A \subset E$, entonces:

$$x \in A' \Leftrightarrow \forall \varepsilon > 0 : (B(x, \varepsilon) \setminus \{x\}) \cap A \neq \emptyset$$

Con lo cual, si tomamos $0 < \varepsilon \leq 1$, entonces

$$B(x, \varepsilon) \setminus \{x\} = \emptyset \Rightarrow A' = \emptyset$$

Proposición (Caracterización de puntos de acumulación)

Podemos caracterizar los puntos de acumulación como:

$$x \in A' \Leftrightarrow \forall \varepsilon > 0 : \exists x_\varepsilon \in A \setminus \{x\} : d(x_\varepsilon, x) < \varepsilon$$

En particular, en \mathbb{R}^n :

$$x \in A' \Leftrightarrow \exists \{x_k\}_{k \in \mathbb{N}} \subset A \setminus \{x\} : \|x - x_k\|_2 \xrightarrow{k \rightarrow \infty} 0$$

Proposición (Caracterización de los cerrados)

Sea (E, d) un espacio métrico y $A \subset E$, entonces ocurre que:

$$A' \subset A \Leftrightarrow A \text{ cerrado}$$

Demostración:

■ \Leftarrow :

Si A es cerrado, entonces $E \setminus A$ es abierto. De esta forma, queremos decir que

$$\begin{aligned} \forall x \notin A : \exists \varepsilon > 0 : B(x, \varepsilon) \subset (E \setminus A) &\Rightarrow B(x, \varepsilon) \cap A = \emptyset \Rightarrow B(x, \varepsilon) \setminus \{x\} \cap A = \emptyset \Rightarrow \\ &\Rightarrow x \notin A' \Rightarrow A' \subset A \end{aligned}$$

■ \Rightarrow :

Supongamos ahora que $A' \subset A$ veamos que el complementario es abierto, luego hay que ver que cualquier punto suyo es interior.

Sea $x \in E \setminus A$, luego $x \notin A'$. Esto quiere decir que:

$$\exists \varepsilon > 0 : (B(x, \varepsilon) \setminus \{x\}) \cap A = \emptyset \Rightarrow B(x, \varepsilon) \cap A = \emptyset \Rightarrow B(x, \varepsilon) \subset (E \setminus A)$$

Definición (Adherencia)

Sea (E, d) un espacio métrico y $A \subset E$, definimos **la adherencia** de A como:

$$\bar{A} = \bigcap_{A \subset F} F \text{ donde } F \text{ es cerrado}$$

Es decir, es el menor cerrado en el cual A está contenido.

Observación:

- Esto siempre está bien definido porque $F = E \supset A$ y E es cerrado.
- $\bar{A} \supset A$
- $A \subset B \Rightarrow \bar{A} \subset \bar{B}$

Ejemplo:

- $A = (0, 1)$ en \mathbb{R} , entonces $\bar{A} = [0, 1]$
- $A = \mathbb{Q}$ en \mathbb{R} , entonces $\bar{A} = \mathbb{R}$

Proposición (Caracterización de la adherencia)

Sea (E, d) un espacio métrico y $A \subset E$, entonces:

$$\bar{A} = A \cup A'$$

Demostración:

Sea $B = A \cup A'$, vamos a probar que $B = \bar{A}$.

■ $B \supset \bar{A}$

Vamos a probar que B es un conjunto cerrado, puesto que si probamos esta propiedad se tiene que $A \subset B \Rightarrow \bar{A} \subset \bar{B} = B$ y quedaría demostrada la primera inclusión:

$$x \in B^c = A^c \cap (E \setminus A') \Rightarrow x \notin B \Rightarrow x \notin A \wedge x \notin A'$$

En primer lugar, por ocurrir que $x \notin \bar{A}$, se tiene que:

$$\exists \varepsilon > 0 : (B(x, \varepsilon) \setminus \{x\}) \cap A = \emptyset$$

Luego usando esa afirmación junto con que $x \notin A$, tenemos que:

$$B(x, \varepsilon) \cap A = \emptyset \Rightarrow B(x, \varepsilon) \subset A^c$$

Queda solo por probar que $B(x, \varepsilon) \subset E \setminus A'$, veámoslo:

$$\begin{aligned} z \in B(x, \varepsilon) &\Rightarrow \exists \delta > 0 : x \notin B(z, \delta) \subset B(x, \varepsilon) \Rightarrow \\ &\Rightarrow (B(z, \delta) \setminus \{z\}) \cap A \subset (B(x, \varepsilon) \setminus \{x\}) \cap A = \emptyset \Rightarrow z \notin A' \Rightarrow \bar{A} \subset B \end{aligned}$$

■ $B \subset \bar{A}$

Como $x \in A \Rightarrow x \in \bar{A}$, todo se reduce a probar que $A' \subset \bar{A}$:

$$A' \subset \bar{A} \Leftrightarrow \forall F \supset A \text{ cerrado}, A' \subset F \Leftrightarrow \forall F \supset A \text{ cerrado}, (E \setminus A') \supset F^c$$

Luego vamos a intentar tratar de probar dicho enunciado equivalente:

$$\begin{aligned} z \notin F &\Rightarrow z \in F^c \Rightarrow \exists \varepsilon > 0 : B(z, \varepsilon) \subset F^c \Rightarrow (B(z, \varepsilon) \setminus \{z\}) \cap A \subset B(z, \varepsilon) \cap A = \emptyset \Rightarrow \\ &\Rightarrow z \notin A' \Rightarrow z \in E \setminus A' \end{aligned}$$

Corolario (Caracterización de los puntos de adherencia)

Si un punto pertenece a la adherencia, entonces cualquier bola centrada en ese punto interseca con el conjunto:

$$x \in \bar{A} \Leftrightarrow \forall \varepsilon > 0 : B(x, \varepsilon) \cap A \neq \emptyset$$

Definición (Densidad)

Sea (E, d) un espacio métrico y $A \subset B \subset E$, se dice que:

$$A \text{ es denso en } B \Leftrightarrow B \subset \bar{A}$$

En particular, A es denso en E si $\bar{A} = E$.

Ejemplo:

- El conjunto \mathbb{Q} es denso dentro de \mathbb{R} .
- El conjunto \mathbb{I} es denso dentro de \mathbb{R} .

Definición (Frontera)

Sea (E, d) un espacio métrico y $A \subset E$, se define la frontera de A como:

$$\partial(A) = \bar{A} \cap \overline{E \setminus A} = \bar{A} \setminus \overset{\circ}{A}$$

Observación:

- $\partial(A)$ es cerrado.
- $A = [0, 1]$ en \mathbb{R} , implica que $\bar{A} = [0, 1]$, luego $\partial(A) = \{0, 1\}$
- $\partial(A) = \partial(A^c)$
- Ejercicio: probar que $\partial(A) = \bar{A} \setminus \overset{\circ}{A} = \bar{A} \cap (E \setminus \overset{\circ}{A})$. Pista: probar que $E \setminus \overset{\circ}{A} = \overline{E \setminus A}$
- Ejercicio: probar que A' es cerrado.
- Ejercicio: estudiar cómo son los abiertos y los cerrados en un espacio con la métrica discreta.

Proposición (Caracterización de la frontera)

Si un punto pertenece a la frontera, entonces cualquier bola centrada en él interseca con el conjunto y con el complementario del conjunto:

$$p \in \partial(A) \Leftrightarrow \forall \varepsilon : B(p, \varepsilon) \cap A \neq \emptyset \text{ y } B(p, \varepsilon) \cap A^c \neq \emptyset$$

Aclaraciones de Clase:

- Vamos a poner un ejemplo donde $\bar{B}(x, r) \neq \overline{B(x, r)}$, tomamos (E, d) espacio métrico donde d es la métrica discreta. Escogemos $x \in E$ y $r = 1$, tenemos que $B(x, 1) = \{x\} \Rightarrow \overline{B(x, 1)} = \{x\}$, pero vemos que $\bar{B}(x, 1) = E \neq \{x\}$.
- Veamos que no siempre $Diam(A) = Diam(\partial A)$, si tomamos $E = [0, 1]$, entonces $\partial E = \emptyset$, luego $Diam(E) = 1 \neq Diam(\emptyset)$. Si tomamos \mathbb{R} y $A = (0, \infty)$, vemos que $\partial A = \{0\}$, luego $Diam(A) = \infty$ pero $Diam(\partial A) = 0$. De nuevo, si tomamos $E = [0, 1]$ y $A = (0, 1]$, tenemos que $\partial(A) = \{0\}$ y se tiene que $Diam(A) = 1 \neq 0 = Diam(\partial(A))$
- Veamos que $Diam(A) = Diam(\bar{A})$, como $A \subset \bar{A}$ entonces $Diam(A) \leq Diam(\bar{A})$. Y la otra desigualdad es muy fácil de probar también.
- Una buena propiedad es $(A \cap B)^\circ = \overset{\circ}{A} \cap \overset{\circ}{B}$. Elemental la demostración.

SUCESIONES COMPLETITUD Y COMPACIDAD

Convergencia y sucesiones sobre espacios métricos

En Análisis de Variable Real definimos los conceptos de sucesiones y convergencia para el espacio métrico $(\mathbb{R}, |\cdot|)$, por tanto, es necesario redefinir dichos conceptos y prescindir de las particularidades de \mathbb{R} para poder hacerlos extensible a cualquier espacio métrico en el que trabajemos.

Definición (Sucesión y Subsucesión)

Sea E un conjunto. Una **sucesión** $\{x_n\}_{n \in \mathbb{N}} = \{x_n\}_n = \{x_n\}$ es una función

$$\begin{aligned} f : \mathbb{N} &\rightarrow E \\ n &\mapsto x_n \end{aligned}$$

Una **subsucesión** o **sucesión parcial** de $\{x_n\}_{n \in \mathbb{N}}$ es una sucesión:

$$\{x_{n_j}\}_{j \in \mathbb{N}} \subset \{x_n\}_{n \in \mathbb{N}} \text{ donde } n_1 < n_2 < \dots < n_j$$

Definición (Convergencia)

Sea (E, d) un espacio métrico, y sea $\{x_n\}_{n \in \mathbb{N}}$ una sucesión de E , se dice que $\{x_n\}_{n \in \mathbb{N}}$ **converge** a $x \in E$ si:

$$\forall \mathcal{U} \text{ abierto } \subset E \text{ tal que } x \in \mathcal{U}, \exists n_0 \in \mathbb{N} : x_n \in \mathcal{U} : \forall n \geq n_0$$

Y entonces, escribimos $\lim_{n \rightarrow \infty} x_n = x$.

Esta definición que se ha dado de convergencia⁴ es equivalente a las siguientes caracterizaciones:

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} : \forall n \geq n_0 : x_n \in B(x, \varepsilon)$$

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} : \forall n \geq n_0 : d(x, x_n) < \varepsilon$$

Proposición (Operaciones con los límites)

Sea $(E, \|\cdot\|)$ un espacio normado, si suponemos que dicha métrica anterior viene dada por la norma especificada ahora, entonces:

- Si $\lim_{n \rightarrow \infty} x_n = x$ y $\lim_{n \rightarrow \infty} y_n = y \Rightarrow \lim_{n \rightarrow \infty} (x_n + y_n) = x + y$
- Sea $\{\lambda_n\} \subset \mathbb{R}$ y $\lambda \in \mathbb{R}$ donde $\lambda_n \xrightarrow{n \rightarrow \infty} \lambda$ y $x_n \xrightarrow{n \rightarrow \infty} x \in E$, entonces se tiene que:
 $\lambda_n \cdot x_n \xrightarrow{n \rightarrow \infty} \lambda \cdot x \in E$

Demostración:

El caso $x = 0$ se deja como ejercicio. Sea $x \neq 0$ y $\varepsilon > 0$. Entonces:

$$\exists n_1, n_2 \in \mathbb{N} : |\lambda_n - \lambda| < \frac{\varepsilon}{2 \cdot \|x\|} : \forall n \geq n_1 \text{ y } \|x_n - x\| < \frac{\varepsilon}{2 \cdot C} : \forall n \geq n_2 : |\lambda_n| \geq C < \infty$$

Sea $n_0 = \max\{n_1, n_2\}$ y $n \geq n_0$

$$\|\lambda_n x_n - \lambda x\| = \|\lambda_n x_n - \lambda_n x + \lambda_n x - \lambda x\| \leq |\lambda_n| \cdot \|x_n - x\| + |\lambda_n - \lambda| \cdot \|x\| \leq C \cdot \frac{\varepsilon}{2C} + \frac{\varepsilon}{2\|x\|} \cdot \|x\| = \varepsilon$$

Proposición (Intercambiabilidad de normas)

Si $\|\cdot\|_1, \|\cdot\|_2$ son normas equivalentes en E , entonces ambas normas tienen las mismas sucesiones convergentes y el límite coincide en ambas.

$$0 \leftarrow C \|x_n - x\|_1 \geq \|x_n - x\|_2 \rightarrow 0$$

Proposición (Convergencia en \mathbb{R}^n)

En $(\mathbb{R}^n, \|\cdot\|_2)$ espacio euclídeo usual, sea $\{x_m\}_{m \in \mathbb{N}} : x_m = (x_m^1, \dots, x_m^n)$ la convergencia se caracteriza como:

$$x_m \xrightarrow{m \rightarrow \infty} x = (x^1, \dots, x^n) \Leftrightarrow \forall j = 1, \dots, n : x_m^j \xrightarrow{m \rightarrow \infty} x^j$$

Es decir, se reduce la convergencia en \mathbb{R}^n a la convergencia de sus componentes en \mathbb{R} .

Demostración:

Trivialmente, se tiene que $|x_m^j - x^j| \leq \|x_m - x\|_2$, luego :

$$x_m \rightarrow x \Rightarrow |x_m^j - x^j| \leq \|x_m - x\|_2 \Rightarrow x_m^j \rightarrow x^j : \forall j = 1, \dots, n$$

⁴Si $\{x_n\}_{n \in \mathbb{N}}$ converge, el límite es único.

Recíprocamente, podemos expresar la norma como:

$$\|x_m - x\|_2^2 = \sum_{j=1}^n |x_m^j - x^j|^2$$

De este modo, por lo probado en la proposición anterior:

$$\lim_{m \rightarrow \infty} \|x_m - x\|_2^2 = \sum_{j=1}^n \lim_{m \rightarrow \infty} |x_m^j - x^j|^2 = 0$$

Proposición

Sea (E, d) un espacio métrico y $\{x_n\}_{n \in \mathbb{N}} \subset E$, podemos caracterizar de nuevo la convergencia en base a subsucesiones:

- $x_n \xrightarrow{n \rightarrow \infty} x \Leftrightarrow \forall \{x_{n_j}\}_{j \in \mathbb{N}} \subset \{x_n\}_{n \in \mathbb{N}} : x_{n_j} \xrightarrow{n \rightarrow \infty} x$
- $x_n \xrightarrow{n \rightarrow \infty} x \Leftrightarrow \forall \{x_{n_j}\}_{j \in \mathbb{N}} \subset \{x_n\}_{n \in \mathbb{N}} : \exists \{y_{m_k}\}_{k \in \mathbb{N}} \subset \{x_{n_j}\}_{j \in \mathbb{N}} : y_{m_k} \xrightarrow{n \rightarrow \infty} x$

Demostración:

- \Rightarrow : Tenemos que $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} : \forall n \geq n_0 : d(x_n, x) < \varepsilon$, luego entonces si consideramos $n_{j_0} \geq n_0 : \forall n_j \geq n_{j_0} \geq n_0 : d(n_j, x) < \varepsilon$ se tiene trivialmente.
 \Leftarrow : Trivial.
- \Rightarrow : Trivial por 1.
 \Leftarrow : Supongamos que $x_n \not\rightarrow x$, entonces $\exists \varepsilon > 0 : \forall m \in \mathbb{N}, \exists n_m \geq m : d(x_{n_m}, x) \geq \varepsilon$. Si consideramos ahora $\{x_{n_m}\}_m \subset \{x_n\}_n$ y escogemos una subsucesión de esta $\{y_{m_k}\}_k \subset \{x_{n_m}\}_m : d(y_{m_k}, x) \geq \varepsilon \Rightarrow y_{m_k} \not\rightarrow x \#$

Definición (Sucesión de Cauchy)

Sea (E, d) un espacio métrico, se dice que $\{x_n\}_{n \in \mathbb{N}} \subset E$ es **sucesión de Cauchy** si:

$$\forall \varepsilon > 0 : \exists n_0 \in \mathbb{N} : \forall m, n \geq n_0 \Rightarrow d(x_m, x_n) < \varepsilon$$

Se dice que (E, d) es **completo** si toda sucesión de Cauchy es convergente.

Observación:

- $(\mathbb{R}, |\cdot|)$ es completo.
- $(\mathbb{Q}, |\cdot|)$ no es completo.
- Sea $E = \mathbb{R} \setminus \{0\}$ no es completo. Por ejemplo, la sucesión $x_n = \frac{1}{n} \in E$ es sucesión de Cauchy y no converge en E .

Concepto de compacidad y propiedades

En este apartado, vamos a definir un concepto muy provechoso que hasta ahora había quedado subyacente solo para los intervalos cerrados y acotados en \mathbb{R} que es la noción de compacidad. Todas las buenas propiedades que tenían los cerrados y acotados en \mathbb{R} como el Teorema de Bolzano-Wierstrass se generalizan a conjuntos compactos para espacios de mayor dimensión.

Definición (Conjunto Acotado)

En (E, d) un espacio métrico, un conjunto $A \subset E$ es **acotado** si:

$$\exists x \in E : \exists r > 0 : A \subset B(x, r) \Leftrightarrow \forall y \in A : d(y, x) < r$$

En el caso $(E, \|\cdot\|)$ espacio normado, $A \subset E$ es **acotado** si:

$$\exists r > 0 : \forall a \in A : \|a\| < r \Leftrightarrow A \subset B(0, r)$$

Proposición

Sea (E, d) un espacio métrico. Entonces, se cumple que:

- Si $\{x_n\}_{n \in \mathbb{N}} \subset E$ es convergente, entonces es una sucesión de Cauchy.
- Si $\{x_n\}_{n \in \mathbb{N}} \subset E$ es una sucesión de Cauchy, entonces es acotada.
- Si $\{x_n\}_{n \in \mathbb{N}} \subset E$ es convergente, entonces es acotada.
- Si $\{x_n\}_{n \in \mathbb{N}} \subset E$ es una sucesión de Cauchy y $\exists \{x_{n_k}\}_{n_k}$ parcial tal que $x_{n_k} \xrightarrow{n_k \rightarrow \infty} x$, entonces $x_n \xrightarrow{n \rightarrow \infty} x$.

Demostración:

- (i) y (iii) son triviales.
- Sea $\varepsilon = 1$. Entonces, $\exists n_0 \in \mathbb{N} : \forall m, n \geq n_0 : d(x_n, x_m) \leq 1$, luego se tiene que $\forall n \geq n_0 : d(x_n, x_{n_0}) \leq 1$. Por tanto, sea $r = \max\{1, d(x_{n_0}, x_k : k = 1, \dots, n_0 + 1)\} + 1$, se tiene que $\forall n \in \mathbb{N} x_n \in B(x_{n_0}, r)$.
- Sea $\varepsilon > 0$, entonces $\exists n_0 \in \mathbb{N} : \forall m, n \geq n_0 : d(x_n, x_m) < \frac{\varepsilon}{2}$. Del mismo modo, para el mismo ε $\exists m \in \mathbb{N} : n_k \geq m : d(x_{n_k}, x) < \frac{\varepsilon}{2}$. Luego, sea $k = \max\{n_0, m\} : \forall n \geq k$ se tiene que por la desigualdad triangular: $d(x_n, x) \leq d(x_{n_j}, x_n) + d(x_{n_j}, x)$. Elegimos $n_j \geq k$ y vemos que:

$$d(x_n, x) \leq d(x_{n_j}, x_n) + d(x_{n_j}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Teorema (de Completitud de \mathbb{R}^n)

El espacio euclídeo habitual (\mathbb{R}^n, d_2) es un espacio métrico completo.

Demostración:

Sea $\{x_m\}_{m \in \mathbb{N}}$ una sucesión de Cauchy en \mathbb{R}^n de la forma $x_m = (x_m^1, \dots, x_m^n) : m \in \mathbb{N}$, entonces:

$$|x_m^j - x_\ell^j| \leq \|x_m - x_\ell\|_2 < \varepsilon : m, \ell \geq n_0 : \forall j \in \{1, \dots, n\}$$

Así que $\forall j \in \{1, \dots, n\}$, la sucesión de la componente $\{x_m^j\}_{m \in \mathbb{N}}$ es sucesión de Cauchy en \mathbb{R} y como \mathbb{R} es completo, $x_m^j \xrightarrow{m \rightarrow \infty} x^j$. Por tanto, por la definición que se dio de convergencia en \mathbb{R}^n , si el límite es $x = (x^1, \dots, x^n) \in \mathbb{R}^n$, entonces componente a componente $x_m \xrightarrow{m \rightarrow \infty} x$

Definición (Conjunto Secuencialmente Compacto)

Sea (E, d) un espacio métrico. Se dice que $A \subset E$ es **secuencialmente compacto** si:

$$\forall \{x_n\}_{n \in \mathbb{N}} \subset A : \exists \{x_{n_k}\}_{n_k} \subset \{x_n\}_n : x_{n_k} \xrightarrow{k \rightarrow \infty} x \in A$$

Observación:

- Si A es secuencialmente compacto, entonces $A' \subset A$ y, por lo tanto, A es cerrado.

- Si A es secuencialmente compacto, entonces A es acotado (porque hay una bola que lo contiene completamente).

En caso contrario, ocurriría que $\forall \varepsilon > 0 : \forall x \in A : A \not\subset B(x, \varepsilon)$, pero si escogemos $\varepsilon = n$, entonces $A \not\subset B(x, \varepsilon)$, luego $\exists y_n \in A : d(y_n, x) \geq n$, es decir, podemos formar $\{y_n\}_{n \in \mathbb{N}} \subset A$.

Claramente $\{y_{n_k}\}_{n_k} \subset \{y_n\}_{n \in \mathbb{N}} \subset A$ convergente, ya que si existiese:

$$d(y_{n_k}, y) < \varepsilon : \forall n_k \geq m \Rightarrow d(y_m, y_n) \geq d(y_m, x) - d(x, y_n) \geq m - d(x, y_n) \xrightarrow{m \rightarrow \infty} \infty \Rightarrow \#$$

Por tanto, llegamos a una contradicción.

Proposición (Caracterización de la topología vía sucesiones)

Sea (E, d) un espacio métrico. Se cumple que:

- Un subconjunto $A \subset E$ es cerrado $\Leftrightarrow \forall \{x_n\}_{n \in \mathbb{N}} \subset A : (\exists \lim_{n \rightarrow \infty} x_n = x \Rightarrow x \in A)$
- Un punto $x \in \bar{A} \Leftrightarrow \exists \{x_n\}_{n \in \mathbb{N}} \subset A : \lim_{n \rightarrow \infty} x_n = x$
- Un punto $x \in A' \Leftrightarrow \exists \{x_n\}_{n \in \mathbb{N}} \subset A : x_n \neq x : \lim_{n \rightarrow \infty} x_n = x$

Demostración:

El conjunto A es cerrado si y sólo si $A' \subset A$. Por tanto, si escogemos una sucesión $\{x_n\}_{n \in \mathbb{N}} \subset A$ de forma que $\lim_{n \rightarrow \infty} x_n = x$, vamos a ver que $x \in A$. Supongamos que no, entonces tenemos una sucesión convergente a x de forma que $\forall n \in \mathbb{N} : x_n \neq x$ porque está contenida en A . En consecuencia, por el apartado 3: $x \in A' \subset A$.

Definición (Conjunto Compacto)

Sea (E, d) un espacio métrico:

- Se dice que un **recubrimiento abierto** de $A \subset E$ es una familia de abiertos en cuya unión se encuentra A :

$$\mathcal{U} = \{\mathcal{U}_i\}_{i \in I} : \mathcal{U}_i \text{ abierto} \wedge A \subset \bigcup_{i \in I} \mathcal{U}_i$$

- Un **subrecubrimiento finito** de A es una familia finita de intervalos abiertos en cuya unión se encuentra A :

$$\mathcal{U} = \{\mathcal{U}_1, \dots, \mathcal{U}_m\} \subset \mathcal{U} : A \subset \bigcup_{i=1}^m \mathcal{U}_i$$

- Se dice que $A \subset E$ es **compacto** si para todo recubrimiento abierto de A se puede extraer un subrecubrimiento finito.

Observación:

- Si A es compacto, entonces A es acotado.

$$\text{Sea } x \in A \subset \bigcup_{n \in \mathbb{N}} B(x, n) \xRightarrow{\exists m \in \mathbb{N}} A \subset B(x, m)$$

- $(0, 1)$ no es compacto en \mathbb{R} .

En efecto, sea $\mathcal{U}_n = (\frac{1}{n}, 1)$. Es claro que $(0, 1) \subset \bigcup_{n \in \mathbb{N}} \mathcal{U}_n$, pero $\nexists m : (0, 1) \subset \mathcal{U}_m$

Definición (Conjunto Totalmente Acotado)

Sea (E, d) un espacio métrico, se dice que $A \subset E$ es **totalmente acotado** si ocurre que:

$$\forall \varepsilon > 0 : \exists \{x_1, \dots, x_m\} \subset A : A \subset B(x_1, \varepsilon) \cup \dots \cup B(x_m, \varepsilon)$$

Observación:

- Si A es totalmente acotado, entonces A es acotado.

Sea $r = \max\{d(x_1, x_j) : j = 2, \dots, n\}$ y sea $R = r + 3\varepsilon$. Entonces $B(x_j, \varepsilon) \subset B(x_1, R) \Rightarrow A \subset B(x_1, R)$

- En \mathbb{R} con la métrica discreta, $\mathbb{R} = B(0, 2)$, es decir, \mathbb{R} es acotado. Sin embargo \mathbb{R} no es totalmente acotado.

Sea $\varepsilon = \frac{1}{2}$. La bola $B(x, \frac{1}{2}) = \{x\}$, es decir, $\mathbb{R} \not\subset B(x_1, \frac{1}{2}) \cup \dots \cup B(x_m, \frac{1}{2})$

Lema (Lema 1)

Sea (E, d) un espacio métrico, si A es compacto, entonces A es cerrado.

Demostración:

Hay que ver que $E \setminus A$ es abierto, es decir, que todo punto es interior: $\forall x \in E \setminus A : \exists \delta > 0 : B(x, \delta) \subset E \setminus A$.

Primero, sea $\mathcal{U}_n = E \setminus \bar{B}(x, \frac{1}{n})$ abierto. Vamos a probar que $A \subset \bigcup_{n \in \mathbb{N}} \mathcal{U}_n$. Tomamos $y \in A$, luego $d(y, x) > 0$ para $x \notin A$, luego $\exists n \in \mathbb{N} : d(y, x) > \frac{1}{n} \Rightarrow y \notin \bar{B}(x, \frac{1}{n}) \Rightarrow y \in \mathcal{U}_n$.

Como A es compacto, podemos extraer un subrecubrimiento finito, luego $A \subset \mathcal{U}_{n_1} \cup \dots \cup \mathcal{U}_{n_k} = \mathcal{U}_{n_k} : n_1 < \dots < n_k$ porque la sucesión de conjuntos es creciente.

$$\forall y \in A : y \in \mathcal{U}_{n_k} \text{ para } \exists n_k \Rightarrow y \notin \bar{B}\left(x, \frac{1}{n_k}\right) \Rightarrow \bar{B}\left(x, \frac{1}{n_k}\right) \cap A = \emptyset$$

Finalmente, $B\left(x, \frac{1}{n_k}\right) \subset E \setminus A$

Lema (Lema 2)

Sea (E, d) un espacio métrico compacto, si $A \subset E$ cerrado, entonces A es compacto.

Demostración:

Escojamos un recubrimiento de abiertos de A , es decir:

$$A \subset \bigcup_{i \in I} \mathcal{U}_i : \mathcal{U}_i \text{ abierto}$$

De este modo, podemos escribir el espacio ambiente como:

$$E \subset (E \setminus A) \cup \bigcup_{i \in I} \mathcal{U}_i$$

Como E es compacto, se puede extraer un subrecubrimiento finito $E \subset (E \setminus A) \cup \mathcal{U}_1 \cup \dots \cup \mathcal{U}_m$ y como $A \subset E$ y no está en su complementario, $A \subset \mathcal{U}_1 \cup \dots \cup \mathcal{U}_m$.

Lema (Lema 3)

Sea (E, d) un espacio métrico y $A \subset E$ secuencialmente compacto, entonces A es totalmente acotado.

Demostración:

Supongamos que A no es totalmente acotado, es decir:

$$\exists r > 0 : \forall \{x_1, \dots, x_m\} \subset A : A \not\subset B(x_1, r) \cup \dots \cup B(x_m, r)$$

Sea $z_1 \in A$, entonces $A \setminus B(z_1, r) \neq \emptyset$. Por tanto, podemos escoger $z_2 \in A \setminus B(z_1, r)$ y entonces $A \setminus (B(z_1, r) \cup B(z_2, r)) \neq \emptyset$. De modo análogo, podemos construir la sucesión:

$$z_n \in A \setminus (B(z_1, r) \cup \dots \cup B(z_{n-1}, r)) \neq \emptyset$$

Y, además, ocurre que:

$$\forall n \in \mathbb{N} : \{z_n\}_{n \in \mathbb{N}} \subset A \text{ donde } d(z_m, z_n) \geq r : m \neq n$$

Luego se tiene que $\{z_m\}_{m \in \mathbb{N}}$ no posee ninguna subsucesión convergente, lo cual contradice que A sea secuencialmente compacto.

Teorema (Bolzano - Weierstrass)

Sea (E, d) un espacio métrico:

$$A \subset E \text{ es compacto} \Leftrightarrow A \text{ es secuencialmente compacto}$$

Demostración:

■ \Rightarrow

Supongamos que A compacto. Por reducción al absurdo, si A no es secuencialmente compacto, entonces $\exists \{x_n\}_{n \in \mathbb{N}} \subset A$ sin parciales convergentes.

Sin pérdida de generalidad, suponemos que todos los términos son distintos unos de otros, es decir, $x_n \neq x_m : n \neq m$ porque si existiese un número infinito de términos con el mismo valor, entonces podríamos coger la parcial formada por dichos términos y sería convergente. Escogiendo el conjunto formado por los puntos de la sucesión $B = \{x_n\}_{n \in \mathbb{N}} \Rightarrow B' = \emptyset \subset B$, luego B es cerrado y como $B \subset A$ que es compacto, por el Lema 2, B es cerrado (y su complementario abierto), por lo que (si lo siguiente no ocurriese $B' \neq \emptyset$):

$$\forall n \in \mathbb{N}, \exists \varepsilon_n > 0 : B(x_n, \varepsilon_n) \cap (\{x_m\}_{m \in \mathbb{N}} \setminus \{x_n\}) = \emptyset$$

$B = \{x_n\}_{n \in \mathbb{N}} \subset \bigcup_{n \in \mathbb{N}} B(x_n, \varepsilon_n)$. Como B es compacto, $B \subset B(x_{n_1}, \varepsilon_{n_1}) \cup \dots \cup B(x_{n_m}, \varepsilon_{n_m})$, lo cual es una contradicción pues B no es un conjunto finito, por lo tanto A es secuencialmente compacto.

■ \Leftarrow

Supongamos ahora que A es secuencialmente compacto. Veamos que A es compacto, es decir, que si $A \subset \bigcup_{i \in I} \mathcal{U}_i$, entonces $A \subset \mathcal{U}_{i_1} \cup \dots \cup \mathcal{U}_{i_m}$.

Como A es secuencialmente compacto, es totalmente acotado por el Lema 3, entonces, dado $\varepsilon > 0$ sabemos⁵ que:

$$A \subset B(z_1, \varepsilon) \cup \dots \cup B(z_m, \varepsilon) : z_1, \dots, z_m \in A \Rightarrow A \subset \mathcal{U}_{i_{z_1}} \cup \dots \cup \mathcal{U}_{i_{z_m}}$$

Ejercicios

- Caracterizar los compactos de un espacio métrico con la distancia discreta
- Ver cuales son ciertos. Sean $A, B \begin{cases} \text{abiertos} \\ \text{cerrados} \\ \text{compactos} \end{cases}$. Entonces $A + B$ también lo es.

Proposición (Constante de Lebesgue)

Si A es secuencialmente compacto, para un recubrimiento $A \subset \bigcup_{i \in I} \mathcal{U}_i$ se tiene que:

$$\exists \varepsilon > 0 : \forall x \in A, \exists i_x \in I : B(x, \varepsilon) \subset \mathcal{U}_{i_x}$$

⁵Se aplica la propiedad de Lebesgue que se ve más adelante en este documento

Demostración:

En efecto, si suponemos lo contrario:

$$\forall n \in \mathbb{N} : \varepsilon_n = \frac{1}{n} : \exists x_n \in A : \forall i \in I : B\left(x_n, \frac{1}{n}\right) \not\subset \mathcal{U}_i$$

Como A es secuencialmente compacto, $\exists \{y_m\}_{m \in \mathbb{N}} \subset \{x_n\}_{n \in \mathbb{N}} : y_m \rightarrow y \in A$. Además, como $y \in A \subset \bigcup_{i \in I} \mathcal{U}_i$, entonces $y \in \mathcal{U}_{i_0}$, es decir, que $\exists \delta > 0 : B(y, \delta) \subset \mathcal{U}_{i_0}$ por ser este último abierto.

Sea $n_0 \in \mathbb{N} : \frac{1}{n_0} < \frac{\delta}{2}$ y $\forall n \geq n_0 : d(y_n, y) < \frac{\delta}{2}$ tenemos que la bola $B(y_n, \frac{1}{n}) \subset B(y, \delta) : \forall n \geq n_0$.

Sea $z \in B(y_n, \frac{1}{n})$, tenemos que:

$$d(z, y_n) < \frac{1}{n} : d(z, y) \leq d(z, y_n) + d(y_n, y) < \frac{1}{n} + \frac{\delta}{2} < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

Así, $B(x_m, \frac{1}{n}) = B(y_n, \frac{1}{n}) \subset B(y, \delta) \subset \mathcal{U}_{i_0}$, lo cual contradice que $B(x_n, \frac{1}{n}) \not\subset \mathcal{U}_i$

Observación:

- $[0, 1]$ es compacto en \mathbb{R} . En efecto, sea $\{x_n\}_{n \in \mathbb{N}} \subset [0, 1]$. Entonces, es una sucesión acotada y tiene por lo tanto una parcial convergente:

$$\{x_{n_k}\}_{n_k} \subset \{x_n\}_n, x_{n_k} \rightarrow x \in \mathbb{R}$$

Como $[0, 1]$ cerrado $\Rightarrow x \in [0, 1] \subset [0, 1]$.

Análogamente, para $\{x \in \mathbb{R}^n : \|x\| < 1\}$ es compacto.

- Sea (E, d) un espacio métrico. Decir que $A \subset E$ es compacto $\Leftrightarrow (A, d)$ es compacto.

Teorema

(E, d) es compacto $\Leftrightarrow E$ es completo y totalmente acotado

Demostración:

- \Rightarrow Como E es compacto, por el Teorema de Bolzano - Weierstrass E es secuencialmente compacto, y por el Lema 3 E es totalmente acotado.

Ahora, sea $\{x_n\}_{n \in \mathbb{N}} \subset E$ una sucesión de Cauchy. Por el Teorema de Bolzano - Weierstrass existe una parcial convergente, y por ser de Cauchy, la sucesión converge.

- \Leftarrow Veamos que es secuencialmente compacto para poder aplicar el Teorema de Bolzano - Weierstrass. Sea $\{x_n\}_n \subset E$, podemos suponer, sin pérdida de generalidad, que $x_n \neq x_m, n \neq m$ por el mismo argumento visto en anteriores demostraciones. Como E es totalmente acotado, $E \subset B(y_1^1, 1) \cup \dots \cup B(y_{m_1}^1, 1)$. Entonces,

$$\{x_n^1\}_{n \in \mathbb{N}} \subset \{x_n\}_{n \in \mathbb{N}}, \exists j : \{x_n^1\}_{n \in \mathbb{N}} \subset B(y_j^1, 1)$$

Iterativamente, $E \subset B(y_1^2, \frac{1}{2}) \cup \dots \cup B(y_{m_1}^2, \frac{1}{2})$. Entonces,

$$\{x_n^2\}_{n \in \mathbb{N}} \subset \{x_n^1\}_{n \in \mathbb{N}}, \exists j : \{x_n^2\}_{n \in \mathbb{N}} \subset B(y_j^2, \frac{1}{2})$$

En general, $\{x_n^m\} \subset \{x_n^{m-1}\}_n$ y $\{x_n^m\} \subset B(y_{j_m}^m, \frac{1}{m})$. Finalmente, sea

$$z_n = x_n^n : \{z_n\}_{n \in \mathbb{N}} \subset \{x_m\}_{m \in \mathbb{N}}$$

Dado $\varepsilon > 0$, sea $n_0 \in \mathbb{N} : \frac{1}{n_0} < \frac{\varepsilon}{2}$. Sean $m, n \geq n_0 : x_m^m = z_m, x_n^n = z_n \in B(y_{j_m}^{n_0}, \frac{1}{n_0}) \Rightarrow d(z_m, z_n) < \frac{2}{n_0} < \varepsilon$. Entonces, $\{z_m\}_m$ es una sucesión de Cauchy.

Como E es completo, $\{z_m\}_{m \in \mathbb{N}}$ es parcial de $\{x_n\}_{n \in \mathbb{N}}$ es convergente, por lo que E es secuencialmente compacto.

Corolario

Sea (E, d) espacio métrico completo. Entonces, $A \subset E$ es compacto $\Leftrightarrow A$ es cerrado y totalmente acotado.

Demostración:

- \Rightarrow Por la proposición anterior, basta probar que A es cerrado.
Sea $x \in A'$, y sea $\{x_n\}_{n \in \mathbb{N}} \subset A : x_n \xrightarrow{n \rightarrow \infty} x$ con $x_n \neq x$ hay que ver que $x \in A$. Como $\{x_n\}_{n \in \mathbb{N}}$ es sucesión de Cauchy en A y A es completo, $x_n \xrightarrow{n \rightarrow \infty} x \in A$.
- \Leftarrow Sabemos que E es completo y A es cerrado y totalmente acotado. Veamos que A es completo.
Sea $\{x_n\}_{n \in \mathbb{N}} \subset A$ una sucesión de Cauchy, entonces $\{x_n\}_{n \in \mathbb{N}} \subset E$ es sucesión de Cauchy y como E es completo ocurre que $x_n \xrightarrow{n \rightarrow \infty} x \in E$, pero además como A es cerrado por hipótesis se tiene $x \in A$.

Teorema (Heine - Borel)

$$A \subset \mathbb{R}^n \text{ compacto} \Leftrightarrow A \text{ es cerrado y acotado}$$

Demostración:

- \Rightarrow Como hemos probado que \mathbb{R}^n es completo, esta implicación es trivial a partir del corolario anterior.
- \Leftarrow . Sea A cerrado y acotado en \mathbb{R}^n , debemos ver que A es secuencialmente acotado.
Sea $\{x_m\}_{m \in \mathbb{N}} \subset A$, entonces por ser acotado se tiene que $\{x_m\}_{m \in \mathbb{N}}$ es acotada en \mathbb{R}^n . Por consiguiente, $\exists \{y_k\}_{k \in \mathbb{N}} \subset \{x_m\}_{m \in \mathbb{N}}$ parcial convergente en \mathbb{R}^n y entonces:

$$y_k \xrightarrow{n \rightarrow \infty} y \in \bar{A} = A \Rightarrow \text{compacto}$$

Ejercicio*: En \mathbb{R}^n , A es acotado $\Leftrightarrow A$ es totalmente compacto.

Ejercicio*: \mathbb{R}^n , A es acotado $\Leftrightarrow \bar{A}$ es compacto.

Teorema (Principio de Conjuntos Encajados)

Sea (E, d) espacio métrico y sea $\{A_n\}_{n \in \mathbb{N}}, A_n \subset E$ y $A_n \neq \emptyset$ compacto tal que $A_{n+1} \subset A_n$. Entonces,

$$\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$$

Demostración:

Sea $x_n \in A_n \subset A_m : m \in \{1, \dots, n\}$. Entonces, $\{x_n\}_{n \in \mathbb{N}} \subset A_1$, y como A_1 es compacto, $\exists \{x_{n_k}\}_{k \in \mathbb{N}}$ convergente $x_{n_k} \rightarrow x \in A_1$.

Fijado $m \in \mathbb{N}, \exists k_m \in \mathbb{N} : m_k \geq m, k \geq k_m$

$$\{x_{n_k}\}_{k \geq k_m} \text{ converge a } x$$

Además, $\{x_{n_k}\} \subset A_m : k \geq k_m$, y como A_m es compacto, $x \in A_m : \forall m \in \mathbb{N}$. Entonces,

$$x \in \bigcap_{m \in \mathbb{N}} A_m \neq \emptyset$$

CONTINUIDAD Y DIFERENCIABILIDAD

Es fundamental hacer extensibles los conceptos de continuidad y derivabilidad a dimensiones mayores que \mathbb{R} y más generalmente para espacios métricos, puesto que cuestiones como la optimización de funciones, la búsqueda de extremos relativos o propiedades de conservación topológicas dependen esencialmente de estos conceptos.

DEFINICIÓN DE CONTINUIDAD

En este apartado, redefinimos el concepto de continuidad de \mathbb{R} haciéndolo extensible ahora a cualquier espacio métrico. Estudiaremos más concretamente \mathbb{R}^n y veremos los resultados más notables en cuanto a topología que otorga la continuidad.

Definición (Límite)

Sean $(E, d), (F, d')$ espacios métricos, $A \subset E$ de forma que $f : A \rightarrow F$ y $x_0 \in A'$ un punto de acumulación, se dice que $L \in F$ es el **límite**¹ de f en x_0 si:

$$\forall \varepsilon > 0, \exists \delta > 0 : x \in A \cap (B(x_0, \delta) \setminus \{x_0\}) : d'(f(x), L) < \varepsilon$$

Y en ese caso escribimos que $\lim_{x \rightarrow x_0} f(x) = L$. De hecho, dicha definición es equivalente a:

$$\lim_{x \rightarrow x_0} f(x) = L \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0 : f(A \cap (B_d(x_0, \delta) \setminus \{x_0\})) \subset B_{d'}(L, \varepsilon)$$

Proposición (Caracterización del límite)

Si $f : A \subset E \rightarrow F$ y $x_0 \in A'$ es un punto de acumulación, entonces:

$$\lim_{x \rightarrow x_0} f(x) = L \Leftrightarrow \forall \{x_n\}_{n \in \mathbb{N}} \subset A \setminus \{x_0\} : (x_n \rightarrow x_0 \Rightarrow \lim_{n \rightarrow \infty} f(x_n) = L)$$

Definición (Continuidad)

Sean $(E, d), (F, d')$ espacios métricos, $A \subset E$ de forma que $f : A \rightarrow F$ y $x_0 \in A$ un punto del conjunto, se dice que f es **continua en** x_0 si:

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) \text{ si } x_0 \in A$$

En el caso en que $x_0 \notin A'$ se definirá continua por defecto.

¹En caso de existir, este debe ser único

Teorema (Caracterización de la continuidad)

Sean (E, d) , (F, d') espacios métricos, $A \subset E, f : A \rightarrow F$. Son equivalentes:

1. f es continua en A
2. Si $\{x_n\}_{n \in \mathbb{N}} \subset A : x_n \xrightarrow{n \rightarrow \infty} x_0 \in A$, entonces $f(x_n) \xrightarrow{n \rightarrow \infty} f(x_0)$
3. Si $\mathcal{U} \subset F$ es abierto, entonces $f^{-1}(\mathcal{U})$ es abierto en A .
4. Si $B \subset F$ es cerrado, entonces $f^{-1}(B)$ es cerrado en A .

Demostración:

- $(i) \Rightarrow (ii)$

Sea $\{x_n\}_{n \in \mathbb{N}} \subset A : x_n \rightarrow x_0 \in A$. Si $x_0 \notin A'$, entonces

$$\exists n_0 \in \mathbb{N} : x_n = x_0 : \forall n \geq n_0 \Rightarrow f(x_n) = f(x_0) \xrightarrow{n \rightarrow \infty} f(x_0) : \forall n \geq n_0$$

Por otro lado, si $x_0 \in A'$, entonces

$$\forall \varepsilon > 0, \exists \delta > 0 : x \in A \cap (B(x_0, \delta) \setminus \{x_0\}) \Rightarrow d'(f(x), f(x_0)) < \varepsilon$$

Luego entonces se tiene que:

$$\exists n_0 \in \mathbb{N} : \forall n \geq n_0 : d(x_n, x_0) < \delta \Rightarrow x_n \in B(x_0, \delta) \Rightarrow d'(f(x_n), f(x_0)) < \varepsilon$$

- $(ii) \Rightarrow (iv)$

Suponemos $B \subset F$ cerrado y tomamos $\{x_n\} \subset (f^{-1}(B) \cap A) : x_n \rightarrow x_0 \in A$ ¿ocurre que $x_0 \in f^{-1}(B)$, es decir, que $f(x_0) \in B$? Por (ii) tenemos que, $f(x_n) \xrightarrow{n \rightarrow \infty} f(x_0)$ y además $\{f(x_n)\} \subset B$, por tanto, por ser B cerrado, $f(x_0) \in \bar{B} = B$.

- $(iv) \Rightarrow (iii)$

Como $\mathcal{U} \subset F$ abierto, entonces $B = F \setminus \mathcal{U}$ es cerrado, luego:

$$f^{-1}(B) = (E \setminus f^{-1}(\mathcal{U})) \cap A \text{ es cerrado} \Rightarrow f^{-1}(\mathcal{U}) \text{ es abierto en } A$$

- $(iii) \Rightarrow (i)$

Sea $x_0 \in A$ y $\varepsilon > 0$ tomamos $\mathcal{U} = B(f(x_0), \varepsilon) \Rightarrow x_0 \in f^{-1}(\mathcal{U})$ abierto $\Rightarrow x_0 \in \overset{\circ}{A}$ de forma que:

$$\exists \delta > 0 : B(x_0, \delta) \subset f^{-1}(\mathcal{U}) \Rightarrow f(B(x_0, \delta)) \subset f(f^{-1}(\mathcal{U})) \subset \mathcal{U} = B(f(x_0), \varepsilon)$$

Teorema

Sean $(E, d), (F, d')$ espacios métricos, $A \subset E$ un compacto y $f : A \rightarrow F$ una función continua, entonces $f(A)$ es compacto. De hecho se cumple en general:

- Toda función constante es continua, es decir, $f(x) = k$ siempre es continua.
- Si la métrica de partida es la discreta, entonces $f(x)$ siempre es continua
- Las relaciones entre continuidad y las propiedades topológicas siguen la siguiente tabla:

$A \longrightarrow f(A)$	$f^{-1}(B) \longleftarrow B$
cerrado \nrightarrow cerrado	cerrado \longleftarrow cerrado
abierto \nrightarrow abierto	abierto \longleftarrow abierto
acotado \nrightarrow acotado	acotado \nrightarrow acotado
compacto \mapsto compacto	compacto \nrightarrow compacto

Sin embargo, tomando el espacio \mathbb{R}^n podemos considerar como cierto que los acotados van a acotados.

Demostración:

Sea $f(A) \subset \bigcup_{i \in I} \mathcal{U}_i$ unión de abiertos de F : $\forall i \in I$ sabemos que:

$$A \subset f^{-1}(f(A)) \subset f^{-1}\left(\bigcup_{i \in I} \mathcal{U}_i\right) = \bigcup_{i \in I} \underbrace{f^{-1}(\mathcal{U}_i)}_{\text{abierto}} \Rightarrow A \subset f^{-1}(\mathcal{U})_{i_1} \cup \dots \cup f^{-1}(\mathcal{U})_{i_m}$$

Entonces, de nuevo ocurre que:

$$f(A) \subset f(f^{-1}(\mathcal{U})_{i_1} \cup \dots \cup f^{-1}(\mathcal{U})_{i_m}) \subset f(f^{-1}(\mathcal{U})_{i_1}) \cup \dots \cup f(f^{-1}(\mathcal{U})_{i_m}) \subset \mathcal{U}_{i_1} \cup \dots \cup \mathcal{U}_{i_m}$$

Teorema

Sean $(E, d_1), (F, d_2), (G, d_3)$ espacios métricos y supongamos $A \subset E, B \subset F$ tal que $A \xrightarrow{f} B \xrightarrow{g} G$ ambas continuas, entonces:

$$g \circ f : A \rightarrow G \text{ es continua}$$

Demostración:

Sea $\mathcal{U} \subset G$ abierto, se tiene $(g \circ f)^{-1}(\mathcal{U}) = f^{-1}(g^{-1}(\mathcal{U}))$ y como g es continua, entonces $g^{-1}(\mathcal{U})$ es abierto de B , por extensión, como f es continua, $f^{-1}(g^{-1}(\mathcal{U}))$ es abierto de A

Proposición

Sean (E, d) espacio métrico, $(F, \|\cdot\|)$ espacio normado, $A \subset E$ y $x_0 \in A'$ un punto de acumulación, se cumple:

- Si $f, g : A \rightarrow F, \exists \lim_{x \rightarrow x_0} f(x) = L_1$ y $\exists \lim_{x \rightarrow x_0} g(x) = L_2$, entonces:

$$\exists \lim_{x \rightarrow x_0} (f(x) + g(x)) = L_1 + L_2$$

- Si $f : A \rightarrow \mathbb{R}$ y $g : A \rightarrow F, \exists \lim_{x \rightarrow x_0} f(x) = L_1 \in \mathbb{R}$ y $\exists \lim_{x \rightarrow x_0} g(x) = L_2 \in F$, entonces:

$$\exists \lim_{x \rightarrow x_0} (f(x) \cdot g(x)) = L_1 \cdot L_2$$

- Si $f : A \rightarrow \mathbb{R}$ y $g : A \rightarrow F, \exists \lim_{x \rightarrow x_0} f(x) = L_1 \neq 0$ y $\exists \lim_{x \rightarrow x_0} g(x) = L_2 \in F$, entonces:

$$\exists \lim_{x \rightarrow x_0} \frac{g(x)}{f(x)} = \frac{L_2}{L_1}$$

Demostración:

- (i), (ii) son fáciles
- (iii)

Sin pérdida de generalidad, supongamos $L_1 > 0$, entonces sea $\varepsilon = \frac{L_1}{3} > 0$ se tiene que

$$\exists \delta > 0 : f((B(x_0, \delta) \setminus \{x_0\}) \cap A) \subset \left(L_1 - \frac{L_1}{\varepsilon}, L_1 + \frac{L_1}{\varepsilon}\right)$$

Por tanto, se ocurre que $\forall x \in B(x_0, \delta) : f(x) > 0$. Veamos que $\lim_{x \rightarrow x_0} \frac{1}{f(x)} = \frac{1}{L_1}$, es decir, $\forall x \in A : x \neq x_0$ y $d(x, x_0) < \delta$ se tiene que $\left|\frac{1}{f(x)} - \frac{1}{L_1}\right| < \varepsilon$.

Por como hemos escogido el delta tenemos que $f(x) > \frac{2 \cdot L_1}{3}$, entonces $\frac{1}{|L_1 \cdot f(x)|} < \frac{3}{2 \cdot L_1^2}$ así que tenemos:

$$\left|\frac{1}{f(x)} - \frac{1}{L_1}\right| = \left|\frac{L_1 - f(x)}{L_1 \cdot f(x)}\right| \leq \frac{3\varepsilon}{2L_1^2} : \forall \varepsilon > 0$$

Corolario (Operaciones con Continuidad)

Sean (E, d) espacio métrico y $(F, \|\cdot\|)$ espacio normado, $A \subset F$ un subconjunto y $x_0 \in A \cap A'$ un punto de acumulación del conjunto:

- Si $f, g : A \rightarrow F$ son continuas en x_0 , entonces $f + g : A \rightarrow F$ es continua en x_0 .
- Si $f : A \rightarrow \mathbb{R}$ y $g : A \rightarrow F$ son continuas en x_0 , entonces $f \cdot g : A \rightarrow F$ es continua en x_0 .
- Si $f : A \rightarrow \mathbb{R}$ y $g : A \rightarrow F$ son continuas en x_0 con $\forall x \in A : f(x) \neq 0$, entonces $\frac{g}{f} : A \rightarrow F$ es continua en x_0 .

Cálculo de límites en \mathbb{R}^n

En \mathbb{R} teníamos la posibilidad de acercarnos al límite por la derecha y por la izquierda. Sin embargo, en \mathbb{R}^n podemos acercarnos desde cualquier dirección posible, luego se complica en gran medida el cálculo de límites en dicha dimensión. Por ello, es necesario definir una nueva serie de criterios que faciliten esta tarea.

Proposición (Límite en \mathbb{R}^n)

Sea (E, d) espacio métrico, $A \subset E$ un subconjunto, $f : A \rightarrow \mathbb{R}^n$ tal que $f = (f_1, \dots, f_n)$, $f_j : A \rightarrow \mathbb{R}$ y $x_0 \in A'$ un punto de acumulación, entonces:

$$\lim_{x \rightarrow x_0} f(x) = L = (L_1, \dots, L_n) \Leftrightarrow \forall j \in \{1, \dots, n\} : \lim_{x \rightarrow x_0} f_j(x) = L_j$$

Es decir, reducimos el cálculo de límites en \mathbb{R}^n al cálculo de cada componente en \mathbb{R} .

Teorema (Regla del Bocado)

Sean $f, g, h : A \subset E \rightarrow \mathbb{R}$, $x_0 \in A'$ un punto de acumulación y supongamos que:

- $\exists \delta > 0 : g(x) \leq f(x) \leq h(x) : \forall x \in (B(x_0, \delta) \setminus \{x_0\}) \cap A$
- $\exists \lim_{x \rightarrow x_0} g(x) = \lim_{x \rightarrow x_0} h(x) = L$

Entonces se tiene que:

$$\exists \lim_{x \rightarrow x_0} f(x) = L$$

Teorema (Límites reiterados)

Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y $(a, b) \in \mathbb{R}^2$ que satisfacen que:

- $\exists \lim_{(x,y) \rightarrow (a,b)} f(x, y) = L$
- $\forall y \in \mathbb{R}, \exists \lim_{x \rightarrow a} f(x, y) = L_1(y)$
- $\forall x \in \mathbb{R}, \exists \lim_{y \rightarrow b} f(x, y) = L_2(x)$

Entonces, existen los límites reiterados y son iguales:

$$\lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x, y) = \lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x, y) = L$$

Demostración:

Por hipótesis, tenemos que:

$$\forall \varepsilon > 0, \exists \delta > 0 : 0 < ||(x, y) - (a, b)|| < \delta \Rightarrow |f(x, y) - L| < \frac{\varepsilon}{2}$$

También sabemos que:

$$\forall y \in \mathbb{R}, \exists \delta_1 = \delta_1(y) > 0 : 0 < |x - a| < \delta_1(y) \Rightarrow |f(x, y) - L_1(y)| < \frac{\varepsilon}{2}$$

Por tanto, solo hay que ver que $L_1(y) \xrightarrow{y \rightarrow b} L$. Suponemos que $|y - b| < \delta$, y sea x tal que $|x - a| < \min\{\delta_1(y), \delta\}$, entonces:

$$|L_1(y) - L| \leq |L_1(y) - f(x, y)| + |f(x, y) - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

El primer sumando es menor que $\frac{\varepsilon}{2}$ por hipótesis y, para el segundo sumando, si tomamos la métrica $||\cdot||_\infty$, observamos que $||(x, y) - (a, b)||_\infty < \delta$. Como la métrica infinito es equivalente a la métrica euclídea, solo diferirá en una constante, luego ya lo tenemos.

Ejemplo:

$$\blacksquare \text{ Sea } f(x, y) = \begin{cases} \frac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

$$\text{Si } y \neq 0, \lim_{x \rightarrow 0} f(x, y) = -1$$

$$\text{Si } x \neq 0, \lim_{y \rightarrow 0} f(x, y) = 1$$

Como los límites reiterados existen pero no coinciden, $\nexists \lim_{(x, y) \rightarrow (0, 0)} f(x, y)$

$$\blacksquare \text{ Sea } f(x, y) = \begin{cases} \frac{|x| + |y|}{\sqrt{x^2 + y^2}} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

$$\text{Si } y \neq 0, \lim_{x \rightarrow 0} f(x, y) = 1$$

$$\text{Si } x \neq 0, \lim_{y \rightarrow 0} f(x, y) = 1$$

Sin embargo, observamos que para los puntos de la recta $y = x$ tenemos:

$$f(x, x) = \frac{2 \cdot |x|}{\sqrt{2} \cdot x^2} = \frac{2}{\sqrt{2}} = \sqrt{2}$$

Más generalmente, para una recta cualquiera que pase por el origen, es decir, $y = \lambda \cdot x, \lambda \in \mathbb{R}$ se tiene:

$$f(x, \lambda \cdot x) = \frac{|x|(\lambda + 1)}{\sqrt{x^2(1 + \lambda^2)}} = \frac{\lambda + 1}{\sqrt{\lambda^2 + 1}}$$

Como $\forall \{x_n\} \subset \mathcal{R} : x_n \rightarrow (0, 0)$ se tiene que $f(x_n) = \frac{\lambda + 1}{\sqrt{\lambda^2 + 1}} \nrightarrow 0$, entonces el límite no existe. Lo cual pone de manifiesto que la existencia de los límites reiterados no implica la existencia del límite.

$$\blacksquare \text{ Sea } f(x, y) = \begin{cases} y \sen \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Fijando la $x \neq 0$, calculamos el límite:

$$\lim_{y \rightarrow 0} y \sen \frac{1}{x} = 0$$

Además, observamos que si fijamos la $y \in \mathbb{R}$, entonces:

$$\nexists \lim_{x \rightarrow 0} y \sen \frac{1}{x}$$

Pero, por la regla del Bocadillo, el límite $\exists \lim_{(x, y) \rightarrow (0, 0)} f(x, y) = 0$, luego se pone de manifiesto la necesidad de que se den las 3 condiciones simultáneamente.

Teorema (Convergencia por coordenadas polares)

Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y tomando $x = r \cos \theta$ y $y = r \sin \theta$, tenemos que:

$$\lim_{(x,y) \rightarrow (0,0)} f(x,y) = L \Leftrightarrow \lim_{r \rightarrow 0^+} f(r \cos \theta, r \sin \theta) = L \text{ uniformemente en } \theta \in [0, 2\pi)$$

Es decir, que la convergencia de la función debe ser independiente de θ .

Demostración:

■ \Rightarrow : inmediato por la definición de límite

■ \Leftarrow :

Sabemos por la definición que se ha dado de convergencia uniforme que:

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall 0 < r < \delta : |f(r \cos \theta, r \sin \theta) - L| < \varepsilon : \forall \theta \in [0, 2\pi)$$

Y queremos ver que $\forall \varepsilon > 0, \exists \delta > 0 : \|(x,y)\| < \delta \Rightarrow |f(x,y) - L| < \varepsilon$. En consecuencia, basta hacer un cambio de variable [COMPLETAR].

Ejemplo:

■ Sea $f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$

$$\lim_{(x,y) \rightarrow (0,0)} f(x,y) = \lim_{r \rightarrow 0^+} \frac{r^4 \cos \theta \sin^3 \theta}{r^2} = \lim_{r \rightarrow 0^+} r^2 \cos \theta \sin^3 \theta = 0$$

Es decir, el límite es uniforme en θ , porque la convergencia a 0, en este caso, es independiente del valor que tome θ .

■ Sea $f(x,y) = \begin{cases} \frac{x}{y} \sin(x^2 + y^2) & y \neq 0 \\ 0 & y = 0 \end{cases}$

$$\lim_{(x,y) \rightarrow (0,0)} f(x,y) = \lim_{r \rightarrow 0^+} \frac{r \cos \theta}{r \sin \theta} \sin r^2 \not\rightarrow 0 \text{ uniformemente en } \theta \in [0, 2\pi)$$

Porque como $\sup_{\theta \in [0, 2\pi)} \left\{ \cot \theta = \frac{\sin \theta}{\cos \theta} \right\} = \infty$, entonces para dicho valor de θ el límite no converge a 0, es decir, no converge uniformemente en θ .

Proposición (Convergencia por curvas)

Sea $f : \mathbb{R}^2 \subset \rightarrow \mathbb{R}$ y un abierto \mathcal{U} de forma que $(a,b) \in \mathcal{U}$, entonces podemos definir²

$$\alpha : [-c, c] \rightarrow \mathcal{U} \text{ continua : } \alpha(0) = (a,b)$$

Entonces podemos decir que:

$$\lim_{(x,y) \rightarrow (a,b)} f(x,y) = L \Rightarrow \lim_{t \rightarrow 0} f(\alpha(t)) = L$$

Por tanto, al generalizar el límite para \mathbb{R}^n , la existencia del límite implica que es el mismo valor para cualquier curva o modo de aproximación al punto, no sólo para direcciones concretas.

Demostración:

Trivial a partir de la definición.

²Donde c puede ser el valor que más nos interese (se suele coger 1 o un valor cercano al punto)

Ejemplo:

$$\blacksquare \text{ Sea } f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Veamos que ocurre con las rectas de la forma $y = \lambda x, \lambda \in \mathbb{R}$:

$$f(x, \lambda x) = \frac{\lambda^2 x^3}{(\lambda^2 + 1)\lambda^2} = \frac{\lambda^2}{(\lambda^2 + 1)} x \xrightarrow{x \rightarrow 0} 0$$

Lo que nos indica que es posible que dicho límite exista, luego consideremos la curva $\alpha(t) = (t^2, t)$ y tenemos entonces que:

$$f(\alpha(t)) = \frac{t^2 \cdot t}{t^4 + t^2} = \frac{t^2}{t^2 + 1} \xrightarrow{t \rightarrow 0} 0$$

Por tanto, teniendo las suficientes garantías de que parece que converge, demostramos:

$$f(r \cos \theta, r \sin \theta) = \frac{r^3 \cos \theta \sin^2 \theta}{r^2} = r \cos \theta \sin^2 \theta \rightarrow 0 \text{ uniformemente en } \theta$$

■ 31 d) COMPLETAR

Teorema

Sea (E, d) espacio métrico, sea $f : A \rightarrow \mathbb{R}$ continua y un compacto $B \subset A$, entonces:

$$\exists x_1, x_2 \in B : f(x_1) = \max_{x \in B} f(x) \quad f(x_2) = \min_{x \in B} f(x)$$

Es decir, que se alcanza el supremo y el ínfimo en las imágenes de los compactos de las funciones continuas.

Demostración:

Sabemos que $f(B) \subset \mathbb{R}$ es compacto y, por el Teorema de Heine - Borel, $f(B)$ es cerrado y acotado. Por ser acotado, sabemos que $\exists M = \sup_{x \in B} f(x)$ y $\exists m = \inf_{x \in B} f(x)$ y por ser cerrado, $M = f(x_1), m = f(x_2) : x_1, x_2 \in B$.

Definición (Continuidad Uniforme)

Sean $(E, d), (F, d')$ espacios métricos. Se dice que $f : A \rightarrow F, A \subset E$ es **uniformemente continua** en A si:

$$\forall \varepsilon > 0, \exists \delta > 0 : d'(f(x), f(y)) < \varepsilon \text{ si } d(x, y) < \delta, \forall x, y \in A$$

Observación:

■ Si f es uniformemente continua en A , entonces f es continua en A

Teorema

Sea $f : A \rightarrow F$ y supongamos que f es continua y A es compacto, entonces f es uniformemente continua.

Demostración:

Sea $\varepsilon > 0$ y sea $x \in A$. Como f es continua en x , $\exists \delta_x > 0 : f(B(x, \delta_x)) \subset B(f(x), \frac{\varepsilon}{2})$

Como $A \subset \bigcup_{x \in A} B(x, \frac{\delta_x}{2})$, entonces $A \subset B(x_1, \frac{\delta_{x_1}}{2}) \cup \dots \cup B(x_m, \frac{\delta_{x_m}}{2})$

Sea $\delta = \min\{\frac{\delta_{x_1}}{2}, \dots, \frac{\delta_{x_m}}{2}\} > 0$, y sean $x, y \in A : d(x, y) < \delta$

Como $x \in A$, $\exists j : 1 \leq j \leq m : B(x_j, \frac{\delta_{x_j}}{2})$. Entonces, $d'(f(x), f(x_j)) < \frac{\varepsilon}{2}$. Por la desigualdad triangular,

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

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

DEFINICIÓN DE DIFERENCIABILIDAD

La diferenciabilidad supone una parte esencial del cálculo porque permite desde calcular planos tangentes en geometría hasta poder hacer un estudio exhaustivo de la monotonía y de los máximos y mínimos de una función. Por todo ello, es necesario generar una definición válida que haga extensible las propiedades de \mathbb{R} a dimensión arbitraria.

Diferencial y propiedades

Cuando trabajábamos en una dimensión, $f : I \rightarrow \mathbb{R}$ con $I \subset \mathbb{R}$ intervalo abierto y $x_0 \in I$, definíamos la derivada como:

$$f'(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + t) - f(x_0)}{t} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

DIBUJO

Donde podemos expresar la recta tangente como:

$$y = f(x_0) + f'(x_0)(x - x_0) \quad \langle (x - x_0, y - f(x_0), (f'(x_0), -1)) \rangle = 0$$

En esta sección vamos a tratar de definir quién juega el papel de $f'(x_0)$ en funciones multivariable y entender como se extrapolan los resultados de derivabilidad a más de una dimensión.

Definición (Derivada Direccional)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto³, $x_0 \in A$ y sea $v \in \mathbb{R}^n$. Se define como la **derivada** de f en la dirección v en el punto x_0 como:

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

También se conoce como **derivada direccional**.

Observación:

El módulo del vector que se emplee para calcular dicha derivada no hace más que dilatar el resultado de la misma, es decir, para calcular la derivada del vector unitario basta con dividir por el módulo del vector:

$$D_{\lambda v} f(x_0) = \lambda D_v f(x_0)$$

Asimismo, fijados el punto x_0 y la dirección v sobre la que se va a calcular la derivada, podemos definir la función $\phi(t) = f(x_0 + tv)$ de modo que la derivada direccional se reduce a una derivada usual sobre 1 variable, en este caso t:

$$D_v f(x_0) = \phi'(0)$$

Ejemplo:

- Sea $f(x, y) = |x^2 - y^2|^{\frac{1}{2}}$ podemos estudiar cuando existe $D_v f(0, 0)$. Sin pérdida de generalidad, suponemos v vector unitario de la forma $v = (\cos \theta, \sin \theta) : \theta \in [0, 2\pi)$ que es trivialmente de dirección arbitraria:

$$D_v f(0, 0) = \lim_{t \rightarrow 0} \frac{f(t \cos \theta, t \sin \theta) - 0}{t} = \frac{|t|}{t} |\cos^2 \theta - \sin^2 \theta|^{\frac{1}{2}}$$

Como el término $\frac{|t|}{t}$ no converge por ambos lados cuando $t \rightarrow 0$, tiene que ocurrir que $\cos^2 \theta = \sin^2 \theta$ para que se anule todo. Es decir, f es derivable en $(0, 0)$ en las direcciones:

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right), \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right)$$

³En general y aunque no se especifique los enunciados de esta sección siempre son para funciones definidas en abiertos

Definición (Derivada Parcial)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto y $x_0 \in A$, definimos **derivada parcial j -ésima** de f en x_0 como:

$$D_j f(x_0) = \frac{\partial f}{\partial x_j}(x_0) = D_{e_j} f(x_0)$$

Donde $e_j = (0, \dots, \underbrace{1}_j, \dots, 0)$ denota a los vectores de la base canónica.

Observación:

Recuperando la observación hecha anteriormente, si consideramos $\phi(t) = f(x_0 + e_j t)$ podemos definir la derivada parcial como:

$$\frac{\partial f}{\partial x_j}(x_0) = \phi'(0)$$

Es decir, que el cálculo de derivadas parciales se reduce a la derivación de la función $\phi(t)$ en $t = 0$ dejando fijas todas las variables que no sean x_j .

Ejemplo:

Sea $f(x, y) = y \sin(xy^2)$, entonces vemos que:

$$D_1 f = \frac{\partial f}{\partial x}(x, y) = y \cdot \cos(xy^2) \cdot y^2 = y^3 \cos(xy^2)$$

$$D_2 f = \frac{\partial f}{\partial y}(x, y) = \sin(xy^2) + 2xy^2 \cos(xy^2)$$

Definición (Gradiente)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto y $x_0 \in A$, si existen todas las derivadas parciales, se define **el gradiente de f en x_0** como:

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

Cabe mencionar que dicha definición genera un **vector** del espacio de partida.

Ejemplo:

Tomando la función de antes $f(x, y) = y \sin(xy^2)$ y habiendo calculado sus derivadas direccionales, tenemos que:

$$\nabla f(x, y) = (y^3 \cos(xy^2), \sin(xy^2) + 2xy^2 \cos(xy^2)) \Rightarrow \nabla f(0, 1) = (1, 0)$$

Observación:

En ocasiones trabajaremos con funciones del tipo $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. En este caso, basta con considerar la función como $f(z) = (f_1(z), \dots, f_m(z))$ y hacer los cálculos relativos sobre cada componente, es decir, sobre la función $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ siendo la derivada direccional resultante:

$$D_v f(x_0) = (D_v f_1(x_0), \dots, D_v f_m(x_0)) \in \mathbb{R}^m$$

En cuyo caso, el gradiente se convierte en una matriz: la fila i -ésima es el gradiente sobre la función f_i .

Definición (Forma lineal)

Sea $L : \mathbb{R}^n \rightarrow \mathbb{R}$ una aplicación lineal, por estar definida de \mathbb{R}^n a \mathbb{R} , entonces decimos que se llama **forma lineal sobre \mathbb{R}^n** de modo que:

$$L(x) = (a_1 \quad \dots \quad a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \dots + a_n x_n$$

Esto puede ser expresado como $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) = M_{1,n}$, es decir, hay una matriz que representa dicha forma lineal (mejor dicho un vector fila) cuyos coeficientes podemos calcular de la siguiente forma:

$$x = \sum_{j=1}^n x_j e_j, L(x) = \sum_{j=1}^n x_j \underbrace{L(e_j)}_{a_j}$$

Proposición

Si $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, entonces L es uniformemente continua en \mathbb{R}^n

Demostración:

Dado un $\varepsilon > 0$, si tomamos $\delta = \frac{\varepsilon}{2|||(a_1, \dots, a_n)|||}$, entonces:

$$|L(x) - L(y)| = |L(x - y)| = |\langle (a_1, \dots, a_n), (x - y) \rangle| \leq ||(a_1, \dots, a_n)||_2 \cdot ||x - y||_2 \leq \varepsilon$$

Para los vectores que verifican $||x - y||_2 < \delta$.

Definición (Hiperplano tangente)

Definimos un **hiperplano afín** como un subespacio afín $H \subset \mathbb{R}^m$ de dimensión $m - 1$, es decir:

$$H = \{x \in \mathbb{R}^m : L(x) = d\} : L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}) \text{ y } d \in \mathbb{R}$$

Luego podemos expresar dicho espacio afín como el conjunto de soluciones de la ecuación definida por la forma lineal L y d . Dicho hiperplano se denomina **tangente en un punto** cuando su vector normal es paralelo al gradiente de la gráfica en un punto y solo interseca en el punto con la misma.

Observación:

Cuando trabajábamos en \mathbb{R} con funciones de una variable del tipo $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ que eran derivables en un punto $x_0 \in I$, podíamos definir el concepto de recta tangente en el punto x_0 como la recta de pendiente $f'(x_0)$ y que pasa por el punto $(x_0, f(x_0))$. De este modo, podemos reescribir el concepto de derivada como:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{f(x_0 + t) - f(x_0)}{t} = f'(x_0) &\Leftrightarrow \lim_{t \rightarrow 0} \frac{f(x_0 + t) - \overbrace{(f(x_0) + t f'(x_0))}^{\text{recta tangente}}}{t} = 0 \Leftrightarrow \\ &\Leftrightarrow \lim_{t \rightarrow 0} \frac{f(x_0 + t) - (f(x_0) + t f'(x_0))}{|t|} = 0 \end{aligned}$$

De modo análogo, podemos definir el siguiente concepto fundamental a la hora de entender el cálculo diferencial.

Definición (Función Diferenciable)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto y $x_0 \in A$, se dice que f es **diferenciable en x_0** si:

$$\exists L_{x_0} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) : \lim_{h \rightarrow 0} \frac{f(x_0 + h) - \overbrace{(f(x_0) + L_{x_0}(h))}^{\text{hiperplano tangente}}}{||h||} = 0$$

Es decir, las funciones diferenciables introducen el concepto de **hiperplano tangente**, que sería el homólogo a las rectas tangentes definidas en \mathbb{R} .

Cabe destacar que esta definición y la definición de derivada coincide para las funciones de una variable, es decir, que dada $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, entonces f es diferenciable en x_0 si y sólo si f es derivable en x_0 y, además, $L_{x_0}(t) = t \cdot f'(x_0)$.

Definición (Diferencial)

Si f es diferenciable en $x_0 \in A$, entonces llamamos **diferencial**⁴ de f en x_0 :

$$df(x_0) = L_{x_0} \in L(\mathbb{R}^n, \mathbb{R})$$

De esta forma, tenemos que:

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

Teorema

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto, $x_0 \in A$ y $v \in \mathbb{R}^n$, si f es diferenciable en x_0 , entonces existe la derivada direccional en cualquier dirección.

$$f \text{ diferenciable} \Rightarrow \forall v \in \mathbb{R}^n : \exists D_v f(x_0)$$

Además, se tiene que:

$$L_{x_0}(v) = D_v f(x_0) \text{ para } \|v\| = 1$$

Demostración:

Como conocemos la certeza de la existencia del límite, podemos aproximarnos desde la dirección o curva que deseemos, luego si consideramos $h = t \cdot v$ con $t \in \mathbb{R}$, tenemos que:

$$\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0) - tL_{x_0}(v)}{|t|} = 0 \Leftrightarrow \underbrace{\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}}_{D_v f(x_0)} = L_{x_0}(v)$$

Corolario

Si f es diferenciable en $x_0 \in A$, entonces:

- La diferencial en x_0 es el gradiente en x_0 :

$$df(x_0) = \nabla f(x_0) \quad df(x_0)(v) = \langle \nabla f(x_0), v \rangle = \nabla f(x_0) \cdot v = \sum_{j=1}^n v_j \frac{\partial f}{\partial x_j}(x_0)$$

Porque, si consideramos $L_{x_0} = (a_1, \dots, a_n)$, por el teorema que hemos demostrado antes, tenemos que:

$$\begin{cases} a_1 \cdot 1 + 0 + \dots + 0 = L_{x_0}(e_1) = D_{e_1} f(x_0) = \frac{\partial f}{\partial x_1} \\ \vdots \\ a_n \cdot 1 + 0 + \dots + 0 = L_{x_0}(e_n) = D_{e_n} f(x_0) = \frac{\partial f}{\partial x_n} \end{cases} \Rightarrow df(x_0) = \nabla f(x_0)$$

- La demostración anterior no depende de que el vector sea unitario:

$$df(x_0)(v) = df(x_0) \left(\|v\| \cdot \frac{v}{\|v\|} \right) = \|v\| \cdot df(x_0) \left(\frac{v}{\|v\|} \right) = \|v\| \cdot D_{\frac{v}{\|v\|}} f(x_0) = \frac{\|v\|}{\|v\|} \cdot D_v f(x_0) = D_v f(x_0)$$

Es decir, cuando trabajábamos en \mathbb{R} el papel de la diferencial (o como hemos visto el gradiente porque son lo mismo) lo hacía $f'(x_0)$ que en el fondo es un vector en \mathbb{R} . Consecuentemente, tiene sentido definir $df(x_0)$ como un vector en \mathbb{R}^n porque al restringirse a \mathbb{R} sigue siendo aplicable dicha definición.

⁴De hecho, de la unicidad del límite se desprende la unidad de la diferencial y, en consecuencia, del gradiente.

Interpretación geométrica de la diferencial

Cuando trabajamos en funciones de una variable, para hallar la ecuación de la recta tangente (que hace las veces de hiperplano tangente en 2 dimensiones) en un punto a la función teníamos en cuenta la siguiente ecuación:

$$\langle (x - x_0, y - f(x_0)), (f'(x_0), -1) \rangle = 0 \Leftrightarrow y = f(x_0) + f'(x_0)(x - x_0)$$

Si tratamos de expresar dicha recta como una forma lineal en \mathbb{R}^2 , entonces tenemos que:

$$\langle (x, y), (f'(x_0), -1) \rangle - \langle (x_0, f(x_0)), (f'(x_0), -1) \rangle = 0 \Rightarrow$$

$$\langle (x, y), (f'(x_0), -1) \rangle = \langle (x_0, f(x_0)), (f'(x_0), -1) \rangle$$

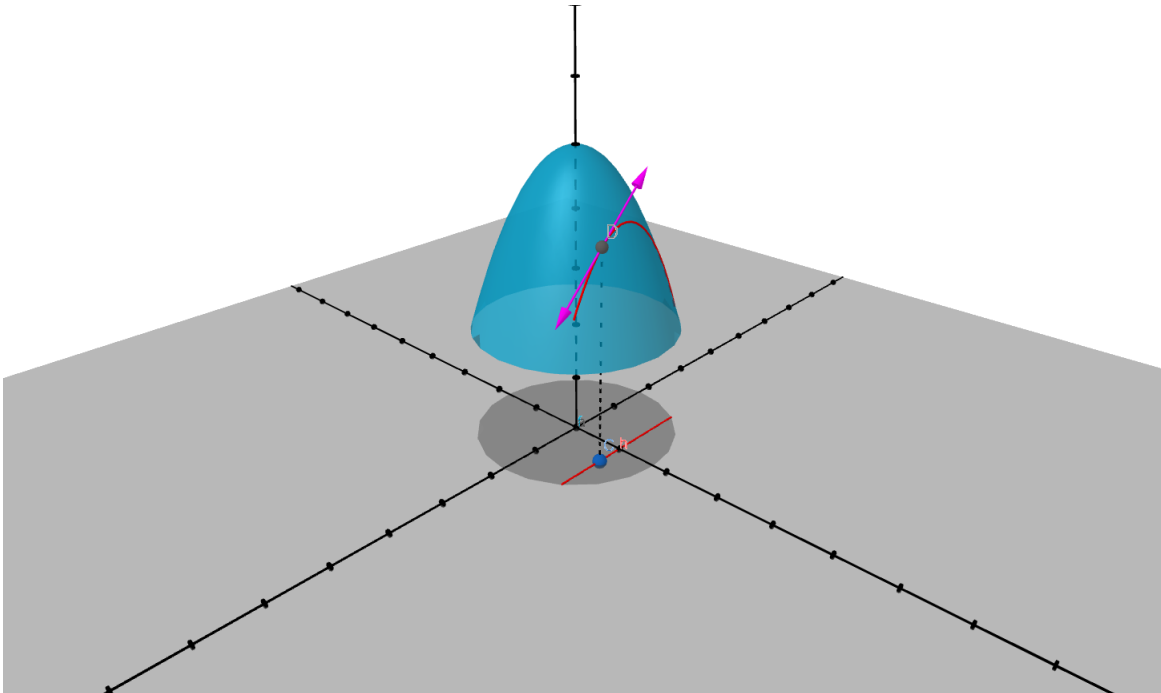
Es decir, que si llamamos $L_{x_0} = (f'(x_0), -1)$, entonces tenemos que la recta puede ser expresada como:

$$\underbrace{L_{x_0}(x, y)}_{\in L(\mathbb{R}^2, \mathbb{R})} = \underbrace{L_{x_0}(x_0, f(x_0))}_{\in \mathbb{R}}$$

De modo análogo, vamos a tratar de determinar dichas ecuaciones y esa forma lineal en el caso de dimensión arbitraria.

Tomemos $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en $x_0 \in A$.

- Si fijamos $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n$, es decir, todas las variables excepto una, estamos reduciendo la función inicial a una función sobre una única variable (la x_j) y los puntos de dicha gráfica son de la forma $(x_1, \dots, x_n, f(x))$.
- Para calcular la derivada de dicha función tenemos que calcular $f'(x_0)$, es decir, $\frac{\partial f}{\partial x_j}(x_0)$. Trivialmente, como el vector tangente en una variable es el vector $(1, f'(x_0))$ extrapolando a más dimensiones tenemos que el vector $(0, \dots, 1, \dots, \frac{\partial f}{\partial x_j}(x_0))$ es el vector tangente a dicha gráfica.
- Este último paso es así porque, en dimensión 2, vale 1 en la variable que se deja libre, las demás tienen que valer 0 porque nos permite restringir el vector al plano que surge de fijar todas las demás variables y con el que cortamos a la gráfica y la derivada es la última coordenada, pero en el caso multivariable la derivada en x_0 es la parcial en x_0 .



Para poner un ejemplo y en referencia al dibujo, consideramos $f(x, y) = -x^2 - y^2 + 5$ restringida al conjunto de puntos $\{x^2 + y^2 \leq 3\}$ que son la cónica azul y la sombra gris que proyecta en el suelo y el punto $x_0 = (1, \frac{1}{2})$ el punto azul de la sombra. Cuando fijamos la y y tomamos la x como variable libre nos estamos restringiendo a los puntos de la recta roja sobre el suelo, que son los que tienen como $y = \frac{1}{2}$ (la de x_0) y recorren x , cuyas imágenes forman la parábola roja en la cónica, que son los puntos de la forma $(x, \frac{1}{2}, f(x, \frac{1}{2}))$. Si calculamos $\frac{\partial f}{\partial x} = -2x$, entonces vemos que el vector tangente a la gráfica en el punto $(1, \frac{1}{2}, \frac{15}{4})$, que es la imagen de x_0 , es el vector formado por $(1, 0, \frac{\partial f}{\partial x}(x_0)) = (1, 0, -2)$.

- Para seguir y poder calcular el hiperplano tangente a la gráfica en dicho punto hay que recordar una noción básica de Álgebra: un hiperplano viene dado por una forma lineal cuyos coeficientes pueden ser considerados como un vector, que será el vector ortogonal al hiperplano en cuestión llamado vector normal.
- Por tanto, tenemos que calcular el vector normal a todos los vectores tangentes en cualquiera de las direcciones posibles. Como habíamos definido dichos vectores tangentes como:

$$v_j = \left(0, \dots, \underbrace{1}_j, \dots, 0, \frac{\partial f}{\partial x_j}(x_0) \right)$$

Si tomamos el vector

$$u = \left(\underbrace{\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0)}_{\nabla f(x_0)}, -1 \right)$$

Entonces el producto escalar resulta como

$$v_j \cdot u = 0 + \dots + \frac{\partial f}{\partial x_j}(x_0) + \dots + 0 - \frac{\partial f}{\partial x_j}(x_0)$$

Luego es ortogonal a todos los demás vectores

A partir de aquí, $x \in \mathbb{R}^n$ denota un vector de dimensión n , no la coordenada x del vector. De modo análogo, $z, f(x_0) \in \mathbb{R}$ se refieren a la coordenada x_{n+1} del vector y del punto x_0 respectivamente.

- Los puntos que definen la variedad afín (el hiperplano tangente) que estamos buscando están definidos por los vectores $(x - x_0, z - f(x_0)) \in \mathbb{R}^n$ donde además $(x - x_0, z - f(x_0)) \cdot (\nabla f(x_0), -1) = 0$ porque tienen que ser perpendiculares al vector normal, por tanto:

$$z = \nabla f(x_0)(x - x_0) + f(x_0)$$

Ecuación que nos resulta muy similar a la de la recta tangente en \mathbb{R}^2 .

- Para expresarlo en forma de forma lineal, basta con despejar $(\nabla f(x_0), -1)$ del producto escalar inicial, es decir:

$$(x, z) \cdot \underbrace{(\nabla f(x_0), -1)}_{L_{x_0}} = (x_0, f(x_0)) \cdot (\nabla f(x_0), -1)$$

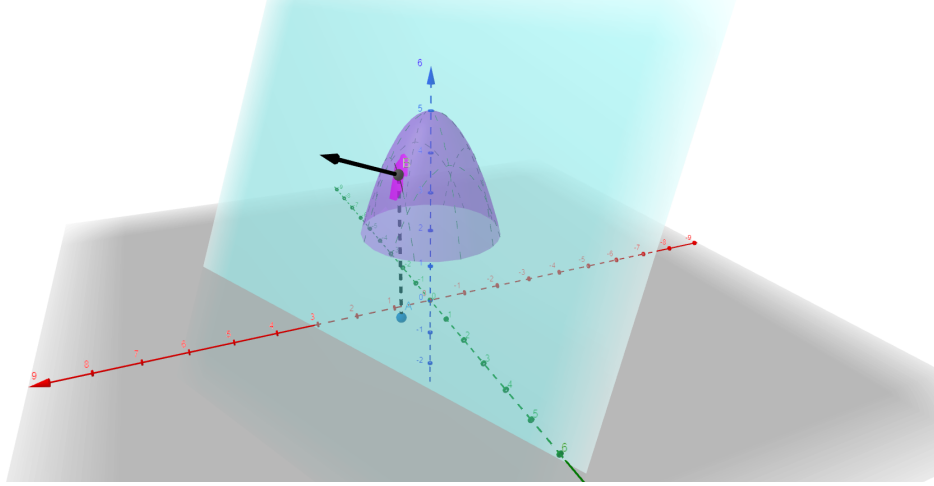
Luego los coeficientes de la forma lineal, son justamente los del vector normal. Para visualizarlo mejor, en caso de encontrarse en dimensión 3, tendríamos que:

$$(x, y, z) \cdot \left(\frac{\partial f}{\partial x}(x_1, x_2), \frac{\partial f}{\partial y}(x_1, x_2), -1 \right) = (x_1, x_2, f(x_1, x_2)) \cdot \left(\frac{\partial f}{\partial x}(x_1, x_2), \frac{\partial f}{\partial y}(x_1, x_2), -1 \right)$$

Donde $x, y, z \in \mathbb{R}$ denotan coordenadas y $x_1, x_2, x_3 \in \mathbb{R}$ denotan las coordenadas del punto x_0 en cuestión.

Para completar el ejemplo anterior, el vector normal $(\nabla f(x_0), -1) = (-2, -1, -1)$ que hemos tomado en positivo y el hiperplano tangente vendría dado por:

$$(x, y, z) \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \left(1, \frac{1}{2}, \frac{15}{4}\right) \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$



Teorema

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en $x_0 \in A$, entonces f es continua en x_0 .

Demostración:

Sea $\varepsilon = 1$ existe $\delta' > 0$ de modo que $\|x - x_0\| < \delta' \Rightarrow |f(x) - f(x_0) - df(x_0)(x - x_0)| < \|x - x_0\|$ por ser diferenciable. De este modo:

$$|f(x) - f(x_0)| \leq |f(x) - f(x_0) - df(x_0) \cdot (x - x_0)| + |df(x_0) \cdot (x - x_0)| \leq \|x - x_0\| \cdot (1 + \|\nabla f(x_0)\|)$$

Luego, para cualquier $\varepsilon > 0$ arbitrario, tomando como $\delta = \min \left\{ \frac{\delta'}{\|\nabla f\| + 1}, \frac{\varepsilon}{1 + \|\nabla f\|} \right\}$, obtenemos que $|f(x) - f(x_0)| < \varepsilon$ para $\|x - x_0\| < \delta$.

Ejemplo:

Sea $f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$ para ver que no es continua veamos que no es diferenciable en el $(0, 0)$.

Tomamos un vector unitario de dirección arbitraria $v = (\cos \theta, \sin \theta)$, entonces resulta :

$$\lim_{t \rightarrow 0} \frac{f((0, 0) + tv) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{f(t \cos \theta, t \sin \theta)}{t} = \lim_{t \rightarrow 0} \frac{1}{t} \cdot \frac{t^3 \cos \theta \sin^2 \theta}{t^2 \cos^2 \theta + t^4 \sin^4 \theta}$$

Basta ahora con distinguir casos para evaluar la continuidad:

■ $\cos \theta = 0$:

$$\frac{1}{t} \cdot \frac{t^3 \cos \theta \sin^2 \theta}{t^2 \cos^2 \theta + t^4 \sin^4 \theta} \xrightarrow{t \rightarrow 0} 0 = D_v f(0, 0)$$

■ $\cos \theta \neq 0$:

$$\frac{\cos \theta \sin^2 \theta}{\cos^2 \theta + t^2 \sin^4 \theta} \xrightarrow{t \rightarrow 0} \frac{\sin^2 \theta}{\cos \theta} = D_v f(0, 0)$$

Entonces, f no es continua en $(0, 0)$. También podíamos verlo por la aproximación por la curva $(y^2, y) \Rightarrow f(y^2, y) = \frac{y^4}{y^4+y^4} = \frac{1}{2} \Rightarrow \nexists \lim_{(x,y) \rightarrow (0,0)} f(x, y)$.

Proposición (Operaciones con Diferenciabilidad)

Sean $f, g : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $x_0 \in A$. Si f, g son diferenciables en x_0 , entonces se cumple:

- $f + g$ es diferenciable y $d(f + g)(x_0) = df(x_0) + dg(x_0)$
- $f \cdot g$ es diferenciable en x_0 y $d(f \cdot g)(x_0) = df(x_0) \cdot g(x_0) + f(x_0) \cdot dg(x_0)$
- Si $f(x_0) \neq 0$, entonces $\frac{1}{f}$ es diferenciable en x_0 y $d\left(\frac{f}{g}\right)(x_0) = \frac{f(x_0) \cdot dg(x_0) - g(x_0) \cdot df(x_0)}{f^2(x_0)}$

Teorema

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $x_0 \in A$. Si $\forall j \in \{1, \dots, n\} : \exists \frac{\partial f}{\partial x_j}(x)$ y son funciones continuas en x_0 , entonces f es diferenciable.

Demostración:

Para la demostración vamos a suponer que $n = 2$, puesto que para más dimensión basta con extrapolar esta demostración a dicha dimensión.

Sean $(x, y) \in \mathbb{R}^2$ y $x_0 = (a, b) \in A$, entonces:

$$f(x, y) - f(a, b) = f(x, y) - f(a, y) + f(a, y) - f(a, b) = (f(x, y) - f(a, y)) + (f(a, y) - f(a, b)) =$$

Considerando en el primer paréntesis la y fija, entonces por el T^a. V. M. se tiene que $\exists \mu \in (x, a) \cup (a, x)$ que verifica las condiciones de abajo. Además, si consideramos en el segundo paréntesis la función como una función sobre una variable; la y , entonces por el mismo teorema $\exists \nu \in (y, b) \cup (b, y)$ que verifica la igualdad de abajo:

$$= \left(\frac{\partial f}{\partial x}(\mu, y) \right) \cdot (x - a) + \left(\frac{\partial f}{\partial y}(a, \nu) \right) \cdot (y - b)$$

Veamos ahora la siguiente expresión:

$$|f(x, y) - f(a, b) - \nabla f(a, b) \cdot (x - a, y - b)| \leq \left| \frac{\partial f}{\partial x}(\mu, y) - \frac{\partial f}{\partial x}(a, b) \right| \cdot |x - a| + \left| \frac{\partial f}{\partial y}(a, \nu) - \frac{\partial f}{\partial y}(a, b) \right| \cdot |y - b|$$

Como las derivadas parciales son funciones continuas sobre \mathbb{R} , entonces:

$$\forall \varepsilon > 0, \exists \delta > 0 : \|(x - a, y - b)\| < \delta \Rightarrow \left| \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(a, b) \right| < \frac{\varepsilon}{2} \text{ y } \left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(a, b) \right| < \frac{\varepsilon}{2}$$

Luego los valores absolutos de las parciales los tenemos acotados y como los términos $|x - a|$ e $|y - a|$ al dividirse por la norma $\|(x - a, y - b)\|$ son menores o iguales que 1, tenemos que:

$$\frac{|f(x, y) - f(a, b) - \nabla f(a, b) \cdot (x - a, y - b)|}{\|(x - a, y - b)\|} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Observación:

Del ejemplo anterior, deducimos que

$$C^1) \Rightarrow \text{Diferenciable} \Rightarrow \forall x_0 \in A : \exists \nabla f(x_0)$$

Pero de ahí no podemos deducir que la función diferencial (que es la función gradiente) sea una función continua:

$$f(x, y) = \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \Rightarrow \forall x_0 \in \mathbb{R}^2 : \exists \nabla f(x_0) \text{ y diferenciable en } (0, 0)$$

Sin embargo, la función $\nabla f(x, y) = \left(\frac{\partial f}{\partial x} f(x, y), \frac{\partial f}{\partial y} f(x, y) \right)$ no es una función continua en $\mathbb{R}^2 \rightarrow \mathbb{R}^2$.

Definición (Matriz Jacobiana)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, definimos el **Jacobiano de F en $x_0 \in A$** a:

$$J_f(x_0) = df(x_0) = \begin{pmatrix} \nabla f_1(x_0) \\ \vdots \\ \nabla f_m(x_0) \end{pmatrix} = \left(\frac{\partial f_j}{\partial x_k}(x_1, \dots, x_n) \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} \in M_{m \times n}$$

De modo análogo, cuando el espacio de llegada es \mathbb{R} , llamamos **Hessiano de f en x_0** a:

$$H_f(x_0) = \begin{pmatrix} D_{11}f(x_0) & \dots & D_{1n}f(x_0) \\ \vdots & & \vdots \\ D_{n1}f(x_0) & \dots & D_{nn}f(x_0) \end{pmatrix} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \right)_{1 \leq i, j \leq n}$$

Ejemplo:

1. Sea $f(x, y) = (x^2y, y^3x, x + y)$, entonces el jacobiano queda como:

$$J_f(x, y) = \begin{pmatrix} 2xy & x^2 \\ y^3 & 3y^2x \\ 1 & 1 \end{pmatrix} \in M_{3 \times 2}$$

2. Sea $g(x, y) = x^2y$, entonces el hessiano queda como:

$$H_g(x, y) = \begin{pmatrix} 2y & 2x \\ 2x & 0 \end{pmatrix}$$

Teorema (Regla de la Cadena)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función diferenciable en $x_0 \in A$ y $g : B \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ de modo que $f(A) \subset B$ y diferenciable en $f(x_0)$, entonces $(g \circ f) : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ es diferenciable en $x_0 \in A$ y su diferencial se calcula como:

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

Tal y como hemos demostrado antes, la diferencial en el punto x_0 se corresponde con el valor del gradiente en dicho punto, luego tendremos que:

$$d(g \circ f)(x_0) = \left(\frac{\partial (g \circ f)_j}{\partial x_k}(x_0) \right)_{\substack{1 \leq j \leq p \\ 1 \leq k \leq n}} \in M_{p \times n}$$

De hecho, como hemos definido la relación entre la diferencial de la composición y las diferenciales de las funciones que se componen, tenemos que cada coordenada de dicha matriz se calcula como:

$$\frac{\partial (g \circ f)_j}{\partial x_k}(x_0) = \sum_{l=1}^m \frac{\partial g_j}{\partial x_l}(f(x_0)) \cdot \frac{\partial f_l}{\partial x_k}(x_0)$$

Por ser la diferencial de la composición el producto de las matrices $dg(f(x_0))$ y $df(x_0)$.

Ejemplo:

Sean $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ y $g : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definidas como:

$$\begin{cases} f(x, y) = (x^2y^3, xy) \\ g(a, b) = (ab, a^2, b^2) \end{cases}$$

Por tanto, la composición $g \circ f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ se definirá como:

$$(g \circ f)(x, y) = (x^3y^4, x^4y^6, x^2y^2)$$

La matriz del gradiente de la composición puede ser calculada fácilmente por el método explicado en secciones anteriores:

$$d(g \circ f)(x, y) = \begin{pmatrix} 3x^2y^4 & 4x^3y^3 \\ 4x^3y^6 & 6x^4y^5 \\ 2xy^2 & 2x^2y \end{pmatrix}$$

Sin embargo, podemos comprobar que el producto de matrices de las diferenciales correspondiente tiene como resultado dicha matriz:

$$dg(f(x, y)) = \begin{pmatrix} xy & x^2y^3 \\ 2x^2y^3 & 0 \\ 0 & 2xy \end{pmatrix} \quad df(x, y) = \begin{pmatrix} 2x^3 & 3x^2y^2 \\ y & x \end{pmatrix}$$

$$dg(f(x, y)) \cdot df(x, y) = \begin{pmatrix} xy & x^2y^3 \\ 2x^2y^3 & 0 \\ 0 & 2xy \end{pmatrix} \cdot \begin{pmatrix} 2x^3 & 3x^2y^2 \\ y & x \end{pmatrix} = \begin{pmatrix} 3x^2y^4 & 4x^3y^3 \\ 4x^3y^6 & 6x^4y^5 \\ 2xy^2 & 2x^2y \end{pmatrix}$$

Ejemplo:

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable podemos demostrar el cálculo de $d(f^2)(x)$ gracias a la regla de la Cadena. Si consideramos $h : \mathbb{R} \rightarrow \mathbb{R}$ definida como $h(x) = x^2$, sabemos que $dh(x) = 2x$, luego:

$$d(f^2)(x) = d(h \circ f)(x) = dh(f(x)) \cdot df(x) = 2f(x) \cdot df(x)$$

Gracias a un resultado como este podríamos demostrar la fórmula de la diferencial de un producto porque $f(x) \cdot g(x) = \frac{1}{4}[(f(x) + g(x))^2 - (f(x) - g(x))^2]$, luego basta con calcular:

$$d(f \cdot g)(x) = \frac{1}{4} [2(f(x) + g(x)) \cdot (df(x) + dg(x)) - 2(f(x) - g(x)) \cdot (df(x) - dg(x))]$$

Que termina siendo la fórmula mencionada en secciones anteriores.

Taylor, extremos y puntos críticos

En este apartado, definimos los conceptos de extremos relativos, el nuevo punto silla y los puntos críticos. Asimismo, dejamos el terreno preparado para formular un Teorema Del Valor Medio en dimensión arbitraria y el esencial Teorema de Taylor. Por tanto, esta sección sienta las bases de la optimización y el cálculo analítico de la monotonía de las funciones estudiadas.

Definición (Conjunto convexo)

Sea $A \subset \mathbb{R}^n$, decimos que A es **convexo** si

$$\forall x, y \in A : L[x, y] = \{tx + (1-t)y : 0 \leq t \leq 1\} \subset A$$

Donde la definición que se ha dado de $L[x, y]$ representa geométricamente el segmento que une ambos puntos.

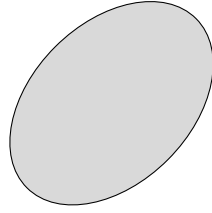
Observación:

Cuando manejábamos funciones de una variable tales como $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ derivable, podíamos afirmar que $\forall x, y \in I, \exists c \in L[x, y] : f(y) - f(x) = f'(c)(y - x)$ gracias al Teorema del Valor Medio. Sin embargo, en general es falso en el caso vectorial, por ejemplo sea $f : \mathbb{R} \rightarrow \mathbb{R}^2$ definida como $f(t) = (t^2, t^3)$ si tomamos $x = 0$ y $y = 1$ podríamos suponer que:

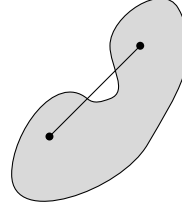
$$(1, 1) = f(1) - f(0) = df(c)(1 - 0) = df(c) : c \in [0, 1]$$

Y, en realidad, vemos que el valor de la diferencial es:

$$df(c) = (2c, 3c^2) = (1, 1) \Leftrightarrow \begin{cases} c = \frac{1}{2} \\ c = \frac{1}{\sqrt{3}} \end{cases} \Rightarrow \#$$



(a) Conjunto Convexo



(b) Conjunto no Convexo

Ejemplo:

Lema

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en el abierto A , entonces $\forall x, y \in A : L[x, y] \subset A$ y $\forall z \in \mathbb{R}^m$ se tiene que:

$$\exists c_z \in L[x, y] : \langle z, f(y) - f(x) \rangle = \langle z, df(c_z) \cdot (y - x) \rangle$$

Demostración:

Al igual que hicimos con el T^a. V. M. en Análisis, vamos a definir una función auxiliar $\psi : [0, 1] \rightarrow \mathbb{R}$ definida como $\psi(t) = \langle z, f(x + t(y - x)) \rangle$. Veamos ahora que:

$$\begin{aligned} \psi'(t) &= \frac{d}{dt} \sum_{j=1}^m z_j \cdot f_j(x + t(y - x)) = \sum_{j=1}^m z_j \cdot \frac{d}{dt} f_j(x + t(y - x)) = \\ &= \sum_{j=1}^m z_j \cdot df_j(x + t(y - x)) \cdot (y - x) = \langle z, df(x + t(y - x)) \cdot (y - x) \rangle \end{aligned}$$

Luego, como ψ es derivable, entonces por el Teorema del Valor Medio se tiene que

$$\psi(1) - \psi(0) = \psi'(\alpha)(1 - 0) : \alpha \in (0, 1) \Rightarrow \langle z, f(y) - f(x) \rangle = \langle z, df(x + \alpha(y - x)) \cdot (y - x) \rangle$$

Y vemos que basta llamar $x + \alpha(y - x) = c_z = \alpha y + (1 - \alpha)x \in L[x, y]$.

Corolario

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en el abierto A , si $L[x, y] \subset A$, entonces se cumple la igualdad:

$$\exists c \in L[x, y] : f(y) - f(x) = df(c) \cdot (y - x)$$

Demostración:

En el lema anterior, basta coger $z = 1$

Definición (Norma de una aplicación lineal)

Sea $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineal definimos **la norma de h** como:

$$\|h\| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|h(x)\|}{\|x\|} = \sup_{\|y\|=1} \|h(y)\|$$

La última igualdad se da porque como el denominador es un escalar, puede entrar dentro de la norma del numerador y por ser $h(x)$ lineal, finalmente se tiene que $\left\| h\left(\frac{x}{\|x\|}\right) \right\|$, es decir, que se reduce al supremo de las imágenes de vectores unitarios.

Observación:

Es un buen ejercicio probar que el espacio de matrices $(M_{m \times n}, \|\cdot\|)$ con la norma definida anteriormente es un espacio normado.

Teorema (Teorema de los incrementos finitos)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en el abierto A , si $L[x, y] \subset A$, se satisface que:

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

Demostración:

Si $f(x) = f(y)$ se cumple trivialmente, así que vamos a suponer que $f(x) \neq f(y)$. Sea $z = \frac{f(y) - f(x)}{\|f(y) - f(x)\|} \in \mathbb{R}^m$, por el lema anterior:

$$\exists c_z \in L[x, y] : \langle z, f(y) - f(x) \rangle = \langle z, df(c_z) \cdot (y - x) \rangle$$

Sustituyendo z en la expresión anterior:

$$\begin{aligned} \langle z, f(y) - f(x) \rangle &= \left\langle \frac{f(y) - f(x)}{\|f(y) - f(x)\|}, f(y) - f(x) \right\rangle = \frac{1}{\|f(y) - f(x)\|} \langle f(y) - f(x), f(y) - f(x) \rangle = \\ &= \frac{1}{\|f(y) - f(x)\|} \cdot \|f(y) - f(x)\|^2 = \|f(y) - f(x)\| \end{aligned}$$

Por la desigualdad de Cauchy - Schwarz

$$\|f(y) - f(x)\| \leq 1 \cdot \|df(c_z) \cdot (y - x)\| \leq \|df(c_z)\| \cdot \|y - x\| \leq \sup_{c \in L[x, y]} \|df(c)\| \cdot \|y - x\|$$

Corolario

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, donde A es convexo y abierto, es diferenciable y $df = 0$ en A , entonces f es constante.

Demostración:

Sean $x, x_0 \in A$, por la convexidad de A , $L[x, x_0] \subset A$ y, por el Teorema de los Incrementos Finitos, se tiene que:

$$\|f(x) - f(x_0)\| \leq \sup_{c \in L[x, x_0]} \|df(c)\| \cdot \|x - x_0\| = 0 \Rightarrow f(x) = f(x_0)$$

Definición (Derivada Segunda)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ donde A es abierto y $u, v \in \mathbb{R}^n$. La **derivada segunda** de F se define⁵ (siempre que existan) como:

$$D_{u,v} = D_u(D_v f)$$

Si tomamos como vectores los canónicos $u = e_i, v = e_j$, entonces:

$$D_{e_i, e_j} f = D_{i,j} f = \frac{\partial^2 f}{\partial x_i \partial x_j} = f_{ij} = f_{x_i x_j}$$

Y si además lo hacemos sobre el mismo vector en cada iteración:

$$\frac{\partial^2 f}{\partial x_i^2} = f_{ii} = f_{x_i x_i}$$

De manera análoga, se definen las derivadas de cualquier orden.

$$D_{i_k, \dots, i_1} f = \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} : i_1, \dots, i_k \in \{1, \dots, n\}$$

⁵Nótese que el orden de la derivación parcial es de dentro hacia fuera, comenzando por la variable más a la derecha

Ejemplo:

Sea $f(x, y) = yx^2 + xy^3$

$$\begin{array}{lll} \frac{\partial f}{\partial x} = 2xy + y^3 & \frac{\partial f}{\partial y} = x^2 + 3xy^2 & \frac{\partial^2 f}{\partial x^2} = 2y \\ \frac{\partial^2 f}{\partial y^2} = 6xy & \frac{\partial^2 f}{\partial y \partial x} = 2x + 3y^2 & \frac{\partial^2 f}{\partial x \partial y} = 2x + 3y^2 \end{array}$$

Definición (Función de Clase k)

Se dice que f es **función de clase k** y se denota por $f \in C^k(A)$ con $A \subset \mathbb{R}^n$ abierto, si $\forall \{i_1, \dots, i_k\} \subset \{1, \dots, n\} : \exists D_{i_k, \dots, i_1} f$ y todas son funciones continuas.

Teorema (de Schwarz - Clairaut)

Sea $f \in C^2(A)$ donde $A \subset \mathbb{R}^n$ es abierto, entonces $\forall i, j \in \{1, \dots, n\}$ se cumple que:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Es decir, las derivadas “cruzadas” coinciden (la diagonal de derecha a izquierda en la matriz de la diferencial).

Demostración:

Podemos suponer que $n = 2$ y $m = 1$ porque por inducción se puede hacer extensible para n . Además si $m > 1$, entonces al tener que hacerlo sobre cada componente y unir en un vector los resultados se tiene trivialmente por haberlo demostrado para $m = 1$. Considerando $(x, y) \in A$ y $h, k \in \mathbb{R}$ definimos las siguientes funciones:

$$g_k(u) = f(u, y+k) - f(u, y) \quad j_h(u) = f(x+h, u) - f(x, u)$$

Y determinamos el siguiente conjunto:

$$A_{h,k} = f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x, y) \Rightarrow \begin{cases} A_{h,k} = g_k(x+h) - g_k(x) \\ A_{h,k} = j_h(y+k) - j_h(y) \end{cases}$$

Por hipótesis, la función g_k es continua y derivable por serlo f y ser suma y composición de f y funciones continuas, luego por el Teorema del Valor Medio, tenemos $c_{h,k} \in (x, x+h)$ y $d_{h,k} \in (y, y+k)$ (en caso de ser negativo se da la vuelta a los intervalos) tal que:

$$g_k(x+h) - g_k(x) = g'_k(c_{h,k}) \cdot (h-0) = \left(\frac{\partial f}{\partial x}(c_{h,k}, y+k) - \frac{\partial f}{\partial x}(c_{h,k}, y) \right) \cdot h = \frac{\partial^2 f}{\partial y \partial x}(c_{h,k}, d_{h,k}) \cdot h \cdot k$$

En concreto, la última igualdad viene de considerar la función $\frac{\partial f}{\partial x}(c_{h,k}, y)$ como una función (continua y derivable por ser $f \in C^2(A)$) sobre la variable y y aplicar también el teorema del valor medio para la misma.

De este modo, si consideramos el mismo razonamiento sobre la función $j_h(u)$, tenemos $c'_{h,k} \in (x, x+h)$ y $d'_{h,k} \in (y, y+k)$ tal que:

$$j_h(y+k) - j_h(y) = j'_h(d'_{h,k}) \cdot (k-0) = \left(\frac{\partial f}{\partial y}(x+h, d'_{h,k}) - \frac{\partial f}{\partial y}(x+h, d'_{h,k}) \right) \cdot k = \frac{\partial^2 f}{\partial x \partial y}(c'_{h,k}, d'_{h,k}) \cdot h \cdot k$$

Luego, como ambas cosas son iguales a $A_{h,k}$, se tiene que:

$$\frac{\partial^2 f}{\partial x \partial y}(c'_{h,k}, d'_{h,k}) \cdot h \cdot k = \frac{\partial^2 f}{\partial y \partial x}(c_{h,k}, d_{h,k}) \cdot h \cdot k \xrightarrow[k \rightarrow 0]{h \rightarrow 0} \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y)$$

Y, en particular, cuando pasamos al límite en el último paso, podemos afirmar la igualdad porque $f \in C^2(A)$ ya que como $c_{h,k}$ y $d_{h,k}$ se aproximan a los valores de x e y , la función toma el mismo valor en el límite que en la imagen a consecuencia de la continuidad que le confiere $f \in C^2(A)$.

Observación:

En general, la igualdad no es cierta si $f \notin \mathcal{C}^2(A)$. Por ejemplo, $f(x, y) = \begin{cases} \frac{xy^3}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \cdot f$ tiene derivadas de orden 2 en \mathbb{R}^n y $\frac{\partial^2 f}{\partial x \partial y}(0, 0) \neq \frac{\partial^2 f}{\partial y \partial x}(0, 0)$.

Corolario

Si $f \in \mathcal{C}^k(A)$ es de clase k y $\sigma \in \mathcal{S}_k$ es el conjunto de permutaciones de $\{1, \dots, k\}$, entonces

$$D_{i_1, \dots, i_k} f = D_{\sigma(\{1, \dots, k\})} f$$

Es decir, el orden de derivación sobre cada dirección no influye en el resultado final.

Teorema (Teorema de Taylor)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto y $f \in \mathcal{C}^{k+1}(A)$. Si tomamos dos puntos $x_0 \in A$ y $h \in \mathbb{R}^n \setminus \{0\}$ de modo que $L[x_0, x_0 + h] \subset A$, entonces $\exists c \in L[x_0, x_0 + h]$ tal que:

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

Ejemplo:

Sea $f(x, y) = x^2 - 2y^2 + 4xy - 3$, calculemos el Polinomio de Taylor de orden 2 en el punto $(1, 1) \in \mathbb{R}^2$. Observamos que por la definición que se ha dado, el polinomio de Taylor de orden dos se puede expresar como:

$$\begin{cases} \nabla f(x, y) = (2x + 4y, -4y + 4x) \\ H_f(x, y) = \begin{pmatrix} 2 & 4 \\ 4 & -4 \end{pmatrix} \\ h = (x, y) - (1, 1) = (x - 1, y - 1) \end{cases} \Rightarrow P_2((x, y), x_0) = f(x_0) + \nabla f(x_0) \cdot h + h \cdot H_f(x_0) \cdot h^t$$

Luego, sustituyendo tenemos:

$$\begin{aligned} P_2((x, y); (1, 1)) &= 0 + \langle \overbrace{(6, 0)}^{\nabla f(x_0)}, (x - 1, y - 1) \rangle + \frac{1}{2} \cdot (x - 1 \quad y - 1) \cdot \overbrace{\begin{pmatrix} 2 & 4 \\ 4 & -4 \end{pmatrix}}^{H_f(x_0)} \cdot \begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix} \\ &= 6(x - 1) + \frac{1}{2} \cdot (2(x - 1) + 4(y - 1) \quad 4(x - 1) - 4(y - 1)) \cdot \begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix} \\ &= 6(x - 1) + \frac{1}{2} (2(x - 1)^2 + 4(x - 1)(y - 1) + 4(x - 1)(y - 1) - 4(y - 1)^2) \\ &= x^2 - 2y^2 + 4xy - 3 \end{aligned}$$

Definición (Extremos y puntos críticos)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $x_0 \in A$, entonces:

- Decimos que x_0 es un **mínimo relativo** si $f(x) \geq f(x_0) : \forall x \in B(x_0, \varepsilon) \subset A$
- Decimos que x_0 es un **mínimo relativo** si $f(x) \leq f(x_0) : \forall x \in B(x_0, \varepsilon) \subset A$
- Decimos que un **extremo es absoluto** si la desigualdad correspondiente es cierta $\forall x \in A$

- Decimos que x_0 es un **punto crítico** si $\nabla f(x_0) = 0$
- Decimos que x_0 es un **punto silla** si es un punto crítico que no es un extremo relativo.

Teorema

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto y supuesto que existe $\nabla f(x_0)$. Si $x_0 \in A$ es un extremo relativo, entonces $\nabla f(x_0) = 0$.

Demostración:

Sea $j \in \{1, \dots, n\}$. Supongamos que x_0 es un máximo relativo (razonamos de forma análoga para el caso del mínimo relativo). Entonces

$$\frac{f(x_0 + t \cdot e_j) - f(x_0)}{t} = \begin{cases} \leq 0 & t > 0 \\ \geq 0 & t < 0 \end{cases} \Rightarrow \frac{\partial f}{\partial x_j}(x_0) = 0$$

Observación:

Si $f \in \mathcal{C}^2(A)$, x_0 , definimos el Hessiano como $H_f(x_0) = \left(\frac{\partial^2 f}{\partial x_j \partial x_k} \right)_{1 \leq j, k \leq n}$, es decir, es una matriz cuadrada simétrica. Si recordamos del curso de Álgebra Lineal, definíamos las formas cuadráticas como las aplicaciones $Q \in M_{n \times n}$ tales que

$$Q : \mathbb{R}^n \rightarrow \mathbb{R} \\ x \mapsto Q(x) = xQx^t$$

Es decir, los Hessianos son formas cuadráticas.

Definición (Forma cuadrática)

Una forma cuadrática $Q \in M_{n \times n}$ se denomina:

- Definida positiva si $Q(x) > 0 : \forall x \in \mathbb{R}^n \setminus \{0\}$
- Definida negativa si $Q(x) < 0 : \forall x \in \mathbb{R}^n \setminus \{0\}$
- Semidefinida positiva si $Q(x) \geq 0 : \forall x \in \mathbb{R}^n$
- Semidefinida negativa si $Q(x) \leq 0 : \forall x \in \mathbb{R}^n$

Teorema (Caracterización de los puntos críticos vía Hessiano)

Sea $f \in \mathcal{C}^2(A)$ con $A \subset \mathbb{R}^n$, $x_0 \in A$ y $\nabla f(x_0) = 0$, se tiene que:

- Si $H_f(x_0)$ es definida negativa, entonces x_0 es un máximo relativo.
- Si $H_f(x_0)$ es definida positiva, entonces x_0 es un mínimo relativo.
- Si x_0 es un máximo relativo, entonces $H_f(x_0)$ es semidefinida negativa.
- Si x_0 es un mínimo relativo, entonces $H_f(x_0)$ es semidefinida positiva.
- Si $H_f(x_0)$ es indefinida, entonces x_0 es un punto silla.

Demostración:

Se muestra la demostración del primer apartado solo, puesto que las demás son completamente análogas a la que siguen.

Sea $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ compacto de \mathbb{R}^n , como la aplicación $H_f(x_0) = Q_{x_0} : \mathbb{R}^n \rightarrow \mathbb{R}$ es continua, alcanza un máximo y mínimo en S . Además, como Q es definida negativa, por hipótesis se tiene que:

$$\exists \tilde{x} = \max\{H_f(x_0)(x)\} \in S : 0 > Q(\tilde{x}) \geq Q(x) : \forall x \in S$$

Sea $-\varepsilon_0 = Q(\tilde{x})$ y teniendo en cuenta a la continuidad de las derivadas parciales de segundo orden de f , se tiene que:

$$\text{Dado } \varepsilon_0 > 0, \exists \delta > 0 : \|x - x_0\| < \delta \Rightarrow |D_{ij}f(x) - D_{ij}f(x_0)| < \frac{\varepsilon_0}{2n^2} : \forall i, j = 1, \dots, n$$

Por otra parte, si $h \in \mathbb{R}^n$ no nulo, se verifica que:

$$Q(h) = \|h\|^2 \cdot Q\left(\frac{h}{\|h\|}\right) \leq \|h\|^2 \cdot Q(\tilde{x}) = -\varepsilon_0 \cdot \|h\|^2$$

Por el Teorema de Taylor, $\exists c \in L[x_0, x_0 + h]$ de forma que para $\|h\| < \delta$ ocurre:

$$\begin{aligned} f(x_0 + h) - f(x_0) &= \frac{1}{2} \cdot h \cdot H_f(c) \cdot h^t = \frac{1}{2} \sum_{i=1}^n D_{ij}f(c) \cdot h_i \cdot h_j \\ &= \frac{1}{2} \left[\left(\sum_{i=1}^n D_{ij}f(c) - D_{ij}f(x_0) \right) \cdot h_i \cdot h_j + \underbrace{\sum_{i=1}^n D_{ij}f(x_0) \cdot h_i \cdot h_j}_{Q(h)} \right] \leq \frac{1}{2} \left[n^2 \frac{\varepsilon}{2n^2} \cdot \|h\|^2 - \|h\|^2 \varepsilon_0 \right] < 0 \end{aligned}$$

El resto de apartados se demuestran de forma análoga.

Observación:

Sea $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R} \in \mathcal{C}^2(A)$ con A abierto y $(x_0, y_0) \in A$, entonces el Hessiano es de la forma:

$$H_f(x_0, y_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Es decir, si denotamos al determinante Hessiano con $\Delta = \det(H_f(x_0, y_0)) = ac - b^2$, entonces para si $h = (h_1, h_2) \in \mathbb{R}^2$, tenemos

$$hH_f(x_0, y_0)h^t = a \cdot h_1^2 + 2b \cdot h_1 \cdot h_2 + c \cdot h_2^2$$

Lo que da lugar al siguiente teorema.

Teorema

Sea $\nabla f(x_0, y_0) = 0$, entonces considerando el determinante del Hessiano:

- $\Delta > 0 = \begin{cases} a > 0 & (x_0, y_0) \text{ mínimo relativo} \\ a < 0 & (x_0, y_0) \text{ máximo relativo} \end{cases}$
- $\Delta < 0 \Rightarrow (x_0, y_0) \text{ es punto silla}$
- $\Delta = 0$

En el fondo esto es extensible a dimensiones mayores que 2 utilizando los criterios del Álgebra Lineal para decidir cuando una forma cuadrática (el Hessiano) es definida positiva, definida negativa, etc.

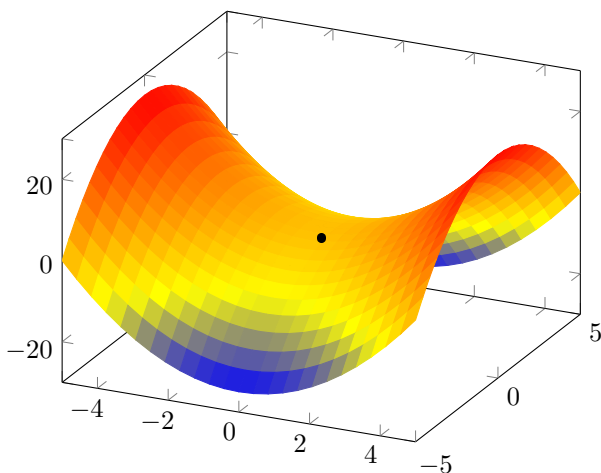
Demostración:

Demostraremos sólo el primer apartado puesto que los demás se basan en una demostración por casos como la que sigue.

Lo que queremos ver es que si $h \neq 0$, entonces $ah_1^2 + 2bh_1h_2 + ch_2^2 > 0$. Sabemos que, por hipótesis, $ac - b^2 > 0$ y que $a > 0$. La clave está en escribir $ah_1^2 + 2bh_1h_2 + ch_2^2 = a \left(h_1 + \frac{b}{a}h_2\right)^2 + \frac{\Delta}{a}h_2^2$ donde dicha igualdad viene de:

$$a \left(h_1 + \frac{b}{a}h_2\right)^2 + \frac{\Delta}{a}h_2^2 = a \left(h_1^2 + \frac{b^2}{a^2}h_2^2 + 2 \cdot \frac{b}{a}h_1h_2\right) + \left(c - \frac{b^2}{a}\right)h_2^2$$

Gracias a las hipótesis de partida, tenemos que $a \left(h_1 + \frac{b}{a}h_2\right)^2 + \frac{\Delta}{a}h_2^2 > 0$ ya que, aunque podría darse que $h_1 + \frac{b}{a}h_2$ pueda anularse, el segundo término no.



Ejemplo de Punto de Silla

Ejemplo:

Estudia los puntos críticos de las siguiente función y determina si son máximos o mínimos locales:

$$f(x, y) = x^3 + y^3 - 3 \cdot axy \Rightarrow \nabla f(x, y) = (3x^2 - 3ay, 3y^2 - 3ax)$$

- Supongamos $a \neq 0$, entonces igualando a 0 ambas componentes:

$$\begin{cases} 3x^2 - 3ay = 0 \\ 3y^2 - 3ax = 0 \end{cases} \Rightarrow \begin{cases} y = \frac{x^2}{a} \\ \frac{x^4}{a^2} - ax = 0 \Rightarrow x(x^3 - a^3) = 0 \end{cases} \Rightarrow \begin{cases} x = 0 & y = 0 \\ x = a & y = a \end{cases}$$

- Si $a = 0$ trivialmente se tiene que sólo es punto crítico el $(0, 0)$.
- Si calculamos el determinante del Hessiano, tenemos que:

$$H_f(x, y) = \begin{pmatrix} 6x & -3a \\ -3a & 6y \end{pmatrix} \Rightarrow \Delta = 36xy - 9a^2$$

- Si $a \neq 0$, tenemos que:
 - $(0, 0) \Rightarrow \Delta = -9a^2 < 0 \Rightarrow$ punto silla
 - $(a, a) \Rightarrow \Delta = 27a^2 > 0 \Rightarrow \begin{cases} a > 0 & \text{mínimo relativo} \\ a < 0 & \text{máximo relativo} \end{cases}$
- Si $a = 0$, entonces $\Delta = 0$, así que no podemos deducir nada de ello. Es necesario **calcularlo a mano** y finalmente sale $f(0, 0) = 0$. completar

TEOREMAS NOTABLES

En este capítulo vamos a enunciar y demostrar tres resultados fundamentales para otros ámbitos de las matemáticas como son las ecuaciones en derivadas parciales o las ecuaciones algebraicas. Su estudio nace de la necesidad de poder operar y trabajar con funciones como si de números se tratase para poder realizar operaciones como dividir (multiplicar por el inverso) o poder expresar dicha función de una forma más conveniente.

TEOREMA DE LA FUNCIÓN INVERSA

Nuestro objetivo ahora es dada $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, cuándo y de qué forma podemos encontrar la solución:

$$f(x) = y : \begin{cases} f_1(x_1, \dots, x_n) = y_1 \\ \vdots \\ f_n(x_1, \dots, x_n) = y_n \end{cases}$$

Cuando trabajamos en \mathbb{R} , si tenemos $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ de forma que $f \in \mathcal{C}^1$ y $x_0 \in I$ un punto donde la derivada no se anula $f'(x_0) \neq 0$, entonces f es estrictamente monótona en un entorno de x_0 y, por lo tanto, invertible en un entorno de $x_0 \in I$. Además, si $g = f^{-1}$ sabemos calcular la derivada:

$$g'(f(x)) = \frac{1}{f'(x)}$$

Luego, con este teorema pretendemos hacer extensible dicha demostración a dimensión arbitraria y como vimos que $df(x_0)$ jugaba el papel de $f'(x_0)$ será razonable pensar que la condición imprescindible será $\det(df(x_0)) \neq 0$.

Definición

Para las posteriores demostraciones y enunciados, vamos a definir dos objetos fundamentales a los que haremos referencia:

$$L_i(\mathbb{R}^n, \mathbb{R}^n) = \{M \in M_{n \times n} : \det(M) \neq 0\}$$

Es decir, las matrices cuadradas invertibles de orden n .

$$\begin{aligned} \mathcal{L}^{-1} : L_i(\mathbb{R}^n, \mathbb{R}^n) &\rightarrow L_i(\mathbb{R}^n, \mathbb{R}^n) \\ M &\mapsto \mathcal{L}^{-1}(M) = M^{-1} \end{aligned}$$

Es decir, la aplicación invertir matriz cuando sabemos que esta tiene inversa.

Observación:

Es importante ver que podemos identificar $M_{n \times n}$ con el espacio euclídeo \mathbb{R}^{n^2} , puesto que podemos realizar la siguiente asignación:

$$M = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \mapsto (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{n1}, \dots, a_{nn})$$

Y de este modo, $L_i(\mathbb{R}^n, \mathbb{R}^n)$ se puede considerar como un subconjunto de \mathbb{R}^{n^2} , luego tiene sentido hablar de abiertos en dicho conjunto.

Lema

El conjunto $L_i(\mathbb{R}^n, \mathbb{R}^n)$ es un abierto en $M_{n \times n} = \mathbb{R}^{n^2}$ y la aplicación \mathcal{L}^{-1} es de clase C^∞ .

Demostración:

Sea $M \in L_i(\mathbb{R}^n, \mathbb{R}^n)$, por definición $\det(M) \neq 0$, es decir, $L_i(\mathbb{R}^n, \mathbb{R}^n) = d^{-1}(\mathbb{R} \setminus \{0\})$ donde d es la aplicación $M \xrightarrow{d} \det(M)$. Trivialmente $\mathbb{R} \setminus \{0\}$ es un abierto y como d es una función continua por ser una combinación polinómica de las entradas de la matriz, la inversa del abierto $(\mathbb{R} \setminus \{0\})$ es abierto.

Análogamente vemos que $\mathcal{L}^{-1} \in C^\infty(L_i(\mathbb{R}^n, \mathbb{R}^n))$. Por ejemplo:

- Si $n = 1$, entonces $\mathcal{L}^{-1}(a) = \frac{1}{a} : a \neq 0$ y, por tanto, $\mathcal{L}^{-1} \in C^\infty$
- Si $n = 2$ y definiendo $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, denotamos $\Delta = \det(M) = ad - bc \neq 0$ así que:

$$\mathcal{L}^{-1}(M) = \begin{pmatrix} \frac{d}{\Delta} & \frac{-b}{\Delta} \\ \frac{-c}{\Delta} & \frac{a}{\Delta} \end{pmatrix}$$

es una función racional, lo que implica $\in C^\infty(L_i(\mathbb{R}^2, \mathbb{R}^2))$

- Y de este modo, basta con ver como se trata el determinante en cada dimensión para poder hacerlo por inducción.

Lema (Teorema del Punto Fijo)

Sea (X, d) espacio métrico completo y sea $f : X \rightarrow X$ tal que:

$$\forall x, y \in X : d(f(x), f(y)) \leq k \cdot d(x, y) \text{ con } 0 < k < 1$$

Es decir, f es una contracción, entonces existe un único punto fijo:

$$\exists! x \in X : f(x) = x$$

Demostración:

Sea $x_1 \in X$ y $x_{n+1} = f(x_n)$, en primer lugar probaremos que es de Cauchy. Sea $n \geq 2$, tenemos que:

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq k \cdot d(x_n, x_{n-1}) \leq k^2 \cdot d(x_{n-1}, x_{n-2}) \leq \dots \leq k^{n-1} d(x_2, x_1)$$

Ahora, si tomamos $m > n \geq 1$, ocurre que:

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \dots + d(x_{n+1}, x_n) \leq (k^{m-2} + \dots + k^{n-1}) \cdot d(x_2, x_1) \\ &= \frac{k^{n-1} - k^{m-1}}{1 - k} \cdot d(x_2, x_1) \xrightarrow{m, n \rightarrow \infty} 0 \end{aligned}$$

Por completitud¹, lo primero, y por continuidad, lo segundo, tenemos que:

$$\exists \lim_{n \rightarrow \infty} x_n = x_0 \quad x_0 \leftarrow x_{n+1} = f(x_n) \rightarrow f(x_0) \Rightarrow x_0 = f(x_0)$$

Por último, falta comprobar que es único. Supongamos que no, es decir, $f(x_0) = x_0$ y $f(y_0) = y_0$, entonces:

$$d(x_0, y_0) = d(f(x_0), f(y_0)) \leq kd(x_0, y_0)$$

Como k está entre 0 y 1, la única forma de que se cumpla es que $d(x_0, y_0) = 0$, es decir, $x_0 = y_0$.

Teorema (de la Función Inversa)

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ donde A es abierto y tomando $x_0 \in A$ si se verifican:

- $f \in \mathcal{C}^p(A)$ con $p \geq 1$
- $\det(df(x_0)) \neq 0$

Entonces se cumplen:

- $\exists \mathcal{U}, \mathcal{V}$ abiertos tal que $x_0 \in \mathcal{U} \subset A$ y $f(x_0) \in \mathcal{V} \subset f(A)$
- $f|_{\mathcal{U}} : \mathcal{U} \rightarrow \mathcal{V}$ es invertible
- $f^{-1} : \mathcal{V} \rightarrow \mathcal{U} \in \mathcal{C}^p(\mathcal{V})$
- $df^{-1}(f(x)) = d(f(x))^{-1} = \frac{1}{df(x)}$

Demostración:

En esta demostración necesitamos simplificar el caso general al caso específico que se resuelve de forma más sencilla, luego los primeros párrafos de esta demostración irán dirigidos a demostrar que no hay pérdida de generalidad al restringirse a este caso especial.

En primer lugar, veamos que podemos suponer que $df(x_0) = I$. En efecto, llamemos $M = df(x_0)$, entonces $\exists M^{-1}$ y por la regla de la cadena²:

$$d(M^{-1} \circ f)(x_0) = dM^{-1}(f(x_0)) \cdot df(x_0) = M^{-1} \cdot M = I$$

Por lo tanto, si demostrásemos el teorema para $M^{-1} \circ f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ y existiese g inversa de esa función, entonces $g \circ M^{-1}$ sería inversa de f .

Del mismo modo, podemos suponer que $f(x_0) = x_0 = 0$, ya que supone hacer un cambio de base en \mathbb{R}^n para hacer que el punto $(x_0, f(x_0))$ sea el origen. Una vez demostrado en esas condiciones, para el caso general bastaría con tomar $h(x) = f(x + x_0) - f(x_0)$ porque $h(0) = 0$ y $dh(0) = df(x_0)$ y entonces una vez calculada la inversa de $h(x)$ en un entorno de 0, la inversa de $f(x)$ sería $f^{-1}(y) = h^{-1}(y - f(x_0)) + x_0$.

■ Invertibilidad

En las condiciones de $df(x_0) = I$ y $x_0 = f(x_0) = 0$, definimos $g(x) = x - f(x)$ de manera que se tiene $g(0) = 0$ y $dg(0) = 0$. Aplicando el Teorema del Valor Medio tenemos que:

$$\exists \delta > 0 : \text{ si } \|x\| < \delta \Rightarrow \|g(x) - g(0)\| \leq \|df(c)\| \cdot \|x\|$$

Como $dg(0) = 0$, entonces $\exists \delta' > 0 : \|dg(x)\| < \varepsilon = \frac{1}{2}$ si $\|x\| < \delta'$, luego como $c \in B(0, \delta')$ volviendo a la expresión anterior:

$$\|g(x)\| = \|g(x) - g(0)\| \leq \|df(c)\| \cdot \|x\| \leq \frac{1}{2} \cdot \|x\| \leq \frac{\delta}{2}$$

¹Además dicho x_0 sabemos que pertenece a X por ser completo

²Ver que $dM^{-1}(f(x_0)) = M^{-1}$ es trivial si calculamos la diferencial a mano (no necesariamente para $f(x_0)$)

De este modo, podemos restringir g a los dominios $g : \bar{B}(0, \delta) \rightarrow \bar{B}(0, \frac{\delta}{2})$ y $g : B(0, \delta) \rightarrow B(0, \frac{\delta}{2})$. Sea $y \in \bar{B}(0, \frac{\delta}{2})$, y definimos:

$$g_y(x) = y + g(x) = y + x - f(x)$$

Para poder tener $y = f(x)$ es necesario que $g_y(x) = x$, luego queremos encontrar un punto fijo para esta función. Tenemos y ya fijo, que es el que queremos invertir, o dicho de otro modo, probar que hay una antiimagen única. Para ello, probaremos que $g_y(x)$ es una contracción. En primer lugar, sea $\|x\| < \delta$, entonces:

$$\|g_y(x)\| = \|y + g(x)\| \leq \|y\| + \|g(x)\| \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta \Rightarrow g_y : \bar{B}(0, \delta) \rightarrow \bar{B}(0, \delta)$$

Dichos conjuntos sobre los que están definida la función son compactos en \mathbb{R}^n y como éste es completo, entonces también son completos. Ahora veamos que la aplicación es contractiva, sean $x_1, x_2 \in \bar{B}(0, \delta)$, tenemos:

$$\|g_y(x_1) - g_y(x_2)\| = \|g(x_1) - g(x_2)\| \stackrel{TV_M}{\leq} \|dg(c)\| \cdot \|x_1 - x_2\| \leq \frac{1}{2} \|x_1 - x_2\|$$

Es decir, que g_y es una contracción con $k = \frac{1}{2}$, y por el Teorema del Punto Fijo, $\exists! x \in \bar{B}(0, \delta) : g_y(x) = x \Rightarrow f(x) = y$, es decir, f es invertible.

■ Regularidad

Veamos que f^{-1} es continua, sean $y_1, y_2 \in \bar{B}(0, \frac{\delta}{2})$ y sean $x_1, x_2 \in \bar{B}(0, \delta)$ tales que $f(x_1) = y_1$ y $f(x_2) = y_2$, tenemos³ que:

$$\|f^{-1}(y_1) - f^{-1}(y_2)\| = \|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\| + \|g(x_1) - g(x_2)\| \leq \|f(x_1) - f(x_2)\| + \frac{1}{2} \|x_1 - x_2\|$$

De este modo, obtenemos $\|x_1 - x_2\| \leq 2 \cdot \|f(x_1) - f(x_2)\|$, así que:

$$\|x_1 - x_2\| = \|f^{-1}(y_1) - f^{-1}(y_2)\| \leq 2 \cdot \|y_1 - y_2\|$$

Es decir, f^{-1} es una función Lipschitz, por lo que $f^{-1} \in \mathcal{C}(\bar{B}(0, \frac{\delta}{2}))$.

Veamos que f^{-1} es diferenciable en $B(0, \frac{\delta}{2})$. Como $df(0) = I$, $\exists \delta > 0 : \|x\| < \delta : \det(df(x)) \neq 0$ por ser las matrices invertibles un conjunto abierto en \mathbb{R}^{n^2} . Tomemos $y_1, y_2 \in B(0, \delta')$ y denotemos por $x_1 = f^{-1}(y_1)$, $x_2 = f^{-1}(y_2)$. Si escogemos un δ' tal que $f(B(0, \delta)) \subset B(0, \delta') \subset B(0, \frac{\delta}{2})$, veamos que la función es diferenciable y que su diferencial es la inversa de la diferencial:

$$\begin{aligned} \frac{\|f^{-1}(y_1) - f^{-1}(y_2) - (df(x_2))^{-1}(y_1 - y_2)\|}{\|y_1 - y_2\|} &= \frac{\|x_1 - x_2 - (df(x_2))^{-1}(f(x_1) - f(x_2))\|}{\|f(x_1) - f(x_2)\|} \\ &= \frac{\|x_1 - x_2\|}{\|f(x_1) - f(x_2)\|} \cdot \frac{\|x_1 - x_2 - (df(x_2))^{-1}(f(x_1) - f(x_2))\|}{\|x_1 - x_2\|} \\ &\leq \frac{2 \cdot \|df(x_2)[df(x_2)(x_1 - x_2) - f(x_1) + f(x_2)]\|}{\|x_1 - x_2\|} \\ &\leq 2\|(df(x_2))^{-1}\| \cdot \frac{\|df(x_2)(x_1 - x_2) - f(x_1) + f(x_2)\|}{\|x_1 - x_2\|} \xrightarrow{x_1 - x_2} 0 \end{aligned}$$

Es decir, $df^{-1}(f(x_2)) = (df(x_2))^{-1}$.

Para terminar, veamos que $f^{-1} \in \mathcal{C}^p$, es decir:

$$df^{-1}(y) = df^{-1}(f(x)) = (df(x))^{-1} = \mathcal{L}^{-1}(df(f^{-1}(y)))$$

Tenemos que $\mathcal{L}^{-1} \in \mathcal{C}^\infty$, $df \in \mathcal{C}^{p-1}$ $f^{-1} \in \mathcal{C}$. Entonces, $f^{-1} \in \mathcal{C}^1$

Así, si $f^{-1} \in \mathcal{C}^1$ y $p \geq 2$, entonces $df^{-1} \in \mathcal{C}^1$ y $f^{-1} \in \mathcal{C}^2$. De forma inductiva, si $f^{-1} \in \mathcal{C}^{p-1}$, entonces $df^{-1} \in \mathcal{C}^{p-1}$ y $f^{-1} \in \mathcal{C}^p$

³El primer menor o igual es por la desigualdad triangular y la definición de g

Observación:

La hipótesis $df(x_0) \neq 0$ en el teorema de la función inversa es fundamental para garantizar la invertibilidad de la diferencial, pero no es imprescindible para garantizar la existencia de una inversa. Por ejemplo, sea $f : \mathbb{R} \rightarrow \mathbb{R}$ donde $f(x) = x^3$. Tenemos que si $x_0 = 0$, existe $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ y, sin embargo, $f'(0) = 0$. Pero como habíamos dicho, f^{-1} no es derivable en 0 porque no se cumpla la hipótesis especificada.

Lema

Sea $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ donde A es abierto y $f \in \mathcal{C}^1(A)$, se tiene que:

$$\forall x \in A : |df(x)| \neq 0 \Rightarrow \forall \mathcal{U} \subset A \text{ abierto} : f(\mathcal{U}) \text{ abierto}$$

En este caso, se dice que f es una aplicación abierta.

TEOREMA DE LA FUNCIÓN IMPLÍCITA

En el Teorema de la Función Implícita queremos estudiar si en la ecuación $x + \sin y + y = 0$ podemos despejar y en función de la variable x , al menos localmente. Claramente $x = -y - \sin x$ nos da la solución si cambiamos el papel de las variables.

Teorema (de la Función Implícita)

Sea $F : A \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ si se verifica que:

- $F \in \mathcal{C}^p(A)$
- $F(x_0, y_0) = 0$
- $\begin{vmatrix} \frac{\partial F_1}{\partial y_1}(x_0, y_0) & \dots & \frac{\partial F_1}{\partial y_m}(x_0, y_0) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1}(x_0, y_0) & \dots & \frac{\partial F_m}{\partial y_m}(x_0, y_0) \end{vmatrix} \neq 0$

Entonces, se cumple que:

- $\exists \mathcal{U} \subset \mathbb{R}^n$ y $\exists \mathcal{V} \subset \mathbb{R}^m$ abiertos de forma que $x_0 \in \mathcal{U}$ y $y_0 \in \mathcal{V}$
- $\exists ! f : \mathcal{U} \rightarrow \mathcal{V}$ donde $f \in \mathcal{C}^p(\mathcal{U})$ tal que $\forall x \in \mathcal{U} : F(x, f(x)) = 0$

Demostración:

Definimos $G(x, y) = (x, F(x, y))$ donde $(x, y) \in A \subset \mathbb{R}^n \times \mathbb{R}^m$, entonces queda definida como:

$$G : A \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m} \Rightarrow G \in \mathcal{C}^p(A)$$

Se prueba fácilmente que $\det(dG(x_0, y_0)) \neq 0$, luego por el Teorema de la Función Inversa:

$$\begin{cases} \exists \mathcal{W} \subset \mathbb{R}^{n+m} : (x_0, y_0) \in \mathcal{W} \\ \exists \mathcal{W}' \subset \mathbb{R}^{n+m} : (x_0, 0) \in \mathcal{W}' \\ G : \mathcal{W} \rightarrow \mathcal{W}' \text{ biyectiva} \end{cases} \Rightarrow G^{-1} \in \mathcal{C}^p(\mathcal{W}')$$

Ahora vamos a restringir aún más dicho dominio. Siempre existe⁴ un cubo centrado en (x_0, y_0) tal que:

$$Q(x_0, y_0) = \underbrace{Q_1(x_0)}_{\subset \mathbb{R}^n} \times \underbrace{Q_2(y_0)}_{\subset \mathbb{R}^m} \subset \mathcal{W}$$

⁴Dicha demostración no procede en este resultado y puede tomarse como un axioma geométrico hasta que se vea en asignaturas posteriores

Denotamos $\mathcal{U}_1 = Q_1(x_0) \subset \mathbb{R}^n$ abierto, y $\mathcal{V} = Q_2(y_0) \subset \mathbb{R}^m$ abierto. De este modo, si llamamos $G(\mathcal{U}_1 \times \mathcal{V}) = \mathcal{W}''$ y restringimos $G : \mathcal{U}_1 \times \mathcal{V} \rightarrow \mathcal{W}'' \subset \mathcal{W}'$ abierto, tenemos por las consecuencias del Teorema de la función Inversa anteriores que:

$$\begin{cases} G^{-1} : \mathcal{W}'' \rightarrow \mathcal{U}_1 \times \mathcal{V} \\ G^{-1} \in \mathcal{C}^p(\mathcal{W}'') \\ G^{-1}(x, z) = (a, b) \Rightarrow (x, z) = G(a, b) = (a, F(a, b)) \Rightarrow a = x \end{cases}$$

Como b debe estar en función de x y de z , tomamos como definición la expresión $b := H(x, z)$, de este modo, la expresión anterior nos queda como:

$$G^{-1}(x, z) = (x, H(x, z)) \text{ donde } H \in \mathcal{C}^p(\mathcal{W}'') \text{ y } H : \mathcal{W}'' \rightarrow \mathcal{V}$$

Y por ser (x, z) la preimagen de ese par, podemos decir que es igual a $(x, F(x, H(x, z)))$ ya que:

$$G^{-1}(x, z) = (x, H(x, z)) \Rightarrow (x, z) = G(x, H(x, z)) = (x, F(x, H(x, z)))$$

Consideremos ahora la proyección sobre las componentes de \mathbb{R}^m , de modo que:

$$\begin{aligned} \pi_2 : \mathbb{R}^n \times \mathbb{R}^m &\rightarrow \mathbb{R}^m \\ (x, y) &\rightarrow \pi_2(x, y) = y \end{aligned}$$

Es sencillo ver que podemos expresar la F inicial como composición de G con dicha proyección π :

$$F(x, H(x, z)) = (\pi_2 \circ G)(x, H(x, z)) = (\pi_2 \circ G \circ G^{-1})(x, z) = z$$

Para que se entienda la estrategia, nos gustaría ahora poder decir que si tomamos $f(x) = H(x, 0)$, entonces tenemos $F(x, f(x)) = F(x, H(x, 0)) = 0$ y como definidos los conjuntos \mathcal{U} y \mathcal{V} solo faltaba probar eso, habríamos terminado. Sin embargo, no tenemos la certeza de que $H(x, 0)$ tome valores en \mathcal{W}'' , condición imprescindible para poder afirmar que G es invertible. Por este detalle y de modo análogo a cuando dijimos que dado un dominio podíamos escoger un cubo dentro de él donde estuviese contenido el punto que queríamos, ahora queremos ver si el “segmento”⁵ está contenido. Esto no está asegurado, es decir, no podemos afirmar que con seguridad $\mathcal{U}_1 \times \{0\} \subset \mathcal{W}''$, pero lo que sí podemos decir es que como $(x_0, 0)$ es interior, entonces hay un entorno $\mathcal{U} \subset \mathcal{U}_1$ al que pertenece x_0 y que sí que cumple que $\mathcal{U} \times \{0\} \subset \mathcal{W}''$. De este modo, definimos la función $f(x)$ como:

$$\begin{aligned} f : \mathcal{U} &\rightarrow \mathcal{V} \\ x &\mapsto H(x, 0) \end{aligned}$$

Donde $f \in \mathcal{C}^p(\mathcal{U})$ y $F(x, f(x)) = 0 : \forall x \in \mathcal{U}$.

Falta solo ver la **unicidad**. Para ello, supongamos que tenemos dos funciones $F(x, f(x)) = 0$ y $F(x, f^*(x)) = 0$, por tanto:

$$F(x, f(x)) = F(x, f^*(x)) \Rightarrow G(x, f(x)) = (x, F(x, f(x))) = (x, F(x, f^*(x))) = G(x, f^*(x))$$

Como G es inyectiva, $f(x) = f^*(x) : \forall x \in \mathcal{U}$

Ejemplo:

Tenemos la función del ejemplo $h(x) = x + \sin y + y$ y queremos ver si en la ecuación $h(x, y) = 0$, es decir, $x + \sin y + y = 0$ podemos despejar y en función de x .

Tomamos como función $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ donde $F(x, y) = x + \sin y + y$ y como $(x_0, y_0) = (0, 0)$. Como $\det(\frac{\partial F}{\partial y}(0, 0)) = |1 + 1| \neq 0$ nos encontramos en las hipótesis del Teorema de la Función Implícita, luego $\exists f : B(x_0, 0) = (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ tal que $f \in \mathcal{C}^\infty$ que cumple $x + \sin f(x) + f(x) = 0$.

⁵Porque es como dejar fijas todas las \mathbb{R}^m componentes y solo movernos por las \mathbb{R}^n

Caracterizar esta función no suele ser en general sencillo, pero lo que sí podemos caracterizar es su polinomio de Taylor:

$$\frac{\partial h}{\partial x}(x, y) = 1 + \cos(f(x)) \cdot f'(x) + f'(x) = 0 : |x| < \varepsilon$$

$$\frac{\partial h}{\partial x}(0, 0) = 1 + 1 \cdot f'(0) + f'(0) = 0 \Rightarrow f'(0) = \frac{-1}{2}$$

$$\frac{\partial^2 h}{\partial x^2}(x, y) = -\sin f(x)(f'(x))^2 + \cos f(x) \cdot f''(x) + f''(x) = 0$$

$$\frac{\partial^2 h}{\partial x^2}(0, 0) = 0 + f''(0) + f''(0) = 0 \Rightarrow f''(0) = 0$$

Por Taylor, para las x dentro del intervalo donde está definida f , tenemos que $f(x) = \frac{-x}{2} + o(x^2)$

MULTIPLICADORES DE LAGRANGE

Dada $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ con A abierto y $f \in \mathcal{C}^1(A)$. Dado $B \subset A$, podríamos querer optimizar $f|_B : B \rightarrow \mathbb{R}$.

Ejemplo:

Tenemos $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ donde $f(x, y) = x^2 + y^2$ y el conjunto en el que queremos optimizar $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. En general, el conjunto $B \subset A \subset \mathbb{R}^n$ vendrá determinado como intersección de conjuntos de la forma $Z_j = \{x \in A : g_j(x) = 0\}$, siendo $g_j \in \mathcal{C}^1(A) : 1 \leq j \leq m < n$

$$B = \bigcap_{j=1}^m Z_j$$

Teorema (Multiplicadores de Lagrange)

Sea $f, g_1, \dots, g_m : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ donde A es abierto y $m < n$, si se verifican:

- $f, g_1, \dots, g_m \in \mathcal{C}^1(A)$
- $B = \{x \in A : g_j(x) = 0 : j = 1, \dots, m\}$
- $\exists x_0 \in B$ que es extremo local en $f|_B$
- $\{\nabla g_1(x_0), \dots, \nabla g_m(x_0)\}$ es un sistema linealmente independiente.

Entonces se cumple que:

$$\exists \lambda_1, \dots, \lambda_m \in \mathbb{R} : \nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \dots + \lambda_m \nabla g_m(x_0)$$

Demostración:

Sin pérdida de generalidad, como el sistema formado por los gradientes es linealmente independiente, podemos suponer⁶ que las últimas m columnas de $dg(x_0)$, siendo $g = (g_1, \dots, g_m)$, nos dan un menor de rango m . Para simplificar, vamos a distinguir explícitamente entre las $n - m$ primeras componentes y las otras m en forma de dos vectores, de forma que para un $x \in \mathbb{R}^n$ este lo vamos a expresar como $x = (u \in \mathbb{R}^{n-m}, v \in \mathbb{R}^m)$ donde $x_0 = (u_0, v_0)$. A modo de notación, también tendremos en cuenta la restricción sobre cada número de componentes como:

$$d_u g(x) = \left(\frac{\partial g_j}{\partial u_k} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n-m}} \in M_{m \times (n-m)} \quad d_v g(x) = \left(\frac{\partial g_j}{\partial v_k} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq m}} \in M_{m \times m}$$

Como son linealmente independientes el menor formado por las últimas m columnas es no nulo y, por tanto, la restricción sobre ese subespacio es biyectiva sobre su imagen, luego $\exists (d_v g(x_0))^{-1}$. Definida previamente la función g , vemos que cumple las hipótesis del Teorema de la Función Implícita:

$$\begin{cases} g : A \subset \mathbb{R}^{n-m} \times \mathbb{R}^m \rightarrow \mathbb{R}^m \\ g \in \mathcal{C}^1(A) \\ \det(d_v g(x)) \neq 0 \end{cases}$$

Por tanto, se derivan las siguientes consecuencias para la igualdad $g(u, v) = 0$:

$$\begin{cases} \exists \mathcal{U} \subset \mathbb{R}^{n-m} \ni u_0, \exists \mathcal{V} \subset \mathbb{R}^m \ni v_0 \\ \exists ! h : \mathcal{U} \rightarrow \mathcal{V} : h(u_0) = v_0 \text{ y } h \in \mathcal{C}^1(\mathcal{U}) \\ g(u, h(u)) = 0 : \forall u \in \mathcal{U} \end{cases}$$

⁶Porque en caso de que no ocurra así basta con reordenar y cambiar las bases

Como $\forall u \in \mathcal{U} : g(u, h(u)) = 0$, si llamamos $x = (u, h(u))$, tenemos que:

$$0 = \underbrace{d_u g(x)}_{\in m \times n-m} + \underbrace{d_v g(x)}_{\in m \times m} \cdot \underbrace{dh(u)}_{\in m \times n-m} \quad (3.1)$$

Como $f(u, h(u))$ tiene un extremo en $u_0 \in \mathcal{U}$, tenemos que:

$$0 = \nabla_u f(x_0) + \nabla_v f(x_0) \cdot d_h(u_0) \quad (3.2)$$

Por la primera ecuación, tenemos que $dh(u) = -(d_v g(x))^{-1} \cdot d_u g(x)$, es decir, $dh(u_0) = -(d_v g(x_0))^{-1} \cdot d_u g(x_0)$. Si sustituimos dicha expresión en la segunda ecuación, tenemos que:

$$0 = \nabla_u f(x_0) - \nabla_v f(x_0) (d_v g(x_0))^{-1} \cdot d_u g(x_0) \quad (3.3)$$

Si definimos $(\lambda_1, \dots, \lambda_m) := \nabla_v f(x_0) (d_v g(x_0))^{-1}$, entonces despejando:

$$\nabla_v f(x_0) = \sum_{j=1}^m \lambda_j \cdot \nabla_v g_j(x_0)$$

Y de la tercera ecuación, tenemos:

$$\nabla_u f(x_0) = \sum_{j=1}^m \lambda_j \cdot \nabla_u g_j(x_0)$$

Luego ya tenemos probadas ambas cosas

Volviendo al ejemplo anterior, el conjunto de ecuaciones a resolver sería:

$$\begin{cases} \nabla f = \lambda \nabla g \\ g = 0 \end{cases} \Rightarrow \begin{cases} (2x, 1) = \lambda(2x, 2y) \\ x^2 + y^2 + 1 = 0 \end{cases} \Rightarrow \begin{cases} x = \lambda x \\ 1 = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

De la primera ecuación concluimos que o bien $x = 0$ o $\lambda = 1$. Si $x = 0$, la solución es $(0, \pm 1)$. Si $\lambda = 1$, $(\frac{\pm\sqrt{3}}{2}, \frac{1}{2})$.