

Cálculo 2

Apuntes de Clase

Segunda Edición

Instituto de Matemática y Estadística

Facultad de Ingeniería.

UDELAR

24 de julio de 2013

Índice general

Prólogo	4
Cambios de Notación	5
6. Funciones de Varias Variables	7
6.0. Introducción	7
6.1. Propiedades Topológicas de \mathbb{R}^n	8
1. Sucesiones en \mathbb{R}^n	13
6.2. Límites y Continuidad	18
6.3. Derivadas Parciales y Direccionales	23
6.4. Diferenciabilidad	26
1. Condición Suficiente de Diferenciabilidad	29
2. Diferencial de una Función	32
3. Regla de la Cadena	33
4. Curvas y Superficies de Nivel	35
6.5. Derivadas de Mayor Orden	37
1. Funciones de Clase C^k	40
2. Diferenciales de Mayor Orden	40
3. Desarrollo de Taylor	41
6.6. Extremos	44
1. Reconocimiento de Puntos Críticos	46
2. Extremos Absolutos	52

6.7. Funciones Implícitas y Aplicaciones	53
1. Extremos Condicionados (o Ligados)	59
2. Extensiones del Método	63
6.8. Funciones de \mathbb{R}^n en \mathbb{R}^m	67
1. Continuidad	67
2. Diferenciabilidad	68
3. Funciones Inversas e Implícitas	72
7. INTEGRALES MÚLTIPLES	77
7.1. Integrales Dependientes de un Parámetro	77
1. Continuidad	77
2. Derivadas	79
3. Integrales Iteradas	81
7.2. Medida de Jordan en el Plano	83
7.3. Integrales Dobles	89
1. Interpretación Geométrica	92
2. Propiedades de la Integral Doble	94
7.4. Cambios de Variable en Integrales Dobles	98
7.5. Integrales Múltiples de Dimensión Mayor que 2	102
7.6. Aplicaciones de las Integrales Múltiples	106
7.7. Integrales Múltiples Impropias	111
Índice de Símbolos	118
Índice de Nombres	119

Prólogo de la Segunda Edición

Las siguientes notas están basadas en el “Análisis Matemático I” de Fernando Paganini, usado como guía para el curso de Cálculo 2 durante el segundo semestre de

2012. Hemos seguido fielmente las notas de F. Pagannini, salvo por algunas modificaciones de notación que listamos en **Cambios de notación**. En algunos casos hemos dado demostraciones alternativas, sacado algún ejemplo y agregado otros, pero siempre intentando mantener el estilo conciso y directo del texto original. Con el fin de mantener cierta compatibilidad en las referencias, hemos mantenido la numeración original de las definiciones, teoremas, proposiciones, etc. y además hemos incluido notas en los márgenes indicando la página correspondiente del libro de F. Pagannini.

La idea es que para una próxima edición, se deje de mantener esta estricta regla de compatibilidad.

Como extra al texto original, hemos incluido hiperlinks a lo largo y ancho del texto, así como un índice alfabético con los conceptos y símbolos definidos. Un debe son alguna de las figuras que por problemas de tiempo no hemos podido incluir.

Eduardo Canale

Asunción enero de 2013.

Cambios de notación

- En lugar de la notación $E(a, r)$ usaremos la notación $B(a; r)$. En concordancia usaremos más el término “bola” que el de “entorno”.
- En lugar de denotar A^c al exterior de A , lo denotaremos A^{ext} .
- En lugar de denotar las sucesiones de vectores como $(x^{(k)})_{k \in \mathbb{N}}$ escribiremos simplemente $(x_k)_{k \in \mathbb{N}}$. A su vez, la coordenada i -ésima de la sucesión $(x_k)_{k \in \mathbb{N}}$, la denotaremos $(x_{k i})_{k \in \mathbb{N}}$ en lugar de $(x_i^{(k)})_{k \in \mathbb{N}}$.
- Para distinguir entre el intervalo abierto (a, b) y el par ordenado (a, b) , a veces escribiremos $(a; b)$ para el intervalo y (a, b) para el par.

Capítulo 6

Funciones de Varias Variables

p189 de [5]

6.0. Introducción

En diversas aplicaciones tiene interés estudiar funciones de varias variables, es decir con dominio en \mathbb{R}^n . Para poner un ejemplo tomado de la física, si se quiere describir la temperatura del aire en una región del espacio, se está ante una función del tipo $T(x, y, z)$ donde (x, y, z) son las coordenadas del punto (vector de \mathbb{R}^3) y T es un número real. Se está ante una función de \mathbb{R}^3 en \mathbb{R} . Si ahora queremos describir la velocidad de cada partícula de aire, tenemos una función $V(x, y, z)$ donde V es un vector de \mathbb{R}^3 . Es decir, tenemos una función de \mathbb{R}^3 en \mathbb{R}^3 .

¿Qué se quiere estudiar de estas funciones? En principio, se pretende entender, en la medida de lo posible, toda la teoría disponible para funciones de una variable: continuidad, derivadas, integración, etc. A lo largo de este capítulo y el siguiente se verá que es posible, con algunas limitaciones, esta extensión. En el Capítulo 6 trataremos el cálculo diferencial de varias variables y en el Capítulo 7 la integración. El planteo más general es estudiar las funciones de \mathbb{R}^n en \mathbb{R}^m , es decir con n variables y tomando valores en \mathbb{R}^m . Por simplicidad, se

difiere esto para el final del capítulo y se empieza por el caso $m = 1$; funciones reales de varias variables. Los principales conceptos aparecen en este caso y no es difícil extenderlos.

p190 de [5]

6.1. Propiedades Topológicas de \mathbb{R}^n

Antes de encarar las teorías de funciones, veremos algunas propiedades de distancias, sucesiones y convergencia que serán herramientas básicas más adelante. Estas propiedades corresponden a una rama de la matemática, la topología, en la que entraremos solo mínimamente para introducir lo indispensable para el desarrollo posterior.

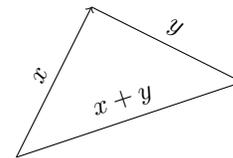
Definición 1. : Si $x \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)$, llamaremos *norma* de x al número real

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

1.0.1. Propiedades

- a) $\|x\| \geq 0$ y $\|x\| = 0 \iff x = 0$,
- b) $\|\lambda x\| \geq |\lambda| \|x\|$, para todo $\lambda \in \mathbb{R}$,
- c) $\|x + y\| \leq \|x\| + \|y\|$ (desigualdad triangular).

Las primeras dos son inmediatas a partir de la Definición 1. Para c), remitimos al curso de Algebra Lineal.



Definición 2. Si $x, y \in \mathbb{R}^m$, se llama *distancia* entre x e y al número real: $d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$

1.0.2. Propiedades

- a) $d(x, y) \geq 0$ y $d(x, y) = 0 \iff x = y$,

b) $d(x, y) = d(y, x),$

c) $d(x, y) \leq d(x, z) + d(z, y).$

Las primeras dos son inmediatas, la tercera se deduce de Propiedades 1 c) así:

$$d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$$

Definición 3. Sea $p \in \mathbb{R}^n$, $r > 0$. Se llama *entorno* o *bola abierta* de *centro* p y *radio* r al conjunto:

$$B(p; r) = \{x \in \mathbb{R}^n : \|x - p\| < r\} = \{x \in \mathbb{R}^n / d(x, p) < r\},$$

también denotado $E(p; r)$.

1.0.2. Observación: También escribiremos $B_p(r)$ para denotar la bola de centro p y radio r y cuando no nos interesa el radio, escribiremos B_p , es decir, B_p significa "alguna bola de centro p ".

1.0.3. Ejemplo:: Para $n = 2$ una bola es el interior un círculo, sin incluir la circunferencia. Para $n = 3$, es el interior una esfera sin incluir la cáscara esférica. ¿qué sería para $n = 1$?

Se llama *entorno reducido* o *bola reducida* de centro p y radio r al conjunto

$$B^*(p; r) = B(p; r) - \{p\}.$$

Es decir, es el entorno excluyendo su centro.

Definición 4. Sea $A \in \mathbb{R}^n$ un conjunto, $x \in \mathbb{R}^n$. A^c denota el *complemento* de A .

- x es un *punto interior* a A sii existe $B_p \subset A$,
- x es un *punto exterior* a A sii existe $B_p \subset A^c$,

- x es un *punto frontera* de A sii no es ni interior ni exterior.

Es decir, un punto es interior a A si esta “rodeado” por A , es exterior si esta “rodeado” por A^c , y frontera si tiene puntos cercanos de A y de A^c .

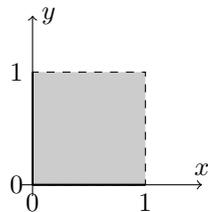
p192 de [5]

1.0.5. Notación

- $\overset{\circ}{A} = \{ \text{puntos interiores a } A \}$
- $A^{ext} = \{ \text{puntos exteriores a } A \}$
- $\partial A = \{ \text{puntos frontera de } A \}$
- $\bar{A} = A \cup \partial A$ se llama conjunto *clausura* de A .

1.0.6. Ejemplos

1) Sea $A = \{(x_1, x_2) \in \mathbb{R}^2 / 0 \leq x < 1, 0 \leq y < 1\}$, entonces



$$\overset{\circ}{A} = \{(x_1, x_2) \in \mathbb{R}^2 / 0 < x < 1, 0 < y < 1\},$$

$$A^{ext} = \{(x_1, x_2) \in \mathbb{R}^2 / |x - 1/2| > 1/2, \text{ o } |y - 1/2| > 1/2\},$$

$$\partial A^{ext} = \text{los lados del cuadrado},$$

$$\bar{A} = \{(x_1, x_2) \in \mathbb{R}^2 / 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

2) Si $A = (a, b) \subset \mathbb{R}$, entonces $\overset{\circ}{A} = A$, $A^{ext} = [a, b]$, $\partial A = \{a, b\}$ y $\bar{A} = [a, b]$.

En el conjunto del Ejemplo 2), se observa que todos los puntos de A son interiores. Esto nos lleva a la siguiente definición:

Definición 5. $A \subset \mathbb{R}^n$ es un *conjunto abierto* sii todos sus puntos son interiores, o sea, si $A = \overset{\circ}{A}$.

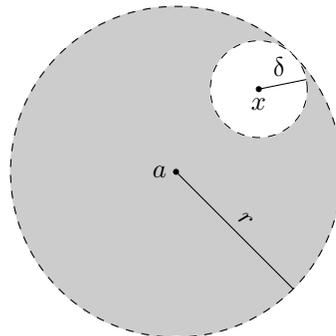
p193 de [5]

1.0.7. Ejemplos

1) Un entorno $B(a; r)$ es un conjunto abierto.

Para demostrarlo, sea $x \in B(a; r)$. Debemos ver que x es interior, o sea que existe $\delta > 0$ tal que $E(x, \delta) \subset B(a; r)$.

Como $x \in B(a; r)$ entonces $d(x, a) < r$ de donde $r - d(x, a) > 0$. Definamos $\delta = r - d(x, a)$. Por lo anterior ya sabemos que $\delta > 0$. Ahora solo resta chequear la inclusión: si $y \in B(x; \delta) \Rightarrow d(x, y) < \delta \Rightarrow d(x, y) + d(x, a) < r$. Por otro lado, por la desigualdad triang., $d(y, a) \leq d(y, x) + d(x, a) = d(x, y) + d(x, a)$. Juntando las dos cosas tenemos que $d(y, a) < r$ o sea que $y \in B(a; r)$.



Debe notarse que en la demostración anterior solo se utilizaron propiedades de la distancia, y no su fórmula explícita. Esto hace pensar en que es posible generalizar lo anterior para una noción de distancia más amplia, en la que se exige solo el cumplimiento de las propiedades 1.0.1. Esto de hecho puede hacerse y conduce a una parte de la topología que es la teoría de **espacios métricos**. No entraremos en ella (ver [2], por ejemplo), pero queda este ejemplo como muestra de las técnicas de trabajo, que pueden también usarse para el siguiente ejemplo.

2) $O = (0, 0, \dots, 0) \in \mathbb{R}^n$. Entonces $\{0\}^\circ$ es un conjunto abierto. Demostración: ejercicio

3) $[a, b] \subset \mathbb{R}^n$ no es un conjunto abierto. De hecho, $a \in [a, b]$ pero no es interior.

Intuitivamente, un conjunto es abierto cuando no contiene a sus bordes (o frontera)(formalmente si $\partial A = \emptyset$ o si $\partial A \neq \emptyset$ pero $\partial A \cap A = \emptyset$. También podríamos escribir $\partial A \subset A^c$). En contraposición, el intervalo cerrado $[a, b]$ del ejemplo 3) contiene a sus bordes. Otra forma de decir esto es que su complemento $(-\infty, a) \cup (b, +\infty)$ es abierto. Obviamente, un conjunto contiene su frontera si su exterior no la contiene (salvo que la misma sea vacía). Esto nos lleva a la siguiente definición:

Definición 6. $A \subset \mathbb{R}^n$ es un *conjunto cerrado* sii A^c es abierto.

1.0.8. Ejemplo

1) $A = \{0\}$, 2) $A = \{x \in \mathbb{R}^n / d(x, a) \leq r\}$.

p194 de [5]

1.0.8. Observación: A es cerrado sii $\partial A \subset A$.

Definición 7. Sea $A \subset \mathbb{R}^n, x \in \mathbb{R}^n$ entonces, diremos que x es un *punto de acumulación* de A sii todo entorno *reducido* de x interseca a A . Al conjunto de puntos de acumulación de A se lo denota A' .

1.0.10. Ejemplos

1) $A = \{1/n : n \in \mathbb{N}\} \subset \mathbb{R}$. 0 es punto de acumulación de A pero 1 no.

2) Todo elemento de \mathbb{R} es punto de acumulación de \mathbb{Q} .

3) Todo elemento de \mathbb{R}^2 es punto de acumulación de \mathbb{Q}^2 .

Proposición 6.1. $A \subset \mathbb{R}^n$ es cerrado sii A contiene a todos sus puntos de acumulación.

Demostración: (\Rightarrow) Por absurdo sea x punto de acumulación de A fuera de A , es decir, $x \in A^c$. Como A es cerrado A^c es abierto, por lo que existirá una bola $B(x; r) \subset A^c$. Pero entonces $B^*(x; r)$ no interseca a A , pues $B^*(x; r) \subset B(x; r)$, contradiciendo que x fuera punto de acumulación de A .

(\Leftarrow) Ejercicio. □

p195 de [5]

Definición 8. $A \subset \mathbb{R}^n$ es *acotado* sii

$$\exists L \in \mathbb{R} / \|x\| \leq L, \forall x \in A.$$

1.0.10. Observación: Equivalentemente, A es acotado sii existe una bola que lo contiene, es decir, si existe B_p tal que $A \subset B_p$. También A es acotado sii $\{\|x\| : x \in A\} \subset \mathbb{R}$ es acotado.

Definición 9. $K \subset \mathbb{R}^n$ es *compacto* sii es cerrado y acotado.

1.0.12. Ejemplos

- 1) (a, b) en \mathbb{R} es acotado pero no compacto.
- 2) $[a, b]$ es compacto.

1. Sucesiones en \mathbb{R}^n

Extendemos el concepto de sucesión real a sucesión en \mathbb{R}^n . De hecho, esta extensión ya se hizo para el caso $n = 2$ al trabajar con sucesiones en los complejos, y aquí el asunto es similar. Repasaremos brevemente los resultados agregando algún resultado topológico.

Definición 10. Una *sucesión* en \mathbb{R}^n es cualquier función de \mathbb{N} en \mathbb{R}^n .

p196 de [5]

1.1.1. Notación:

Si el nombre de la función es x , es decir $x : \mathbb{N} \rightarrow \mathbb{R}^n$, entonces para cada natural k , $x(k) \in \mathbb{R}^n$, es decir $x(k)$ será un vector $x(k) = (x(k)_1, x(k)_2, \dots, x(k)_n)$. Para alivianar la escritura, escribiremos x_k en lugar de $x(k)$ y $x_{k,i}$, en lugar de $x(k)_i$. Dependiendo del contexto la notación x_k puede significar tanto la sucesión (es decir toda la función) como el término k -ésimo. Para desambiguar esto último, si queremos referirnos a la sucesión escribiremos (x_k) o $(x_k)_{k \in \mathbb{N}}$.

Observemos que una sucesión en \mathbb{R}^n equivale a n sucesiones reales. Por ejemplo, $((1/k, 0, (-1)^k))_{k \in \mathbb{N}}$ es una sucesión en \mathbb{R}^3 formada por las tres sucesiones reales: $(1/k)_{k \in \mathbb{N}}$, $(0)_{k \in \mathbb{N}}$ y $((-1)^k)_{k \in \mathbb{N}}$.

Definición 11. Diremos que una sucesión $(x_k)_{k \in \mathbb{N}}$ *converge* a un $p \in \mathbb{R}^n$ y escribiremos $x_k \rightarrow p$ o también,

$$\lim_{k \rightarrow +\infty} x_k = p,$$

si

$$d(x_k, p) \xrightarrow[k \rightarrow +\infty]{} 0.$$

O sea, si $\|x_k - p\| \rightarrow 0$ cuando $k \rightarrow +\infty$.

1.1.1. Observación: 1) Como la única variable aquí es k , es redundante escribir “cuando $k \rightarrow +\infty$ ”.

2) Es fácil ver que $x_k \rightarrow p$ sii $\forall B_p, \exists k_0 / \forall k \geq k_0, x_k \in B_p$.

sii $\forall \varepsilon > 0, \exists k_0 / \forall k \geq k_0, d(x_k, p) < \varepsilon$.

De esta forma se extiende en forma natural la noción de límite de sucesiones en \mathbb{R} . Obsérvese que de hecho, la definición en \mathbb{R}^n se basa en la definición en \mathbb{R} . Al igual que en \mathbb{R} , diremos que la sucesión es **convergente** cuando converge a algún p al cual llamaremos **límite** de la sucesión.

En cuanto al cálculo de límites, este se reduce al caso de \mathbb{R} por la siguiente:

Proposición 6.2. Si $p = (p_1, \dots, p_n)$ y $x_k = (x_{k1}, \dots, x_{kn})$ entonces

$$x_k \rightarrow p \iff x_{ki} \rightarrow p_i, \forall i = 1, \dots, n.$$

Demostración. (\Rightarrow) Sea i entre 1 y n cualquiera. Tenemos

$$0 \leq |x_{ki} - p_i| = \sqrt{|x_{ki} - p_i|^2} \leq \sqrt{|x_{k1} - p_1|^2 + \dots + |x_{kn} - p_n|^2} =$$

$$\|x_k - p\| = d(x_k, p) \rightarrow 0$$

por lo tanto $|x_{ki} - p_i| \rightarrow 0$ por lo tanto $x_{ki} \rightarrow p_i$.

(\Leftarrow)

$$d(x_k, p) = \|x_k - p\| = \sqrt{|x_{k1} - p_1|^2 + \dots + |x_{kn} - p_n|^2} \rightarrow 0,$$

ya que es cada término de la suma dentro de la raíz tiende a 0 y la función raíz es continua en 0 y vale 0. □

1.1.3. Ejemplo

La sucesión $(1/k, 0, (-1)^k)$ no tiene límite en \mathbb{R}^3 porque la 3a. coordenada no lo tiene.

1.1.3. Propiedades:: Sean $a, b, (x_k), (y_k) \in \mathbb{R}^n, \lambda, (\lambda_k) \in \mathbb{R}$, tales que $x_k \rightarrow a, y_k \rightarrow b, \lambda_k \rightarrow \lambda$, entonces

1) $x_k + y_k \rightarrow a + b$; 2) $\lambda_k x_k \rightarrow \lambda a$; 3) $\|x_k\| \rightarrow \|a\|$.

Estas propiedades son corolarios inmediatos de la Proposición 6.2.

Proposición 6.3. Sea x_k una sucesión convergente incluida en un conjunto cerrado A . Entonces el límite de x_k pertenece a A .

Demostración. Sea p el límite de x_k . Supongamos por absurdo que $p \notin A$, entonces sea $B(p; r) \subset A^c$. Como $x_k \rightarrow p$ existe k_0 tal que si $k \geq k_0, x_k \in B(p; r)$, por lo tanto $x_{k_0} \in B(p; r)$, por lo tanto $x_{k_0} \notin A$ contradiciendo una de las hipótesis. \square

Esta proposición nos dice que los conjuntos cerrados son “cerrados” bajo la “operación de tomar límites”. O sea, que no podemos “salir” de ellos tomando límites.

Definición 12 (sucesión parcial). Una *sucesión parcial* o *subsucesión* de una sucesión (x_k) es cualquier otra sucesión (y_n) , con $y_n = x_{k_n}$ para cierta sucesión (k_n) estrictamente creciente de naturales.

El concepto es el mismo que para sucesiones en \mathbb{R} : se forma una nueva sucesión eligiendo algunos (infinitos) elementos de (x_k) . Se plantea nuevamente aquí la relación entre sucesiones convergentes (con límite) y sucesiones acotadas. Los resultados son análogos a los de \mathbb{R} : toda sucesión convergente es acotada; toda sucesión acotada tiene una parcial convergente. La demostración del primero es sencilla y totalmente análoga a la de \mathbb{R} .

Veremos a continuación la demostración del segundo resultado, pues es más delicada aunque puede reducirse al caso real. Particularizando en el caso

$n = 2$, esto incluye a las sucesiones complejas.

Teorema 6.4. *Toda sucesión acotada de \mathbb{R}^n tiene una parcial convergente.*

Demostración. Sea (x_k) una sucesión acotada den \mathbb{R}^n . Como ella es acotada sus "sucesiones coordenadas" $(x_{k1}), (x_{k2}), \dots, (x_{kn})$, también lo serán, ya que para cada i se cumple $\|x_{ki}\| \leq \|x_k\|$.

Las n sucesiones coordenadas son sucesiones acotadas en \mathbb{R} . Para cada una vale el teorema correspondiente en \mathbb{R} , por lo que cada una tiene una parcial convergente. Uno puede sentirse tentado a pensar que con esas parciales tenemos una parcial convergente de la original (x_k) . Pero podría ocurrir que esa parciales de cada coordenada, elegidas en forma independiente, no tuvieran índices comunes. Por ejemplo podría suceder que la subsucesión convergente de la primer coordenada, i.e. (x_{k1}) fuera la de los términos pares, i.e. $(x_{(2k)1})$, mientras que la de la segunda coordenada la de los términos impares $(x_{(2k+1)2})$. Pero ni los subíndices pares y los impares tiene porqué servirme para construir una subsucesión convergente de la original.

Se puede, sin embargo, modificar el argumento de la siguiente forma: tomamos una subsucesión convergente de la primer coordenada. Sea ésta $(x_{k_n 1})$ siendo k_n una sucesión creciente de naturales. Ahora considero la subsucesión $(x_{k_n 2})$ de la segunda coordenada. Obviamente ésta subsucesión es acotada, por lo que tendrá una subsucesión convergente $(x_{k_{n_m} 2})$ siendo n_m otra sucesión creciente de naturales. Ahora bien, la sucesión $(x_{k_{n_m} 1})$ es subsucesión de $(x_{k_n 1})$, la cual es convergente, por lo tanto también convergerá. En resumen, la sucesión

$$(x_{k_{n_m} 1}, x_{k_{n_m} 2})$$

será una sucesión convergente de \mathbb{R}^2 . Por inducción podemos seguir el proceso y construir una sucesión convergente de \mathbb{R}^n que sea subsucesión de la original.

Corolario 6.5. *Si $(x_k) \subset K$ compacto, entonces existe $p \in K$ y una subsucesión*

(x_{k_n}) tal que $x_{k_n} \rightarrow p$.

Demostración. K compacto $\Rightarrow (x_k)$ acotada \Rightarrow existe (x_{k_n}) que converge a un cierto $p \in \mathbb{R}^n$. Como K es cerrado, $p \in K$. \square

\square

En \mathbb{R} se vio una versión equivalente al Teorema 6.4, que aquí también es válida:

Corolario 6.6 (Bolzano-Weierstrass). *Todo conjunto infinito y acotado de \mathbb{R}^n tiene un punto de acumulación.*

Demostración. Sea A conjunto infinito. Puedo elegir una sucesión $(x_k) \subset A$ tal que $x_k \neq x_{k'}$ para $k \neq k'$.

Por ser A acotado, existe (x_{k_i}) parcial convergente a un cierto p .

Dado $B(p; r)$, existe i_0 tal que $\forall i \geq i_0, (x_{k_i}) \in B(p; r)$.

Entonces $B(p; r)$ tiene infinitos puntos de A , y por tanto el entorno reducido $B^*(p; r)$ tiene algún punto de A . Como r era arbitrario, p es de acumulación de A . \square

Puede extenderse aquí la definición de sucesión de Cauchy, en forma análoga a la de \mathbb{R} .

Definición 13. (x_k) es una *sucesión de Cauchy* sii dado $\varepsilon > 0$, existe k_0 tal que para todo $k, k' \geq k_0$, $d(x_k, x_{k'}) < \varepsilon$.

Se deduce en forma sencilla que \mathbb{R}^n , al igual que \mathbb{R} , es **completo**, es decir que toda sucesión de Cauchy converge.

Para ello basta observar que como $|x_{k,i} - x_{k',i}| \leq d(x_k - x_{k'})$ las sucesiones coordenadas son de Cauchy y luego convergen, entonces (x_k) converge por la Proposición 6.2.

6.2. Límites y Continuidad

Consideramos una función $f(x_1, \dots, x_n)$ con dominio $D \subset \mathbb{R}^n$ y que toma valores en \mathbb{R} . Escribiremos $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

Por ejemplo: $f(x, y) = x/(x^2 + y^2)$ es una función de dos variables.

Salvo indicación en contrario, tomaremos como **dominio** al mayor conjunto posible donde tiene sentido la expresión. En el ejemplo, $D = \mathbb{R}^2 \setminus (0, 0)$.

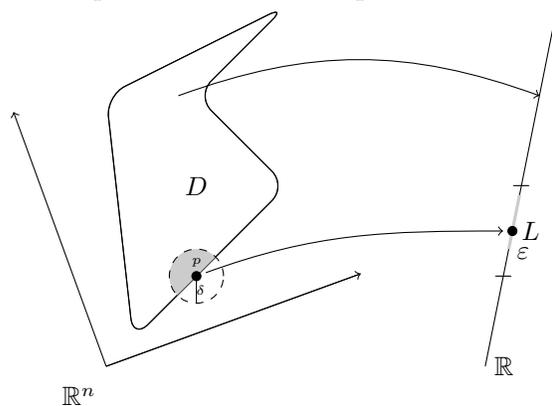
Definición 14 (Límite). $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ función, p punto de acumulación de D , entonces el **límite** es L y escribimos $\lim_{x \rightarrow p} f(x) = L$ (o $f(x) \rightarrow L$ cuando x tiende a p) si

$$\forall \varepsilon > 0, \exists \delta > 0 / \forall x \in B^*(p; \delta) \cap D, f(x) \in B(L; \varepsilon).$$

2.0.4. Observación: La definición es equivalente a las siguientes formulaciones

- $\forall B_L \exists B_p^* / \forall x \in B_p^* \cap D, f(x) \in B_L$;
- $\forall \varepsilon > 0, \exists \delta > 0 / f(B^*(p; \delta) \cap D) \subset B(L; \varepsilon)$
- $\forall B_L \exists B_p^* / f(B_p^* \cap D) \subset B_L$.

Notemos que $B(L; \varepsilon)$ no es más que el intervalo abierto $(L - \varepsilon; L + \varepsilon) \subset \mathbb{R}$.



Definición 15 (Continuidad). $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$ *continua en* $p \in D$ sii

$$\forall B_{f(p)} > 0, \exists B_p > 0 / f(B_p \cap D) \subset B_{f(p)}.$$

(equivalentemente $\forall \varepsilon > 0, \exists \delta > 0 / f(B(p; \delta) \cap D) \subset B(f(p), \varepsilon)$).

Esto equivale (para el caso en que p punto de acumulación de D) a decir

$$\lim_{x \rightarrow p} f(x) = f(p).$$

Se dice que f es *continua* si es continua en p para todo $p \in D$.

Proposición 6.7. Sea p punto de acumulación de D .

(1) $\lim_{x \rightarrow p} f(x) = L \iff \forall (x_k)$ sucesión en $D \setminus \{p\} / x_k \rightarrow p \Rightarrow f(x_k) \rightarrow L$,

(2) f continua en $p \iff p \in D$ y $\forall (x_k)$ sucesión en $D / x_k \rightarrow p \Rightarrow f(x_k) \rightarrow f(p)$.

Demostración. (1) (\Rightarrow) Sea $\varepsilon > 0$. Elegimos $\delta > 0 / f(B^*(p; \delta) \cap D) \subset B(L; \varepsilon)$

Como $x_k \rightarrow p$ y $x_k \neq p, \exists k_0 / \forall k \geq k_0, x_k \in B^*(p; \delta) \cap D$. Por lo tanto, $\forall k \geq k_0, f(x_k) \in B(L; \varepsilon)$, de donde $f(x_k) \rightarrow L$.

(\Leftarrow) Supongamos por absurdo que $f(x)$ no tiene límite L . Entonces $\exists \varepsilon > 0 / \forall \delta, f(B^*(p; \delta) \cap D) \not\subset B(L; \varepsilon)$, por lo tanto existirá al menos un $x_\delta \in B^*(p; \delta) \cap D$ tal que $f(x_\delta) \notin B(L; \varepsilon)$. Si tomemos $\delta = 1/k$ y llamemos x_k a x_δ , entonces tendremos que $x_k \rightarrow p$ con $x_k \neq p$. Por hipótesis, si $p \neq x_k \rightarrow p$ se debe cumplir $f(x_k) \rightarrow L$, pero $f(x_k) \notin B(L; \varepsilon)$ para ningún k . Absurdo.

(2) (\Rightarrow) Análogo a la parte (1). (\Leftarrow) Por hipótesis $\forall (x_k)$ en $D \setminus \{p\}, x_k \rightarrow p \Rightarrow f(x_k) \rightarrow f(p)$, por lo tanto, por (1) $\lim_{x \rightarrow p} f(x) = f(p) \Rightarrow f$ continua en p . \square

Nótese que la parte (2) de este teorema se puede leer así: f es continua en p sii $\forall (x_k) / x_k \rightarrow p$,

$$\lim f(x_k) = f(\lim x_k),$$

o sea que las funciones continua son aquellas que “conmutan con la operación de límite”.

2.0.6. Ejemplos

1) $f(x, y) = x^2 y e^{(x-y)}$. $\lim_{(x,y) \rightarrow (1,1)} f(x, y) = 1$, y f es continua $(1, 1)$.

Una forma sencilla de demostrarlo es por la Proposición 6.7:

Si $(x_k, y_k) \rightarrow (1, 1)$, entonces $x_k \rightarrow 1$ y $y_k \rightarrow 1$, de donde $x_k^2 y_k e^{(x_k - y_k)} \rightarrow 1$ por las propiedades conocidas de sucesiones reales.

2) $f(x, y) = xy/(x^2 + y^2)$. Se verifica análogamente que f es continua en $\mathbb{R}^2 \setminus \{(0, 0)\}$. Queremos estudiar el límite cuando $(x, y) \rightarrow (0, 0)$. Si procedemos como en el ejemplo anterior, obtenemos un límite indeterminado:

$$\lim_{k \rightarrow +\infty} \frac{x_k y_k}{x_k^2 + y_k^2}.$$

con $x_k \rightarrow 0$ y $y_k \rightarrow 0$, que no podemos determinar sin conocer la forma explícita de $(x_k), (y_k)$. Una forma de proceder es la siguiente: Si $x_k = y_k$, entonces

$$\frac{x_k y_k}{x_k^2 + y_k^2} = \frac{x_k^2}{2x_k^2} = \frac{1}{2} \rightarrow \frac{1}{2}.$$

Pero si $x_k = 0, y_k = 1/k$, una cuenta similar nos da límite 0. Al tener límites distintos según la sucesión que tomemos, concluimos que el límite no existe.

Otra forma más general sería restringir f a la recta $y = mx$, entonces

$$f(x, mx) = \frac{m}{1 + m^2},$$

valor constante en la recta y que varía con la pendiente m (si $m = 0$, vale 0 y si $m = 1$, vale $1/2$, por ejemplo). Esto implica que $f(x, y)$ no tiene límite en $(0,0)$: en todo entorno toma los valores entre 0 y $1/2$, ya que la función $h(m) = m/(1 + m^2)$ es continua.

3) $f(x, y) = xy^2/(x^2 + y^4)$. La situación es similar a la del ejemplo 2). Sin

embargo, restringiendo f a rectas $y = mx$,

$$f(x, mx) = \frac{m^2 x^3}{x^2 + m^2 x^4} \xrightarrow{x \rightarrow 0} 0, \quad \text{para todo } m$$

Podríamos pensar en este caso que existe el límite de f y vale 0, ya que por todas las rectas se llega al mismo límite. Sin embargo, si nos restringimos a la parábola $x = y^2$, $f(y^2, y) = 1/2$. Cualquier sucesión (y_k, y_k^2) que tienda a $(0, 0)$, por ejemplo $(1/k, 1/k^2)$, cumplirá $f(y_k, y_k^2) = 1/2 \rightarrow 1/2$. Por la Proposición 6.7, f no puede tener límite en $(0, 0)$.

Comentamos, como condición de los ejemplos 2) y 3), que tomar límites por caminos particulares puede servir para probar que f no tiene límite (si se obtienen valores distintos según el camino), pero nunca para probar que lo tiene.

$$4) f(x, y) = \frac{\sin(x^2 + y^2)}{x^2 + y^2}. \quad \lim_{(x,y) \rightarrow (0,0)} f(x, y) = 1.$$

Efectivamente, si $(x_k, y_k) \rightarrow (0, 0)$, con $(x_k, y_k) \neq (0, 0)$, entonces si $z_k = x_k^2 + y_k^2$ tenemos que $z_k \rightarrow 0$, y $z_k \neq 0$, por lo tanto $(\sin z_k)/z_k \rightarrow 1$ de donde

$$\frac{\sin(x_k^2 + y_k^2)}{x_k^2 + y_k^2} = \frac{\sin z_k}{z_k} \rightarrow 1.$$

No hace falta aquí ninguna restricción de camino para determinar el límite.

2.0.6. Observación: Las dificultades encontradas en el cálculo de límites nos muestran que no es siempre posible reducir el problema “bidimensional” del límite de dos variables a un problema “unidimensional”. Veremos otros ejemplos de esta dificultad más adelante. Si bien estas situaciones aparecen, como hemos visto, en funciones sencillas (cocientes de polinomios), no será en general el interés del curso extenderse demasiado en este tipo de dificultades: nos restringiremos, en general a funciones continuas.

Proposición 6.8. Si f, g son funciones continuas en p , entonces también lo son $f + g, f \cdot g, y$ (si $f(p) \neq 0$) $1/f$.

Proposición 6.9. $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$ continua en p y $g : D'(\subset \mathbb{R}) \rightarrow \mathbb{R}$ continua en $q = f(p)$ con $f(D) \subset D'$, entonces $g \circ f$ (función compuesta) continua en p .

Ambas proposiciones son consecuencia inmediata de Proposición 6.7.

Combinando estas proposiciones puede deducirse la continuidad de muchas funciones. Por ejemplo, $f(x, y) = x^2 + y \exp(x/(x^2 + y^2))$ es continua en $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Definición 16. $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$ es *uniformemente continua* en D sii

$$\forall \varepsilon > 0, \exists \delta > 0 / \forall x, x' \in D, \|x - x'\| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon.$$

La definición es análoga a la vista para funciones de una variable (Capítulo 3 de [5]), y será útil en la parte de integración. En el Capítulo 3 de [5] se vieron los teoremas de Weierstrass (Teorema 6) y Cantor (Teorema 7) sobre funciones continuas en un intervalo cerrado $[a, b]$. Estos teoremas se extienden con idéntica demostración para funciones de varias variables si en lugar de $[a, b]$ ponemos un conjunto K compacto. De hecho, en la demostración se utiliza fuertemente el teorema que dice que una sucesión acotada en $[a, b]$ tiene una parcial con límite en $[a, b]$. En el caso de \mathbb{R}^n , ya hemos probado (Corolario 6.5) que una sucesión en un compacto tiene una parcial con límite en él.

p206 de [5]

Enunciamos entonces los teoremas dejando la demostración al lector:

Teorema 6.10 (Weierstrass). *Una función continua en un compacto tiene máximo y mínimo.*

Teorema 6.11 (Cantor). *Una función continua en un compacto es uniformemente continua.*

6.3. Derivadas Parciales y Direccionales

La intención es generalizar el concepto de derivada a funciones de varias variables trabajaremos en general, por simplicidad, con funciones de dos variables aunque no es difícil de extender lo que se diga para funciones de n variables.

Una ventaja de trabajar en \mathbb{R}^2 es que se dispone de una representación gráfica de la función $z = f(x, y)$ como una superficie en \mathbb{R}^3 . Para funciones de una variable se define la derivada $f'(x)$ como el

$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x},$$

si éste existe. Es decir, se divide el incremento de la función por el incremento de la variable, tomando límite con este último tendiendo a 0. La existencia de derivada se interpreta gráficamente con la existencia de la tangente, y se asocia intuitivamente con la “suavidad” de la función. Por otra parte, el signo de la derivada da idea del crecimiento de la función.

Esta segunda propiedad claramente no puede extenderse en varias variables (no tiene sentido hablar de crecimiento). Buscamos entonces una extensión del primer concepto.

Si se quiere “calcar” la definición, aparece una dificultad: no se puede dividir por el incremento de la variable ¹, que en este caso es un vector $(\Delta x, \Delta y)$.

Una primera solución es restringir estos incrementos, por ejemplo, a una variable. Esto conduce a la siguiente:

Definición 17. Sea (x, y) punto interior a D , $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$. Llamamos *derivada*

¹En realidad si vemos el vector como un complejo, sí se puede hacer la división. Ese encare da lugar al importante estudio de las *funciones holomorfas*, empero no es generalizable a dimensiones más altas.

parcial de f respecto a x en el punto (x, y) al

$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x},$$

si existe, y lo denotamos $f_x(x, y)$ o $\frac{\partial f}{\partial x}(x, y)$. Análogamente,

$$f_y(x, y) = \frac{\partial f}{\partial y}(x, y) = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y},$$

si existe.

De hecho, si fijamos $y = y_0$ y escribimos $g(x) = f(x, y_0)$, función de x , se tiene que:

$$f_x(x, y_0) = g'(x).$$

Esto nos permite calcular derivadas parciales siguiendo las formulas usuales de derivación, considerando a la otra variable como un parámetro.

p208 de [5]

3.0.7. Ejemplo: $f(x, y) = \exp(x^2y)$. Entonces $f_x(x, y) = \exp(x^2y)2xy$ y $f_y(x, y) = \exp(x^2y)x^2$.

3.0.8. Interpretación geométrica: La derivada parcial $f_x(x_0, y_0)$ es la pendiente de la curva intersección del plano $y = y_0$ con la superficie $z = f(x, y)$. Se podría pensar que la existencia de derivadas parciales nos garantiza cierta “suavidad” o regularidad de las funciones. Sin embargo, un ejemplo sencillo nos muestra lo contrario: Si $f(x, y) = 1$ si $x = 0$ o $y = 0$ y 0 en otro caso, entonces $f_x(0, 0) = f_y(0, 0) = 0$, pero la función es discontinua.

Da la impresión que la definición de derivada parcial prioriza indebidamente las direcciones de los ejes. Esto nos lleva a la siguiente:

Definición 18. Sea $\vec{V} \neq 0$ vector de \mathbb{R}^2 , $\vec{V} = (V_1, V_2)$, $(x, y) \in \overset{\circ}{D}$. Llamamos *derivada direccional* de f respecto a \vec{V} en el punto (x, y) a

$$\frac{\partial f}{\partial \vec{V}}(x, y) = \lim_{h \rightarrow 0} \frac{f(x + hV_1, y + hV_2) - f(x, y)}{h},$$

si existe.

Esto equivale a restringir f a la recta por $p = (x, y)$ con dirección \vec{V} , definir $\varphi(h) = f(p + h\vec{V}) = f(x + hV_1, y + hV_2)$, y calcular $\varphi'(0)$.

El método anterior con la función $\varphi(h)$ da una forma de calcular $\partial f / \partial \vec{V}$, pero veremos otro más sencillo más adelante, para los casos usuales.

La existencia de derivadas direccionales en todas las direcciones ¿es garantía de regularidad? Una vez más, la respuesta es negativa. Basta tomar $f(x, y) = 1$ si $y = x^2$ y 0 en otro caso. Para dicha función, en el origen $((x, y) = (0, 0))$, para cualquier vector \vec{V} la función $\varphi(h) = f(\vec{0} + h\vec{V})$ es idénticamente nula en algún entorno suficientemente chico de $h = 0$. Por lo tanto

$$\frac{\partial f}{\partial \vec{V}}(0, 0) = 0,$$

pero f es discontinua en $(0, 0)$.

Una vez más, como en el caso de los límites, es imposible reducir el problema a uno unidimensional. Restringirnos a rectas nunca nos dará garantías sobre el comportamiento global.

Para definir un concepto equivalente al de función derivable de una variable, hace falta un enfoque diferente, que desarrollamos en la sección siguiente, que sin embargo, está íntimamente relacionado con el de derivadas parciales.

Para terminar, veamos los conceptos anteriores para funciones de más de dos variables del estilo $f(x_1, \dots, x_n)$:

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

$$\frac{\partial f}{\partial \vec{V}}(x_1, \dots, x_n) = \lim_{h \rightarrow 0} \frac{f(x_1 + hV_1, x_2 + hV_2, \dots, x_n + hV_n) - f(x_1, \dots, x_n)}{h},$$

En ambos casos la igualdad tiene sentido sólo si el límite existe.

3.0.9. Ejemplo: Si $f(x, y, z) = x^2ze^{yz}$, entonces

$$\begin{aligned}f_x &= 2xze^{yz}, \\f_y &= x^2z^2e^{yz}, \\f_z &= x^2(1 + zy)e^{yz}.\end{aligned}$$

6.4. Diferenciabilidad

Volvemos a la definición de derivada en una variable,

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x},$$

que puede reescribirse de la siguiente forma (ver capítulo 3 de [5])

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + r(\Delta x), \quad \text{donde } \frac{r(\Delta x)}{\Delta x} \xrightarrow{\Delta x \rightarrow 0} 0.$$

Generalizamos esta nueva versión para $n = 2$.

Definición 19. Sea $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$, $p = (x_0, y_0)$ interior a D . Decimos que f es *diferenciable* en p si existen $A, B \in \mathbb{R}$ tales que :

$$\lim_{(\Delta x, \Delta y) \rightarrow 0} \frac{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - A\Delta x - B\Delta y}{\|(\Delta x, \Delta y)\|} = 0,$$

4.0.10. Observación: Al numerador en el límite anterior se le llama **resto de orden 1** de f en (x_0, y_0) , y lo denotaremos $r(\Delta x, \Delta y)$, es decir

$$r(\Delta x, \Delta y) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) - A\Delta x - B\Delta y.$$

De la definición deducimos que

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = A\Delta x + B\Delta y + r(\Delta x, \Delta y), \quad \text{con } \lim_{(\Delta x, \Delta y) \rightarrow 0} \frac{r(\Delta x, \Delta y)}{\|(\Delta x, \Delta y)\|} = 0.$$

La definición dice, que el incremento de f puede aproximarse por una función lineal de los incrementos de las variables tal que el error cometido $r(\Delta x, \Delta y)$ tiende a 0 con mayor orden que $\|(\Delta x, \Delta y)\|$.

Esto admite una interpretación geométrica: Si $x = x_0 + \Delta x$, $y = y_0 + \Delta y$, la fórmula anterior es equivalente a :

$$f(x, y) = f(x_0, y_0) + A(x - x_0) + B(y - y_0) + r(\Delta x, \Delta y).$$

La ecuación $z = f(x_0, y_0) + A(x - x_0) + B(y - y_0)$ representa un plano, que aproxima a la superficie $z = f(x, y)$ con error $r(\Delta x, \Delta y)$.

Puede decirse que la superficie es "aproximable por un plano", que será el plano tangente en el punto $(x_0, y_0, f(x_0, y_0))$. Esto es análogo al caso de una variable, en que la función se aproxima por la recta tangente, y nos da idea que la definición dada es una adecuada generalización.

Es claro además, que para funciones de una variable la derivabilidad y la diferenciabilidad son conceptos equivalentes.

La versión geométrica descrita nos da más garantías de regularidad. Intuitivamente, algo, que está cerca de un plano no puede ser muy irregular.

Veremos en el siguiente teorema que ahora sí tenemos garantizada la continuidad, además de la existencia de derivadas parciales y direccionales. Además veremos que las constantes A y B son justamente las derivadas parciales.

Teorema 6.12. *Sea f diferenciable en $p = (x_0, y_0)$ con*

$$\lim_{(\Delta x, \Delta y) \rightarrow 0} \frac{f(x, y) - f(p) + A(x - x_0) + B(y - y_0)}{\|(x - x_0, y - y_0)\|} = 0.$$

Entonces

1. f continua en $p = (x_0, y_0)$,

2. f tiene derivadas parciales en p , y valen

$$\frac{\partial f}{\partial x}(p) = A, \quad \frac{\partial f}{\partial y}(p) = B.$$

3. Si $\vec{V} = (V_1, V_2)$, f tiene derivada direccional en p , y vale

$$\frac{\partial f}{\partial \vec{V}}(p) = \frac{\partial f}{\partial x}(p)V_1 + \frac{\partial f}{\partial y}(p)V_2.$$

Demostración. Por hipótesis, si $\Delta x = x - x_0$ y $\Delta y = y - y_0$, entonces

$$f(x, y) = f(p) + A\Delta x + B\Delta y + r(\Delta x, \Delta y), \quad \text{con} \quad \lim_{(\Delta x, \Delta y) \rightarrow 0} \frac{r(\Delta x, \Delta y)}{\|(\Delta x, \Delta y)\|} = 0.$$

De donde $\lim_{(\Delta x, \Delta y) \rightarrow 0} r(\Delta x, \Delta y) = 0$. Por lo tanto

1) $\lim_{(\Delta x, \Delta y) \rightarrow 0} f(x, y) = f(p)$, pero $\lim_{(\Delta x, \Delta y) \rightarrow 0} f(x, y) = \lim_{(x, y) \rightarrow p} f(x, y)$, de donde entonces f es continua en p .

2) Tomando $\Delta y = 0$, se tiene

$$\frac{f(x_0 + \Delta x, y_0) - f(p)}{\Delta x} = A + \frac{r(\Delta x, 0)}{\Delta x} \quad (*)$$

Como

$$\left| \frac{r(\Delta x, 0)}{\Delta x} \right| = \left| \frac{r(\Delta x, 0)}{\|(\Delta x, 0)\|} \right| \xrightarrow{\Delta x \rightarrow 0} 0.$$

Por lo tanto el límite cuando $\Delta x \rightarrow 0$ del cociente de la izquierda en (*) existe y es igual a A . Como dicho límite cuando existe es igual a $f_x(p)$, entonces $f_x(p) = A$.

Análogamente para $f_y(p)$.

3) Tomando $(\Delta x, \Delta y) = h\vec{V} = (hV_1, hV_2)$, tenemos

$$\frac{f(p + h\vec{V}) - f(p)}{h} = AV_1 + BV_2 + \frac{r(hV_1, hV_2)}{h}$$

como

$$\left| \frac{r(hV_1, hV_2)}{h} \right| = \|\vec{V}\| \cdot \left| \frac{r(hV_1, hV_2)}{\|(hV_1, hV_2)\|} \right| \xrightarrow{h \rightarrow 0} 0,$$

tenemos $(\partial f / \partial \vec{V})(p) = AV_1 + BV_2$. \square

El teorema anterior nos dió, para una función diferenciable, una expresión para la derivada direccional:

$$\frac{\partial f}{\partial \vec{V}}(p) = f_x(p) \cdot V_1 + f_y(p) \cdot V_2 = (f_x(p), f_y(p)) \cdot \vec{V}.$$

Definición 20. Al vector $(f_x(p), f_y(p))$ se le llama *vector gradiente* de la función f y se lo denota $\nabla f(p)$.

Tenemos entonces que si f es diferenciable, entonces

$$\frac{\partial f}{\partial \vec{V}}(p) = \nabla f(p) \cdot \vec{V}$$

y

$$f(p + \vec{V}) - f(p) = \nabla f(p) \cdot \vec{V} + r(\vec{V}), \quad \text{con } \frac{r(\vec{V})}{\|\vec{V}\|} \xrightarrow{\vec{V} \rightarrow 0} 0$$

En la expresión anterior, si nos restringimos a vectores \vec{V} de norma 1 ($\|\vec{V}\| = 1$), vemos que el máximo de $\frac{\partial f}{\partial \vec{V}}(p)$ se obtiene en la dirección de $\nabla f(p)$: el gradiente nos da la dirección de mayor “crecimiento” de f .

La anterior definición se extiende para f de n variables, $f(x_1, \dots, x_n)$. En este caso

$$\nabla f(p) = (f_{x_1}(p), \dots, f_{x_n}(p)).$$

1. Condición Suficiente de Diferenciabilidad

Los resultados anteriores, y otros que surgen, nos dan buenas propiedades para funciones diferenciables. Ahora bien, ¿cómo saber si una función es diferenciable? Se puede, por supuesto, recurrir a la definición, pero existen criterios

más sencillos. Veremos a continuación una condición suficiente aplicable en los casos usuales. Antes de ello corresponde observar que vale el álgebra usual para funciones diferenciales, es decir, la suma, producto y división de funciones diferenciables da lugar a una función diferenciable. Lo dejamos como ejercicio

Ejercicio 6.4.1. Si f y g son funciones diferenciables en p , entonces también lo son $f + g$, fg , y (si $f(p) \neq 0$) $1/f$.

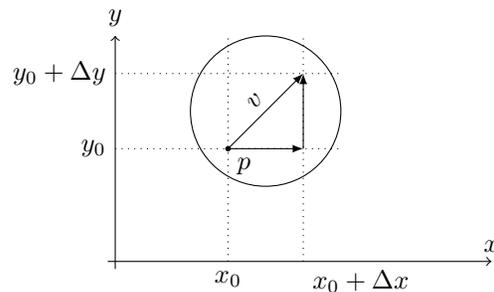
Teorema 6.13 (Condición suficiente de diferenciability). Sea $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$, p interior a D .

Si las derivadas parciales f_x, f_y existen en un entorno de p , y son continuas en p , entonces f es diferenciable en p .

Demostración. Sea $\delta > 0$ tal que tal que f_x y f_y existen en $B(p; \delta)$. Supongamos que $p = (x_0, y_0)$ y sea $v = (\Delta x, \Delta y)$ con $\|v\| < \delta$. Entonces si escribimos

$$r(v) = f(p+v) - f(p) - f_x(p)\Delta x - f_y(p)\Delta y, \quad (*)$$

para probar la diferenciability, debemos ver que $\frac{r(v)}{\|v\|} \xrightarrow{v \rightarrow 0} 0$.



Escribiremos el incremento $f(p+v) - f(p)$ como la suma de dos incrementos, uno vertical y otro horizontal, así:

$$f(p+v) - f(p) = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] + [f(x_0 + \Delta x, y_0) - f(x_0, y_0)] \quad (**)$$

En el primer incremento la variable x permanece constante (igual a $x_0 + \Delta x$),

mientras que en el segundo incremento es la variable y que permanece constante (igual a y_0). Como las derivadas parciales existen, podemos aplicar el teorema de Lagrange a $\varphi(y) = f(x_0 + \Delta x, y)$ en $[y_0, y_0 + \Delta y]$ ya que $\varphi'(y) = f_y(x_0 + \Delta x, y)$. Su aplicación nos dará un $\theta \in (0; 1)$ tal que $\varphi(y_0 + \Delta y) - \varphi(y_0) = \varphi'(y_0 + \theta\Delta y)\Delta y$, de donde

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0) = f_y(x_0 + \Delta x, y_0 + \theta\Delta y)\Delta y. \quad (***)$$

Juntando (*), (**) y (***) obtenemos la siguiente expresión para $r(v)$:

$$\begin{aligned} r(v) &= f_y(x_0 + \Delta x, y_0 + \theta\Delta y)\Delta y + f(x_0 + \Delta x, y_0) - f(p) - f_x(p)\Delta x - f_y(p)\Delta y \\ &= \left[\frac{f(x_0 + \Delta x, y_0) - f(p)}{\Delta x} - f_x(p) \right] \Delta x + [f_y(x_0 + \Delta x, y_0 + \theta\Delta y) - f_y(p)]\Delta y \end{aligned}$$

De donde

$$\frac{r(v)}{\|v\|} = \underbrace{\left[\frac{f(x_0 + \Delta x, y_0) - f(p)}{\Delta x} - f_x(p) \right]}_{(I)} \underbrace{\frac{\Delta x}{\|v\|}}_{(i)} + \underbrace{[f_y(x_0 + \Delta x, y_0 + \theta\Delta y) - f_y(p)]}_{(II)} \underbrace{\frac{\Delta y}{\|v\|}}_{(ii)}$$

Haciendo $v \rightarrow 0$ tenemos que $(I) \rightarrow 0$ por existir $f_x(p)$, $(II) \rightarrow 0$ por ser f_y continua en p , (i) y (ii) acotados entre -1 y 1 , de donde toda la expresión tiende a cero. \square

4.1.0. Ejemplo: La función $f(x, y) = x^2 e^{xy}$ es diferenciable en cualquier punto, ya que sus derivadas parciales, $f_x = (x^2 y + 2x)e^{xy}$ y $f_y = x^3 e^{xy}$ son funciones continuas.

Tomando $p = (1, 0)$, por ejemplo, se obtiene

$$f(x, y) = f(1, 0) + f_x(1, 0)(x-1) + f_y(1, 0)y + r(x-1, y) = 1 + 2(x-1) + y + r(x-1, y).$$

De donde $z = 1 + 2(x-1) + y$ es la ec. del plano tangente a $z = f(x, y)$ en el pto $(1, 0)$.

p215 de [5]

p216 de [5]

2. Diferencial de una Función

Sea f diferenciable en $p = (x_0, y_0)$, Δf el incremento de f , es decir $\Delta f = f(x + \Delta x, y + \Delta y) - f(x_0, y_0)$. Se tiene

$$\Delta f = f_x(p)\Delta x + f_y(p)\Delta y + r(\Delta x, \Delta y), \quad \text{con } \frac{r(\Delta x, \Delta y)}{\|(\Delta x, \Delta y)\|} \xrightarrow{(\Delta x, \Delta y) \rightarrow (0,0)} 0.$$

El incremento Δf de la función depende de los incrementos Δx , Δy y puede ser aproximado cerca de p , despreciando el resto, por la función lineal df

$$df(\Delta x, \Delta y) = f_x(p)\Delta x + f_y(p)\Delta y.$$

Definición 21. $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ diferenciable en p . Llamamos *diferencial* de f en el punto p , a la función lineal de las variables Δx , Δy ,

$$df|_p(\Delta x, \Delta y) = f_x(p)\Delta x + f_y(p)\Delta y.$$

p217 de [5]

4.2.0. Observación: Es usual denotar dx , dy a los incrementos Δx , Δy , y obtener entonces

$$df|_p(dx, dy) = f_x(p)dx + f_y(p)dy.$$

Conviene señalar, sin embargo, que en la expresión anterior dx , dy son exactamente los incrementos Δx , Δy , mientras que df es sólo una aproximación al incremento Δf , tanto más ajustada cuanto más pequeños sean Δx , Δy . En las aplicaciones muchas veces se identifica Δf con df diciendo que se trata de incremento “infinitamente pequeños”. Este procedimiento no tiene sentido estricto en este enfoque, pero resulta poderoso del punto de vista intuitivo y los resultados en general se pueden formalizar por paso al límite.

3. Regla de la Cadena

Se llama habitualmente regla de la cadena a la regla de diferenciación de funciones compuestas. Para funciones de un variable, la regla es la siguiente si $g(y)$, $f(x)$ son derivables la función compuesta $g(f(x))$ es derivable y $(g(f(x)))' = g'(f(x)) \cdot f'(x)$.

Veremos como extender esta regla a funciones de varias variables.

Consideramos en primera instancia la función compuesta $g(t) = f(x(t), y(t))$ compuesta de $f(x, y)$ con la función vectorial de una variable² $(x(t), y(t))$, que supondremos derivable. Nos preguntamos que hipótesis pedirle a $f(x, y)$ para asegurar que $g(t)$ sea derivable. Las consideraciones anteriores hacen pensar en la diferenciabilidad. Efectivamente:

Teorema 6.14 (Regla de la cadena). Sean $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ diferenciable en $p = (x_0, y_0) \in \overset{\circ}{D}$,
 $\alpha : I(\subset \mathbb{R}) \rightarrow D$, $\alpha(t) = (x(t), y(t))$, derivable en t_0 , con $\alpha(t_0) = p$ y $\alpha(I) \subset D$.
 entonces $g(t) = (f \circ \alpha)(t) = f(x(t), y(t))$ es derivable en t_0 , con

$$g'(t_0) = f_x(p) \cdot x'(t_0) + f_y(p) \cdot y'(t_0) = \nabla f(p) \cdot \alpha'(t_0).$$

(notar que g es una función de \mathbb{R} en \mathbb{R})

Demostración. Para demostrar que g es derivable en t_0 consideramos su cociente incremental:

$$\frac{g(t_0 + h) - g(t_0)}{h} = \frac{f(\alpha(t_0 + h)) - \overbrace{f(\alpha(t_0))}^p}{h} = \frac{f(p + \Delta\alpha) - f(p)}{h},$$

donde $\Delta\alpha = \alpha(t_0 + h) - \alpha(t_0)$. Como f es diferenciable en p entonces

$$f(p + \Delta p) = f(p) + \nabla f(p) \cdot \Delta p + r(\Delta p), \quad \text{con } \frac{r(\Delta p)}{\|\Delta p\|} \xrightarrow{\Delta p \rightarrow 0} 0$$

²también llamada *curva*.

Por lo tanto

$$\frac{f(p + \Delta\alpha) - f(p)}{h} = \nabla f(p) \cdot \frac{\Delta\alpha}{h} + \frac{r(\Delta\alpha)}{h}.$$

p219 de [5]

Como $\Delta\alpha/h \rightarrow \alpha'(p)$ cuando $h \rightarrow 0$, basta demostrar que $r(\Delta\alpha)/h \rightarrow 0$. Para ello vemos que, por un lado, si $\Delta\alpha = 0$, entonces $r(\Delta\alpha) = 0$. Por otro lado si $\Delta\alpha \neq 0$, entonces $\|\Delta\alpha\| \neq 0$ y

$$\frac{r(\Delta\alpha)}{h} = \underbrace{\frac{r(\Delta\alpha)}{\|\Delta\alpha\|}}_{(I)} \cdot \underbrace{\frac{\|\Delta\alpha\|}{h}}_{(II)}.$$

Cuando $h \rightarrow 0$, $\Delta\alpha \rightarrow 0$, por ser α continua en t_0 , por lo tanto el cociente (I) tiende a cero cuando $h \rightarrow 0$. El cociente (II) está acotado, pues

$$\left\| \frac{\|\Delta\alpha\|}{h} \right\| = \left\| \frac{\Delta\alpha}{h} \right\| \xrightarrow{h \rightarrow 0} \|\alpha'(t_0)\|.$$

□

Una generalización de lo anterior consiste en componer $f(x, y)$ con una función vectorial $(x(u, v), y(u, v))$ de dos variables, obteniendo $g(u, v) = f(x(u, v), y(u, v))$.

Teorema 6.15. Sea $f(x, y)$ diferenciable en $p = (x_0, y_0)$,

$x(u, v)$ tiene derivadas parciales en (u_0, v_0) y $x(u_0, v_0) = x_0$,

$y(u, v)$ tiene derivadas parciales en (u_0, v_0) y $y(u_0, v_0) = y_0$,

entonces, $g(u, v) = f(x(u, v), y(u, v))$ tiene derivadas parciales en (u_0, v_0) y valen:

$$(1) \quad \frac{\partial g}{\partial u} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial u}$$

$$(2) \quad \frac{\partial g}{\partial v} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial v},$$

donde se sobreentienden los puntos de evaluación de las derivadas.

p220 de [5]

Demostración. Sea $v = v_0$ fijo. Entonces $g(u, v_0) = f(x(u, v_0), y(u, v_0))$ es una función compuesta como en el Teorema 6.14, y se cumplen todas las hipóte-

sis. (1) es entonces una consecuencia inmediata del Teorema 6.14, y se procede análogamente para (2). \square

4.3.0. Observaciones:

- a) Si se pide además que las funciones $x(u, v)$, $y(u, v)$ sean diferenciables, puede demostrarse que g lo es rehaciendo una demostración como la del Teorema 6.14.
- b) Sin embargo, será suficiente en general que $f(x, y)$, $x(u, v)$, $y(u, v)$ tengan derivadas parciales continuas. En ese caso, $g(u, v)$ también las tiene, como surge inmediatamente del Teorema 6.15 y por lo tanto es diferenciable.

4.3.1. Ejemplo: Se tiene la función $f(x, y)$ definida en el plano, se la quiere representar usando coordenadas polares. Es decir, obtener $g(\rho, \theta) = f(\rho \cos \theta, \rho \sin \theta)$.

Usando el Teorema 6.15, las derivadas parciales de g son

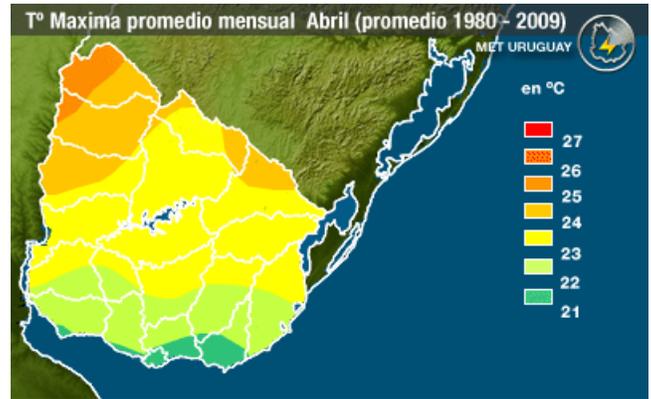
$$\begin{aligned} (1) \frac{\partial g}{\partial \rho} &= \frac{\partial f}{\partial x} \cdot \cos \theta + \frac{\partial f}{\partial y} \cdot \sin \theta, \\ (2) \frac{\partial g}{\partial \theta} &= -\rho \frac{\partial f}{\partial x} \cdot \sin \theta + \rho \frac{\partial f}{\partial y} \cdot \cos \theta. \end{aligned}$$

4. Curvas y Superficies de Nivel

Una forma de representar en el plano una función de dos variables es dando las **curvas de nivel**, es decir los conjuntos $f^{-1}(k) = \{(x, y) / f(x, y) = k\}$ donde la función toma el nivel k .

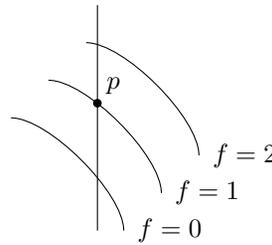
Pensando en la representación espacial, esto equivale a cortar la superficie $z = f(x, y)$ con el plano $z = k$.

Un ejemplo típico son los **mapas de relieve**, en que se da la altura en función de dos coordenadas geográficas.



Otro ejemplo son las **isotermas** (curvas de temperatura constante) en un mapa meteorológico.

Veamos una propiedad de las curvas de nivel: observando la figura, si estamos en p , parece claro que la función crece más rápido si nos movemos en forma perpendicular a las curvas de nivel. Ya sabemos que si $\nabla f \neq 0$, ∇f indica la dirección de máximo crecimiento de f . Esto motiva la siguiente:

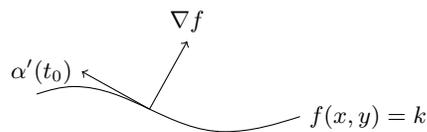


Proposición 6.16. Sea $p = (x_0, y_0)$, f diferenciable, $\nabla f(p) \neq 0$, $\alpha(t)$ curva diferenciable tal que $\forall t \in I$, $\alpha(t) \in \{(x, y) / f(x, y) = k\}$, entonces $\alpha'(t_0) \perp \nabla f(p)$.

p222 de [5]

Demostración. Sea $g(t) = f(\alpha(t))$, $g(t) = k \forall t \in I$ por lo tanto $g'(t_0) = 0$.

Por la regla de la cadena, $g'(t) = \nabla f(p) \cdot \alpha'(t_0)$. □



4.4.0. Observación: a) Lo anterior supuso que la curva de nivel pudo parametrizarse en la forma $\alpha(t), t \in I$. Esto aún no se ha visto, pero más adelante veremos el Teorema 6.22 de la función implícita, que mostrará que si f tiene derivadas parciales continuas y $\nabla f(p) \neq 0$, esto es posible.

b) Los resultados anteriores se pueden extender para funciones de tres variables, introduciendo las superficies de nivel $\{(x, y, z) \in \mathbb{R}^3 / f(x, y, z) = k\}$.

Se demuestra en forma similar que ∇f es perpendicular a las superficies de nivel.

4.4.1. Ejemplo: Veamos un ejemplo de b): si $f(x, y, z) = x^2 + y^2 + z^2$, los conjuntos $\{(x, y, z) / x^2 + y^2 + z^2 = k\}$, con $k > 0$, representan esferas de centro en el origen. Si $P = (x_0, y_0, z_0)$, $\nabla f(P) = (2x_0, 2y_0, 2z_0) = 2(P - O)$ proporcional al radio vector y por lo tanto perpendicular a la esfera.

6.5. Derivadas de Mayor Orden

Supongamos que $f(x, y)$ tiene derivadas parciales en todos los puntos interiores a D . Tenemos definidas, entonces, dos funciones de dos variables, $f_x(x, y)$ y $f_y(x, y)$. Si éstas a su vez tienen derivadas parciales, obtenemos las derivadas segundas:

$$\begin{aligned} f_{xx} &= \frac{\partial f_x}{\partial x} = \frac{\partial^2 f}{\partial x^2}, & f_{xy} &= \frac{\partial f_x}{\partial y} = \frac{\partial^2 f}{\partial y \partial x}, \\ f_{yx} &= \frac{\partial f_y}{\partial x} = \frac{\partial^2 f}{\partial x \partial y}, & f_{yy} &= \frac{\partial f_y}{\partial y} = \frac{\partial^2 f}{\partial y^2}. \end{aligned}$$

Análogamente, pueden definirse derivadas terceras f_{xxx}, f_{xyx} , etc., si las funciones anteriores tienen derivadas parciales.

5.0.3. Ejemplo

$f(x, y) = x^2 e^{xy}$. Derivando se obtiene

$$f_x = (2x + x^2 y)e^{xy}, f_y = x^3 e^{xy}, \quad (6.1)$$

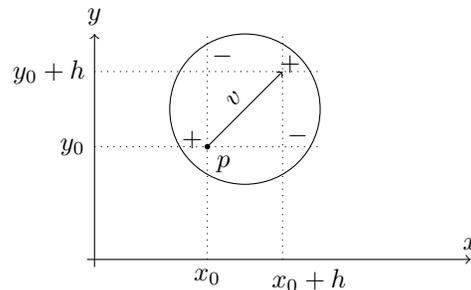
$$f_{xx} = (x^2 y^2 + 4xy + 2)e^{xy}, f_{yy} = x^4 e^{xy}, \quad (6.2)$$

$$f_{xy} = f_{yx} = (3x^2 + x^3 y)e^{xy}. \quad (6.3)$$

Se observa en el ejemplo que las derivadas “cruzadas” f_{xy} , f_{yx} , son iguales. Para funciones con suficiente regularidad este hecho es general.

Teorema 6.17. Si f_{xy} , f_{yx} existen en un B_p , y son continuas en p , entonces $f_{xy}(p) = f_{yx}(p)$.

Demostración. Sea $p = (x_0, y_0)$ y $v = (h, h)$ tal que $a + v \in B_p$.



Calculamos el siguiente “incremento doble” de la función f en los vértices del cuadrado de lado h y vértice inferior izquierdo p :

$$\varphi(h) = f(x_0 + h, y_0 + h) - f(x_0, y_0 + h) - f(x_0 + h, y_0) + f(x_0, y_0).$$

Agrupando la expresión anterior, se la puede escribir en dos formas:

$$(1) \quad \varphi(h) = [f(x_0 + h, y_0 + h) - f(x_0 + h, y_0)] - [f(x_0, y_0 + h) - f(x_0, y_0)],$$

$$(2) \quad \varphi(h) = [f(x_0 + h, y_0 + h) - f(x_0, y_0 + h)] - [f(x_0 + h, y_0) - f(x_0, y_0)].$$

Trabajamos con la forma (1). Llamando $\psi(x)$ al incremento $\psi(x) = f(x, y_0 + h) - f(x, y_0)$, se tiene que

$$\varphi(h) = \psi(x_0 + h) - \psi(x_0) = h\psi'(x_0 + \theta h), \quad \theta \in (0; 1),$$

donde la última igualdad es por Lagrange aplicado a ψ en $[x_0, x_0 + h]$.

$$\Rightarrow \varphi(h) = h [f_x(x_0 + \theta h, y_0 + h) - f_x(x_0 + \theta h, y_0)] \quad \theta \in (0; 1).$$

Aplicando nuevamente Lagrange a la función $\lambda(y) = f_x(x_0 + \theta h, y)$ en $[y_0, y_0 + h]$:

$$\varphi(h) = h^2 \cdot f_{xy}(x_0 + \theta h, y_0 + \theta' h) \quad \theta, \theta' \in (0; 1).$$

De donde, $\frac{\varphi(h)}{h^2} = f_{xy}(x_0 + \theta h, y_0 + \theta' h) \xrightarrow{h \rightarrow 0} f_{xy}(p)$ por ser f_{xy} continua en p . El razonamiento anterior puede repetirse en forma análoga partiendo de la forma (2). En ese caso se llega a $\frac{\varphi(h)}{h^2} \xrightarrow{h \rightarrow 0} f_{yx}(p)$.

Como el límite es único, se deduce $f_{xy}(p) = f_{yx}(p)$.

□

p225 de [5]

5.0.3. Notas: Puede parecer que el teorema anterior no sea muy útil, ya que para verificar las hipótesis hace falta conocer las derivadas segundas para saber si son continuas. Sin embargo, en muchos casos usuales se sabe de antemano que existen derivadas parciales continuas de todos los órdenes, salvo a lo sumo en algún punto. Por ejemplo, para $f(x, y) = x^2 e^{\operatorname{sen} y} / (x^2 + y^2)$ se sabe que $f_{xy}(x, y) = f_{yx}(x, y) \quad \forall (x, y) \neq (0, 0)$.

Existen, por otra parte, versiones, del teorema anterior con hipótesis más débiles: *Basta pedir que, además de f_x, f_y , exista y sea continua f_{xy} para que exista también f_{yx} , y $f_{xy} = f_{yx}$. Ver [1].*

Por último, se pueden dar ejemplos en que no se cumple $f_{xy} = f_{yx}$, aunque

existan las dos. Dos posibles ejemplos son

$$1) f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases} \quad 2) f(x, y) = \begin{cases} xy & \text{si } x^2 < y^2, \\ -xy & \text{si } x^2 > y^2, \\ 0 & \text{si } x^2 = y^2. \end{cases}$$

En ambos casos $f_{xy}(0, 0) = 1$, $f_{yx}(0, 0) = -1$. (Ejercicio)

p226 de [5]

1. Funciones de Clase C^k

Definición 22. $f(x, y)$ es de *clase C^k* si tiene derivadas parciales continuas hasta orden k . Si eso para todo k entonces la función es de *clase C^∞* .

2. Diferenciales de Mayor Orden

Al definir diferenciabilidad vimos la idea de aproximar el incremento Δf de la función por un término lineal (de 1er grado) en Δx , Δy , que llamamos diferencial df .

Nos proponemos ahora generalizar el desarrollo de Taylor que permite, en funciones de una variable con k derivadas, aproximar Δf por un polinomio de grado k .

Introducimos previamente algunas notaciones: Sea f de clase C^k .

Habíamos denotado $df|_p(\Delta x, \Delta y) = f_x(p)\Delta x + f_y(p)\Delta y$

Ahora escribiremos

$$d^2 f|_p(\Delta x, \Delta y) = f_{xx}(p)\Delta x^2 + 2f_{xy}(p)\Delta x\Delta y + f_{yy}(p)\Delta y^2,$$

$$d^3 f|_p(\Delta x, \Delta y) = f_{xxx}(p)\Delta x^3 + 3f_{x^2y}(p)\Delta x^2\Delta y + 3f_{xy^2}(p)\Delta x\Delta y^2 + f_{y^3}(p)\Delta y^3,$$

$$d^k f|_p(\Delta x, \Delta y) = f_{x^k}(p)\Delta x^k + \cdots + C_i^k f_{x^{k-i}y^i}(p)\Delta x^{k-i}\Delta y^i + \cdots + f_{y^k}(p)\Delta y^k,$$

donde C_i^k es el **coeficiente binomial** que es igual a $\frac{k!}{i!(k-i)!}$.

En general $d^k f|_p(\Delta x, \Delta y)$, llamado **diferencial de orden k** de f en p , es un

p227 de [5]

polinomio homogéneo de grado k en los incrementos $\Delta x, \Delta y$, en particular

$$\frac{d^k f|_p(v)}{\|v\|^k} = d^k f|_p(u) \quad \text{con } u = \frac{v}{\|v\|}.$$

Una forma de recordar la fórmula es trabajar formalmente con el binomio de Newton.

$$d^k f(\Delta x, \Delta y) = \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^k f$$

identificando los productos $\frac{\partial^{k-i}}{\partial x^{k-i}} \frac{\partial^i}{\partial y^i}$ con $\frac{\partial^k}{\partial x^{k-i} \partial y^i}$.

3. Desarrollo de Taylor

Recordamos brevemente el desarrollo de una función de una variable en un punto p . Si $f(x)$ tiene k derivadas en p , puede escribirse

$$f(p+\Delta x) = f(p) + f'(p)\Delta x + \frac{f''(p)}{2}\Delta x^2 + \dots + \frac{f^{(k)}(p)}{k!}\Delta x^k + r_k(\Delta x), \quad \lim_{\Delta x \rightarrow 0} \frac{r_k(\Delta x)}{(\Delta x)^k} = 0.$$

Si además f tiene derivada continua de orden $k+1$, se puede escribir

$$r_k(\Delta x) = \frac{f^{(k+1)}(a + \theta\Delta x)(\Delta x)^{k+1}}{(k+1)!}, \quad \text{con } \theta \in (0;1) \quad (\text{resto de Lagrange.})$$

Para funciones de dos variables, hay resultados análogos. En este caso los polinomios son de dos variables $\Delta x, \Delta y$, y el término $d^i f|_p(\Delta x, \Delta y)$ juega el papel del término $f^{(i)}(p)\Delta x$.

5.3.0. Nota: Se tienen los siguientes resultados:

(A) Si f es k -veces diferenciable en p (f y sus derivadas parciales hasta orden $k-1$ son diferenciables en p), entonces, si $\Delta p = (\Delta x, \Delta y)$ se tiene:

$$f(p + \Delta p) = f(p) + df|_p(\Delta p) + \frac{1}{2}d^2 f|_p(\Delta p) + \dots + \frac{1}{k!}d^k f|_p(\Delta p) + r_k(\Delta p),$$

$$\text{con } \lim_{\Delta p \rightarrow 0} \frac{r_k(\Delta p)}{\|\Delta p\|^k} = 0.$$

(B) Si f es de clase C^{k+1} , se puede escribir además

$$r_k(\Delta p) = \frac{1}{(k+1)!} d^{k+1} f|_{p+\theta\Delta p}(\Delta p), \quad 0 < \theta < 1.$$

Nos limitamos aquí a demostrar (B), que puede deducirse fácilmente del caso de una variable, y cubre los casos usuales. Para (A), ver [4].

Teorema 6.18. *Sea f de clase C^{k+1} en $B(p; \delta)$, $v \in B(0; \delta)$. Entonces existe $\theta \in (0; 1)$ tal que*

$$f(p+v) = f(p) + df|_p(v) + \frac{1}{2} d^2 f|_p(v) + \cdots + \frac{1}{k!} d^k f|_p(v) + r_k(v),$$

$$\text{donde } r_k(v) = \frac{1}{(k+1)!} d^{k+1} f|_{p+\theta v}(v).$$

$$\text{De lo anterior se deduce que } \lim_{v \rightarrow 0} \frac{r_k(v)}{\|v\|^k} = 0.$$

Demostración. Sea $g : [0, 1] \rightarrow \mathbb{R}$, $g(t) = f(p+tv) = f(x_0 + t\Delta x, y_0 + t\Delta y)$, donde $v = (\Delta x, \Delta y)$. Entonces $g(0) = f(p)$, $g(1) = f(p+v)$. Derivamos $g(t)$ usando la regla de la cadena, hasta orden $k+1$

$$g'(t) = f_x(p+tv)\Delta x + f_y(p+tv)\Delta y = df|_{p+tv}(v)$$

$$\begin{aligned} g''(t) &= f_{xx}(p+tv)\Delta x^2 + f_{xy}(p+tv)\Delta x\Delta y + f_{yx}(p+tv)\Delta y\Delta x + f_{yy}(p+tv)\Delta y^2 \\ &= d^2 f|_{p+tv}(v). \end{aligned}$$

Procediendo por inducción, no es difícil probar que $g^i(t) = d^i f|_{p+tv}(v)$. Se utilizó en todos los pasos que f es de clase C^{k+1} para la regla de la cadena.

Entonces g es una función de una variable con derivadas continuas hasta orden $k+1$ en $[0, 1]$. La podemos desarrollar por Taylor en $[0, 1]$ y obtener

$$g(1) = g(0) + g'(0) + \frac{1}{2} \frac{g''(0)}{2} + \cdots + \frac{g^{(k)}(0)}{k!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}, \quad 0 < \theta < 1.$$

Sustituyendo con las expresiones obtenidas para las derivadas de g , se obtiene el desarrollo de la tesis. Para terminar el teorema, veamos que $\lim_{v \rightarrow 0} \frac{r_k(v)}{\|v\|^k} = 0$:

$$\frac{r_k(v)}{\|v\|^k} = \|v\| \frac{d^{k+1}f|_{p+\theta v}}{\|v\|^{k+1}}(v) = \|v\| d^{k+1}f|_{p+\theta v}(u) \quad \text{con } u = \frac{v}{\|v\|} \in \partial B(0; 1).$$

Como las derivadas parciales de orden $k + 1$ de f son continuas, el diferencial de orden $k + 1$ de f será una función continua. Además las derivadas parciales de f están evaluadas en puntos del segmento $[p, p + v] = \{p + tv/t \in [0, 1]\}$ que une p y $p + v$, que es un compacto de \mathbb{R}^n (es claramente acotado, y no es difícil ver que es cerrado) y u varía en el compacto $\partial B(0; 1)$. Por lo tanto $d^{k+1}f|_{p+\theta v}(u)$ es una función continua en $[p, p + v] \times \partial B(0; 1)$ compacto (de \mathbb{R}^{n+n}), por ende acotada. Como $\|v\| \rightarrow 0$ entonces $\frac{r_k(v)}{\|v\|^k} \xrightarrow{v \rightarrow 0} 0$ como queríamos. \square

p230 de [5]

5.3.1. Ejemplo: Hallamos el desarrollo de Taylor de orden 2 de $f(x, y) = \log(1 + x + 2y)$ en el punto $p = (2, 1)$.

$$\begin{aligned} f_x(x, y) &= \frac{1}{1 + x + 2y}, & f_y(x, y) &= \frac{2}{1 + x + 2y}, \\ f_{xx}(x, y) &= \frac{-1}{(1 + x + 2y)^2}, & f_{xy} &= f_{yx} = \frac{-2}{(1 + x + 2y)^2}, & f_{yy} &= \frac{-4}{(1 + x + 2y)^2}. \end{aligned}$$

Se obtiene:

$$f(2 + \Delta x, 1 + \Delta y) = \log 5 + \frac{1}{5}\Delta x + \frac{2}{5}\Delta y - \frac{1}{50}(\Delta x^2 + 4\Delta x\Delta y + 4\Delta y^2) + r(\Delta x, \Delta y)$$

$$\text{con } \lim_{(\Delta x, \Delta y) \rightarrow (0, 0)} \frac{r(\Delta x, \Delta y)}{\Delta x^2 + \Delta y^2} = 0.$$

5.3.2. Nota: En funciones de más de dos variables, el desarrollo es válido y tiene la misma forma. En el caso $f(x_1, \dots, x_n)$ debe interpretarse

$$d^k f(\Delta x_1, \dots, \Delta x_n) = \left(\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right)^p f,$$

polinomio homogéneo de grado p que se obtiene desarrollando formalmente

la expresión anterior.

p231 de [5]

6.6. Extremos

Una función de una variable tiene un **máximo relativo** en un punto si la función toma en él un valor mayor que en los demás puntos de un entorno; tiene un **máximo absoluto** en un punto si toma en él un valor mayor que en él.

Los conceptos son análogos en varias variables.

Definición 23. Sea $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$, $p \in D$. Diremos que f tiene en p un

- $\left\{ \begin{array}{l} \text{máximo} \\ \text{mínimo} \end{array} \right\}$ *relativo* si $\exists B_p \subset D / \forall v \in B_p, f(v) \left\{ \begin{array}{l} \leq \\ \geq \end{array} \right\} f(p)$.
- $\left\{ \begin{array}{l} \text{máximo} \\ \text{mínimo} \end{array} \right\}$ *absoluto* si $\forall v \in D, f(v) \left\{ \begin{array}{l} \leq \\ \geq \end{array} \right\} f(p)$.

Notar que con la definición anterior un **extremo absoluto** (máximo o mínimo) puede no ser **extremo relativo**, porque puede no darse en un punto interior³. Por ejemplo, para la función $f : [0, 1] \rightarrow \mathbb{R}$, $f(x) = x^2$, el máximo absoluto ocurre en la frontera de $[0, 1]$. Esta distinción será cómoda para dar condiciones en términos de derivadas. Para la función f anterior, la derivada (lateral) en el extremo absoluto no se anula (vale $f'(1) = 2$), pero en un extremo relativo, si hay derivada, esta se anula. Desarrollamos ahora el caso de varias variables.

Proposición 6.19 (Condición Necesaria de Extremo Relativo).

$$\left. \begin{array}{l} f \text{ tiene máximo relativo en } p \\ f \text{ tiene derivadas parciales en } p \end{array} \right\} \Rightarrow \nabla f(p) = (0, 0).$$

Demostración. Sea $p = (x_0, y_0)$ y supongamos, por ejemplo, que f tiene un máximo relativo en p . Entonces existe B_p tal que $\forall v \in B_p, f(v) \leq f(p)$. Por lo tanto $\forall (x, y_0) \in B_p, f(x, y_0) \leq f(x_0, y_0)$, o sea que la función $g(x) = f(x, y_0)$ posee

³Esta definición de extremo relativo no es estándar. A veces no se exige que el punto p esté en $\overset{\circ}{D}$ y se da la siguiente definición: f presenta un máximo (mínimo) relativo en p ssi $\exists B_p / \forall v \in B_p \cap D, f(v) \leq (\geq) f(p)$. Con ésta definición todo extremo absoluto es relativo.

p232 de [5]

un máximo relativo en x_0 . La existencia de f_x en p asegura la existencia de g' en x_0 , ya que $g'(x_0) = f_x(p)$. Además al ser p interior, también lo es x_0 en el dominio de g . Por el teorema correspondiente al que estamos demostrando pero para funciones de una variable, tenemos que $g'(x_0) = 0$, por lo que $f_x(p) = 0$. Análogamente $f_y(p) = 0$. \square

6.0.3. Ejemplo: $f(x, y) = x^2 + y^2$, $D = \{(x, y) / x^2 + y^2 \leq 1\}$ f tiene un mínimo relativo (que además es absoluto) en $(0, 0)$, y $\nabla f = (2x, 2y)$ se anula en $(0, 0)$.

En este caso el máximo absoluto de f ocurre en la frontera de D ($\partial D = \{(x, y) / x^2 + y^2 = 1\}$), por lo que no es un máximo relativo, y no hay allí anulación de ∇f .

Definición 24. p es *punto crítico* de f sii f es diferenciable en p y $\nabla f(p) = 0$.

6.0.4. Observación: $\nabla f(p)$ puede existir, pero f no ser diferenciable en p .

Por la Proposición 6.19, los extremos relativos son puntos críticos. Sin embargo, no todos los puntos críticos son extremos relativos.

6.0.5. Ejemplo: (paraboloide hiperbólico o silla de montar)

p233 de [5]

$$f(x, y) = x^2 - y^2.$$

El único punto crítico de f es $(0, 0)$, ya que $\nabla f = (2x, -2y)$, pero no es un extremo.

De hecho, $f(x, 0) = x^2 > 0 = f(0, 0)$ si $x \neq 0$ por lo que en toda bola de centro $(0, 0)$ la función toma valores por encima y por debajo de $f(0, 0) = 0$.

6.0.6. Notas: Ya en funciones de una variable la anulación de f' no garantiza la existencia de un extremo. Para estudiar la existencia de extremos, en ese caso, o se recurre a un estudio de crecimiento (signo de f') o se estudia la derivada 2^{da}. El primer método no es trasladable⁴ a \mathbb{R}^n , como ya se dijo. El

⁴Un interesante hecho relacionado con esto es la existencia de funciones de \mathbb{R}^2 a \mathbb{R} , diferencia-

desarrollo de Taylor introducido en la sección anterior nos permitirá extender el 2^{do} método.

1. Reconocimiento de Puntos Críticos

Supongamos que p es punto crítico de f , y que f tiene desarrollo de Taylor de orden 2. Tenemos $df|_p(v) = 0$, entonces

$$f(p+v) - f(p) = \frac{1}{2}d^2f|_p(v) + r(v), \quad \text{con } \frac{r(v)}{\|v\|^2} \xrightarrow{v \rightarrow 0} 0.$$

Si consideramos, intuitivamente, que el resto $r(v)$ es despreciable, entonces el signo de $d^2f|_p(v)$ determina el signo de $\Delta f = f(p+v) - f(p)$.

p234 de [5]

Si $d^2f(v)$ es positivo $\forall v$, $\Delta f \geq 0$ y f tendrá un mínimo en p . Naturalmente, esta versión intuitiva debe ser rigorizada pero nos da una idea de que es relevante saber estudiar el signo de d^2f , que es un polinomio homogéneo de 2^{do} grado, en los incrementos $\Delta x_1, \dots, \Delta x_n$ de las variables. ($\Delta x, \Delta y$ para el caso $n = 2$).

Un polinomio de 2^{do} grado en n variables, homogéneo, es llamado forma cuadrática en \mathbb{R}^n , y será objeto de estudio en el curso de Álgebra Lineal. Aquí nos limitaremos a dar algunas definiciones y estudiar el signo de una forma cuadrática en el caso $n = 2$.

Estudiamos la forma de d^2f :

$$\begin{aligned} d^2f_p(\Delta x, \Delta y) &= f_{xx}(p)\Delta x^2 + 2f_{xy}(p)\Delta x\Delta y + f_{yy}(p)\Delta y^2 = \\ &= (\Delta x, \Delta y) \begin{pmatrix} f_{xx}(p) & f_{xy}(p) \\ f_{xy}(p) & f_{yy}(p) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = v \cdot H \cdot v^t, \end{aligned}$$

donde v es el vector fila $(\Delta x, \Delta y)$, v^t su transpuesta, y H es llamada **matriz Hessiana** de f en p , denotada a veces $f''(p)$.

bles con un mínimo relativo no absoluto y un solo punto crítico (obviamente en el mínimo relativo). Tales ejemplos no existen en funciones de \mathbb{R} a \mathbb{R} , por la existencia de un orden total en \mathbb{R}

Definición 25. Una función $\phi : \mathbb{R}^2 \rightarrow \mathbb{R} / \phi(c) = v \cdot H \cdot v^t$, donde v es un vector fila de \mathbb{R}^2 , y H es una matriz simétrica, se llama *forma cuadrática* en \mathbb{R}^2 .

Una forma cuadrática se dice:

1. *Definida positiva* sii $\phi(v) > 0 \quad \forall v \neq 0$,
2. *Definida negativa* sii $\phi(v) < 0 \quad \forall v \neq 0$,
3. *Semidefinida* $\left\{ \begin{array}{l} \text{positiva} \\ \text{negativa} \end{array} \right\}$ sii $\forall v, \phi(v) \left\{ \begin{array}{l} \geq \\ \leq \end{array} \right\} 0, \exists u \neq 0 : \phi(u) = 0$.
4. *Indefinida* $v_1, v_2 : \phi(v_1) > 0, \phi(v_2) < 0$.

p235 de [5]

6.1.1. Ejemplos

$$1) \phi(\Delta x, \Delta y) = \Delta x^2 + 2\Delta x\Delta y + \Delta y^2 = (\Delta x + \Delta y)^2. H = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

ϕ es semidefinida positiva: $\phi(\Delta x, \Delta y) \geq 0, \phi(-1, 1) = 0$ con $(-1, 1) \neq 0$.

$$2) \phi(\Delta x, \Delta y) = \Delta x^2 + \Delta x\Delta y + \Delta y^2, \quad H = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}.$$

Para estudiar el signo, nos restringimos a la recta $\Delta y = m\Delta x$. Sustituyendo tenemos:

$$\phi(\Delta x, m\Delta x) = \Delta x^2(m^2 + m + 1).$$

Como $\forall m, m^2 + m + 1 > 0$ (ya que el discriminante es $-3 < 0$, por lo que no tiene raíces reales), $\phi(\Delta x, m\Delta x) \geq 0$ y vale 0 sólo en el origen $(0, 0)$.

Esto es válido para cualquier m , por lo tanto para cualquier recta por el origen. (para la recta $\Delta x = 0$, que faltaría, $\phi = \Delta y^2$ y también es cierto). $\Rightarrow \phi(\Delta x, \Delta y) \geq 0, \phi(\Delta x, \Delta y) = 0$ solo en $(0, 0) \Rightarrow \phi$ es definida positiva.

3) $\phi(\Delta x, \Delta y) = \Delta x^2 + 4\Delta x\Delta y + \Delta y^2 \phi(1, 1) = 6 > 0, \phi(1, -1) = -2 < 0 \Rightarrow \phi$ indefinida.

Veremos a continuación un método general para clasificar formas cuadráticas de \mathbb{R}^2 , que surge de extender el procedimiento del Ejemplo 2).

Proposición 6.20. Sea

$$\phi(\Delta x, \Delta y) = \alpha\Delta x^2 + 2\beta\Delta x\Delta y + \gamma\Delta y^2,$$

p236 de [5]

forma cuadrática de \mathbb{R}^2 . Entonces, si $H = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$,

1. Si $\det H > 0$ y $\begin{cases} \alpha > 0 \\ \alpha < 0 \end{cases}$, entonces ϕ es definida $\begin{cases} \text{positiva} \\ \text{negativa} \end{cases}$.
2. Si $\det H < 0$, entonces es indefinida.
3. Si $\det H = 0$, entonces ϕ es semidefinida.

Demostración. Restringimos la forma a las rectas por el origen $\Delta y = m\Delta x$ y $\Delta x = 0$:

$$\phi(\Delta x, m\Delta x) = \begin{cases} \Delta x^2(\alpha + 2\beta m + \gamma m^2), & \Delta y = m\Delta x \\ \gamma \Delta y^2 & \Delta x = 0. \end{cases}$$

y consideramos el discriminante Δ del polinomio de 2^{do} grado $p(m) = \alpha + 2\beta m + \gamma m^2$, que será $\Delta = 4(\beta^2 - \alpha\gamma) = -4 \det H$.

1) Si $\det H > 0$, entonces $\Delta < 0$ y p no tendrá raíces reales, por lo que tendrá signo constante igual al de $p(0) = \alpha$ y al de $\lim_{m \rightarrow \infty} p(m)/m^2 = \gamma$. Por lo tanto ϕ es definida positiva si $\alpha > 0$ y negativa si $\alpha < 0$.

2) Si $\det H < 0 \Rightarrow \Delta > 0 \Rightarrow p(m)$ tiene dos raíces reales distintas por tanto cambia de signo $\Rightarrow \exists m_1, m_2 : p(m_1) > 0, p(m_2) < 0 \Rightarrow \phi(1, m_1) > 0, \phi(1, m_2) < 0$ y ϕ es indefinida.

3) si $\Delta = 0$, y el polinomio $p(m)$ será de la forma $\gamma(m - m_0)^2 \geq 0$, para cierto $m_0 \in \mathbb{R}$. Por lo tanto, ϕ nos queda de la forma

$$\phi(\Delta x, \Delta y) = \gamma \begin{cases} \Delta x^2(m - m_0)^2 & \text{si } \Delta y = m\Delta x, \\ \Delta y^2 & \text{si } \Delta x = 0. \end{cases}$$

Expresión que es o bien nula o bien tiene el mismo signo que γ . Además $\phi(1, m_0) = 0$, siendo $(1, m_0) \neq (0, 0)$. Por lo tanto ϕ es semidefinida. \square

6.1.1. Observación: En el caso 1. en lugar de considerar el signo de α se puede considerar el signo de γ .

Ahora estamos en condiciones de dar un criterio de reconocimiento de puntos críticos.

Teorema 6.21. *f tiene un punto crítico en p, y admite desarrollo de Taylor de orden 2 en p, entonces:*

- 1) Si $d^2 f|_p$, es definida positiva, f tiene un mínimo relativo en p,
- 2) Si $d^2 f|_p$, es definida negativa, f tiene un máximo relativo en p,
- 3) Si $d^2 f|_p$, es indefinida, f no tiene máximo ni mínimo en p (se dice en este caso que f tiene un **punto de silla** en p).

(en el caso $d^2 f|_p$ semidefinida, el criterio no permite decidir)

Demostración. Sea $\phi(v) = d^2 f|_p(v)$, por Taylor (y como $df|_p = 0$),

$$f(p+v) - f(p) = \frac{1}{2}\phi(v) + r(v), \quad \text{con } \varepsilon(v) = \frac{r(v)}{\|v\|^2} \xrightarrow{v \rightarrow 0} 0,$$

$$\Rightarrow f(p+v) - f(p) = \|v\|^2 \left(\frac{1}{2} \frac{\phi(v)}{\|v\|^2} + \varepsilon(v) \right) = \|v\|^2 \left(\frac{1}{2} \phi(u) + \varepsilon(v) \right) \quad (*)$$

donde $u = v/\|v\|$, vector de módulo 1 (**versor**) colineal con v, pertenece a $S = \partial B(0; 1)$, compacto de \mathbb{R}^2 .

1) Como $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función continua (polinomio), tiene máximo y mínimo en S.

Si ϕ es definida positiva, su mínimo m debe ser positivo (pues $m = \phi(u_1)$, para cierto $u_1 \in S$), $\Rightarrow \phi(u) \geq m > 0 \forall u \in S$. Si volvemos a la expresión (*),

Como $\varepsilon(v) \xrightarrow{v \rightarrow 0} 0$, podemos elegir $B_0 : \text{si } v \in B_0, |\varepsilon(v)| < m/2$.

$\Rightarrow \text{Si } v \in B_0, (\frac{1}{2}\phi(u) + \varepsilon(v)) > (\frac{1}{2}m - m/2) = 0 \Rightarrow f(p+v) - f(p) > \|v\|^2 \cdot 0 = 0$

Por lo tanto, si B_0 tiene radio δ , entonces si $q \in B(p; \delta)$, $f(q) \geq f(p)$, de donde f tiene un mínimo relativo en p.

2) Si ϕ es definida negativa, un argumento similar muestra que f tiene un máximo relativo en p.

3) Si ϕ es indefinida, $\exists v_1, v_2/\phi(v_1) > 0, \phi(v_2) < 0$. Sean $u_1 = v_1/\|v_1\|, u_2 =$

$v_2/\|v_2\|$, entonces $u_1, u_2 \in S$ y $\phi(u_1) > 0, \phi(u_2) < 0$.

Escribamos la expresión (*) con $v = tu_1$.

$$f(p + tu_1) - f(p) = t^2 \left(\frac{1}{2}\phi(u_1) + \varepsilon(tu_1) \right)$$

Como $\varepsilon(tu_1) \xrightarrow{t \rightarrow 0} 0$, existe $\delta_1 > 0$ tal que para $|t| < \delta_1$, $|\varepsilon(tu_1)| < \frac{1}{2}\phi(u_1)$, de forma que el segundo factor es positivo, por lo tanto, para $0 \neq |t| < \delta_1$, $f(p + tu_1) > f(p)$.

En forma análoga, existe $\delta_2 > 0$ tal que para $0 \neq |t| < \delta_2$, $f(p + tu_2) < f(p) \Rightarrow$ en todo entorno de p , hay puntos donde f toma valores mayores y menores que en p , por lo tanto f no tiene un extremo relativo en p . \square

6.1.3. Notas

a) La demostración nos da una idea de por qué no puede decidirse el tipo de punto si $d^2f|_p$ es semidefinida. En ese caso, existe u_0 tal que $\phi(u_0) = 0$. Para v colineal con u_0 , $f(p + v) - f(p) = r(v)$, cuyo signo no conocemos sin usar un desarrollo de Taylor de orden mayor.

b) En el caso 3) (ϕ indefinido) el punto p se llama **punto silla**. El nombre proviene del ejemplo ya visto $f(x, y) = x^2 - y^2$. Se puede ver que en general el comportamiento local de f en torno a un punto de silla es similar al del ejemplo.

c) Hemos trabajado en $n = 2$, porque en este caso tenemos un criterio de clasificación de formas cuadráticas sencillo, basado en el determinante de la matriz H .

Sin embargo, el Teorema 6.21 es totalmente general, válido para cualquier n . En el curso de Algebra Lineal se verán métodos para clasificar formas cuadráticas de \mathbb{R}^n . Se podrá, entonces por el Teorema 6.21, reconocer puntos críticos en el caso general.

Para $f(x_1, x_2, \dots, x_n) \in C^2$ la forma cuadrática es $d^2f|_p(v) = vHv^t$, con

$$H = \begin{pmatrix} f_{x_1x_1}(p) & f_{x_1x_2}(p) & \cdots & f_{x_1x_n}(p) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(p) & f_{x_nx_2}(p) & \cdots & f_{x_nx_n}(p) \end{pmatrix}.$$

6.1.4. Ejemplos

1. Estudiar los puntos críticos de $f(x, y) = x^3 + y^3 - 9xy$. $f_x = 3x^2 - 9y$,
 $f_y = 3y^2 - 9x$.
Resolviendo $\nabla f = 0$ se obtienen los puntos críticos $(0, 0)$ y $(3, 3)$.

Para clasificarlos, obtendremos la matriz H en cada caso.

$$H = \begin{pmatrix} f_{xx}(p) & f_{xy}(p) \\ f_{xy}(p) & f_{yy}(p) \end{pmatrix} = H = \begin{pmatrix} 6x & -9 \\ -9 & 6y \end{pmatrix}.$$

En el punto $(0, 0)$, $H = \begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix} \Rightarrow \det H < 0 \Rightarrow d^2f|_{(0,0)}$ indefinida $\Rightarrow (0, 0)$ punto silla.

En el punto $(3, 3)$, $H = \begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix} \Rightarrow \left. \begin{array}{l} \det H > 0 \\ \alpha = 18 > 0 \end{array} \right\} \Rightarrow d^2f|_{(3,3)}$ definida positiva $\Rightarrow (3, 3)$ es un mínimo relativo.

2. Idem para $f(x, y) = x^3 - 3xy^2$. $f_x = 3x^2 - 3y^2$, $f_y = -6xy \Rightarrow (0, 0)$ es el único punto crítico.

En $(0, 0)$, $H = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ y d^2f es semidefinida (es nula).

En este caso el criterio del Teorema 6.21 no identifica el tipo de punto.

Sin embargo, un argumento directo nos muestra que no es un extremo:

$f(x, 0) = -x^3$, por lo tanto f toma valores positivos negativos en todo entorno de $(0, 0)$, $f(0, 0) = 0$.

3. Consideremos $f(x, y) = x^2 + y^2(x+1)^3$, su gradiente es $\nabla f(x, y) = (f_x, f_y) = (2x + y^2 3(x+1)^2, 2y(x+1)^3)$, que si es nulo entonces $2y(x+1)^3 =$

$0 \Rightarrow y = 0$ o $x = -1$. Si $y = 0 \Rightarrow 2x = 0 \Rightarrow x = 0$. Si $x = -1 \Rightarrow f(-1, y) = -2 \neq 0$, por lo tanto el único punto crítico es $(0, 0)$. Clasifiquémoslo:

$$H = \begin{pmatrix} 2 + y^2 6(x+1)^2 & 2y 3(x+1)^2 \\ 2y 3(x+1)^2 & 2(x+1)^3 \end{pmatrix} \Rightarrow H(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Por lo tanto $(0, 0)$ es un mínimo relativo. Sin embargo no es un mínimo absoluto ya que $f(x, 1)$ es un polinomio de grado 3.

2. Extremos Absolutos

Ya se dio la definición anteriormente (Definición 23). Un caso particular en que se puede asegurar la existencia de extremo (máximo y mínimo) absolutos es cuando el dominio D es compacto y es continua (teorema de Weierstrass). Supondremos, en lo que sigue, f diferenciable. Como ya se dijo, un extremo absoluto que se da en un punto interior es también relativo, y por lo tanto un punto crítico. Si se quiere entonces encontrar los extremos absolutos de una función diferenciable en un dominio D compacto, basta entonces estudiar:

1. Los puntos críticos (que estarán $\overset{\circ}{D}$).
2. La frontera de D .

6.2.0. Ejemplo: Sea $f(x, y) = x^2 + y^2 - xy - x - y$ y

$$D = \{(x, y) / x \geq 0, y \geq 0, x + y \leq 3\}$$

El dominio es el triángulo de vértices $O = (0, 0)$, $A = (3, 0)$ y $B = (0, 3)$.

Puntos críticos $\left. \begin{array}{l} f_x = 2x - y - 1 = 0 \\ f_y = 2y - x - 1 = 0 \end{array} \right\}$ punto crítico $(1, 1)$ interior a D . Además $f(1, 1) = -1$.

Frontera Segmento OA . $y = 0, 0 \leq x \leq 3, f(x, 0) = x^2 - x, (f(x, 0))' = 2x - 1,$

$$f(0) = 0, f(1/2) = -1/4, f(3) = 6:$$

Máximo de f en OA es 6 en $A = (3, 0)$

Mínimo de f en OA es $-1/4$ en $(1/2, 0)$

Segmento OB (análogo)

Máximo de f en OB es 6 en $B = (0, 3)$

Mínimo de f en OB es $-1/4$ en $(0, 1/2)$

Segmento AB . $y = 3 - x, 0 \leq x \leq 3,$

$$f(x, 3-x) = 3(x^2 - 3x + 2) \xrightarrow{\frac{d}{dx}} 3(2x - 3), f(0, 3) = 6 = f(3, 0), f\left(\frac{3}{2}, \frac{3}{2}\right) = -3/4.$$

Máximo de f en AB es 6 en A y B

Mínimo de f en AB es $-3/4$ en $(3/2, 3/2)$

Se sabe que f tiene máximo y mínimo absoluto en D . Luego debe ser alguno de los puntos encontrados. Comparando, se deduce:

Máximo absoluto de f en D : 6 en los puntos $A = (3, 0)$ y $B = (0, 3)$,

Mínimo absoluto de f en D : -1 en el punto $(1, 1)$. (es además mínimo relativo)

p243 de [5]

6.7. Funciones Implícitas y Aplicaciones

En muchos problemas se tiene dada una ecuación $f(x, y) = 0$ verificada por dos variables, y se quiere "despejar" $y = \varphi(x)$ (o $x = \psi(y)$). Esto quiere decir encontrar una función $\varphi(x)$ tal que los puntos (x, y) cumplan $f(x, y) = 0$ sii $y = \varphi(x)$.

De un modo más geométrico, poner la curva $f(x, y) = 0$ como el gráfico de una función. Esto no siempre es posible: la ecuación $x^2 + y^2 - 1 = 0$, por ejemplo, no es el gráfico de ninguna función $y = \varphi(x)$ o $x = \psi(y)$.

Sin embargo, si nos conformamos con una solución local, es posible, en un entorno de un punto que verifica la ecuación, identificar a la curva con el gráfico de alguna de las cuatro funciones $y = \sqrt{1 - x^2}$, $y = -\sqrt{1 - x^2}$, $x = \sqrt{1 - y^2}$ o $x = -\sqrt{1 - y^2}$.

Nos preguntamos si lo anterior es general: es decir, si toda ecuación $f(x, y) = 0$, con f suficientemente regular, podrá describirse localmente por una función $y = \varphi(x)$ o $x = \psi(y)$.

El siguiente ejemplo muestra que no: con $f(x, y) = x^2 - y^2$, $f(x, y) = 0$

representa el par de rectas $y = x, y = -x$. En un entorno de O , no es el gráfico de una función. Hace falta, entonces, dar alguna hipótesis para asegurarlo. Si miramos el problema en forma geométrica, $f(x, y) = 0$ equivale a cortar el plano $z = 0$ con $z = f(x, y)$.

p244 de [5]

En un punto p de la intersección, supongamos que existe el plano tangente. Si este plano no es horizontal, cortará al plano $z = 0$ en una recta, donde siempre puede despejarse y en función de x o viceversa. Es natural que para la curva $f(x, y) = 0$ ocurra lo mismo.

Veremos a continuación un teorema que asegura que si f es suficientemente regular, y el plano tangente no es horizontal en cierto punto p , se puede escribir en la forma deseada.

Teorema 6.22 (Teorema de la función implícita). *Sea $f : U(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ de clase C^1 en U tal que $f(x_0, y_0) = 0$ con $(x_0, y_0) \in U$ y $f_y(x_0, y_0) \neq 0$. Entonces: existe un rectángulo abierto $I \times J \subset U$ de centro (x_0, y_0) y una función $\varphi : I \rightarrow J$ de clase C^1 tal que:*

$$\forall (x, y) \in I \times J \text{ se cumple: } f(x, y) = 0 \iff y = \varphi(x).$$

O sea que $f^{-1}(0) \cap I \times J$ es el gráfico de una función $\varphi : I \rightarrow J$.

En otras palabras, la ecuación $f(x, y) = 0$ determina localmente a y en función de x , alrededor de (x_0, y_0) . Además se tiene

$$\varphi'(x) = -\frac{f_x(x, \varphi(x))}{f_y(x, \varphi(x))}.$$

Demostración. Supongamos, para fijar ideas, que $f_y(x_0, y_0) > 0$. Como f_y es continua, existen una bola $B_1 \subset U$ de centro (x_0, y_0) y radio r , tal que $f_y(x, y) > 0 \quad \forall (x, y) \in B$. Entonces, para cualquier $0 < \varepsilon < r$, si $J = (y_0 - \varepsilon; y_0 + \varepsilon)$, el segmento $\bar{J} \times \{x_0\}$ estará incluido en B . Además la función $F(y) = f(x_0, y)$ será estrictamente creciente $\bar{J} \times \{x_0\}$, por lo tanto, como $F(y_0) = 0$, tendremos

p245 de [5]

$f(x_0, y_0 - \varepsilon) = F(y_0 - \varepsilon) < 0$ y $f(x_0, y_0 + \varepsilon) = F(y_0 + \varepsilon) > 0$. La continuidad de f nos permite tomar bolas B_1 y B_2 de centros $(x_0, y_0 - \varepsilon)$ y $(x_0, y_0 + \varepsilon)$ incluidas en B , donde la función será negativa y positiva respectivamente. Además, si δ el menor de los radios de B_1 y B_2 , se cumplirá

$$(*) \quad f(x, y_0 - \varepsilon) < 0 \quad \text{y} \quad f(x, y_0 + \varepsilon) > 0, \quad \forall x \in I = (x_0 - \delta; x_0 + \delta).$$

Resulta entonces que, para cada $x \in I$, la función $F(y) = f(x, y)$ es estrictamente creciente en \bar{J} , y continua en él. De (*) sabemos que $F(y_0 - \varepsilon) < 0$ y $F(y_0 + \varepsilon) > 0$, de donde se deduce, por Bolzano, que para cada $x \in I$, la función $F(y)$, tiene una única raíz, que llamaremos $\varphi(x)$, en \bar{J} . Es decir, para cada $x \in I$ $F(y) = 0 \iff y = \varphi(x)$ o sea para cada $x \in I$, $f(x, y) = 0 \iff y = \varphi(x)$. Como $F(y_0 - \varepsilon) < 0$ y $F(y_0 + \varepsilon) > 0$, es claro que $\varphi(x) \in J \subset \bar{J}$.

Hemos demostrado que existe una función $\varphi : I \rightarrow J$ tal que $\forall (x, y) \in I \times J$ se tiene:

$$(**) \quad f(x, y) = 0 \iff y = \varphi(x).$$

Vamos a probar ahora que φ es continua y derivable en I .

Sea $x \in I$, h tal que $x + h \in I$ y $k = \varphi(x + h) - \varphi(x)$. Por Lagrange aplicado a $g(t) = f(x + th, \varphi(x) + tk)$ en $[0, 1]$ sabemos que existe $\theta \in (0, 1)$ tal que

$$\underbrace{f(x + h, \varphi(x) + k)}_{(A)} - \underbrace{f(x, \varphi(x))}_{(B)} = f_x(\tilde{p}) \cdot h + f_y(\tilde{p}) \cdot k. \quad \text{con } \tilde{p} = (x + \theta h, \varphi(x) + \theta k).$$

Afirmo que tanto (A) como (B) con nulos. Efectivamente por (**) sabemos que (B) es nulo. Respecto de (A), por la definición de k tenemos $\varphi(x) + k = \varphi(x + h)$, por lo tanto, nuevamente por (**) tenemos (A) nulo. En definitiva $f_x(\tilde{p}) \cdot h + f_y(\tilde{p}) \cdot k = 0$, de donde podemos despejar k y obtener

$$(**) \quad k = -h \frac{f_x(\tilde{p})}{f_y(\tilde{p})} \quad \text{con } \tilde{p} = (x + \theta h, \varphi(x) + \theta k).$$

Como f_x y f_y son continuas en $\bar{I} \times \bar{J}$ y f_y no nula, su cociente f_x/f_y también será una función continua en $\bar{I} \times \bar{J}$ compacto y por lo tanto acotada. Al ser f_x/f_y acotada, tomando límites en (***) vemos que $k \rightarrow 0$ cuando $h \rightarrow 0$. Por lo tanto $\tilde{p} \rightarrow (x, \varphi(x))$ cuando $h \rightarrow 0$. Despejando k/h de (***) y tomando límites obtenemos

$$\frac{k}{h} = -\frac{f_x(\tilde{p})}{f_y(\tilde{p})} \xrightarrow{h \rightarrow 0} -\frac{f_x(x, \varphi(x))}{f_y(x, \varphi(x))}.$$

Viendo la definición de k tenemos que k/h es el cociente incremental que define $\varphi'(x)$, por lo tanto dicha derivada existe y vale

$$(***) \quad \varphi'(x) = -\frac{f_x(x, \varphi(x))}{f_y(x, \varphi(x))}.$$

Nos falta ver que φ' es continua. Primero observemos que φ es continua pues $k \rightarrow 0$ cuando $h \rightarrow 0$. Además, como f_x y f_y son continuas y f_y no nula, la fórmula (***) nos dice que φ' es continua, como queríamos demostrar. \square

p246 de [5]

7.0.1. Observación:

1) Si $f_y(x_0, y_0) = 0$ entonces no se puede afirmar nada.

Por ejemplo $f(x, y) = x^2 + y^2 - 1$ en $(x_0, y_0) = (1, 0)$.

Si $f_x(x_0, y_0) \neq 0$ entonces la ecuación $f(x, y) = 0$ determina localmente a x como función de y . Podemos resumir los dos casos de la siguiente manera si $\nabla f(p) \neq 0$ entonces existe B_p entorno de $p = (x_0, y_0)$ y $\alpha : (-\epsilon; \epsilon) \subset \mathbb{R} \rightarrow \mathbb{R}^2$ tal que: $\alpha(0) = p$, $\alpha'(t) \neq 0$ y $\forall (x, y) \in B_p$, $f(x, y) = 0 \iff \exists t \in I : (x, y) = \alpha(t)$.

Si (x_0, y_0) es punto crítico puede suceder que la ecuación no determina función alguna alrededor de (x_0, y_0) .

2) Si la función dada f es de clase C^k entonces se concluye fácilmente que φ también es C^k . En ese caso, para hallar las derivadas sucesivas de φ se pueden aplicar las reglas de derivación, y la regla de la cadena en la formula de φ' o usar el siguiente argumento:

Sabemos que $F(x) = f(x, \varphi(x)) = 0, \forall x \in I \Rightarrow$

p247 de [5]

$$0 = F'(x) = f_x(x, \varphi(x)) + f_y(x, \varphi(x)) \cdot \varphi'(x) \forall x \in I \Rightarrow$$

$$0 = F'' = f_{xx} + f_{xy} \varphi' + (f_{xy} + f_{yy} \cdot \varphi') \cdot \varphi' + f_y \cdot \varphi''$$

(Las derivadas parciales están evaluadas en $(x, \varphi(x))$ y las derivadas de φ en x) Sustituyendo φ' por su expresión, de la última identidad se puede despejar φ'' .

3) El teorema anterior puede generalizarse para funciones de más de dos variables.

Si $f(x_1, \dots, x_n, y)$ es una función de $n+1$ variables de clase C^1 , $f(x_{10}, \dots, x_{n0}, y_0) = 0$ y $f_y(x_{10}, \dots, x_{n0}, y_0) \neq 0$, entonces existe $I \times J$, (I entorno de (x_{10}, \dots, x_{n0}) en \mathbb{R}^n , J entorno de y_0) y $\varphi : I \rightarrow J$ de clase C^1 tales que

$$\forall (x_{10}, \dots, x_{n0}, y_0) \in I \times J, f(x_1, \dots, x_n, y) = 0 \iff y = \varphi(x_1, \dots, x_n).$$

Además, como $f(x_1, \dots, x_n, \varphi(x_1, \dots, x_n)) \equiv 0$ en I , obtenemos por diferenciación que:

$$\frac{\partial \varphi}{\partial x_i} = \frac{-f_{x_i}(p)}{f_y(p)},$$

donde $p = (x_1, \dots, x_n, \varphi(x_1, \dots, x_n))$.

7.0.2. Ejemplos:

p248 de [5]

1. Consideremos la circunferencia $x^2 + y^2 = 1$ cerca del punto $p = (1/\sqrt{2}, 1/\sqrt{2})$.

Consideramos la función $f(x, y) = x^2 + y^2 - 1$ de clase C^2 que en p vale 0. Como $f_y(x, y) = 2y$, entonces $f_y(p) = 2/\sqrt{2} \neq 0$, por lo tanto podemos aplicar el Teorema 6.22 y deducir la existencia una $\varphi : I \rightarrow J$ para un entorno I de $1/\sqrt{2}$ tal que $x^2 + \varphi^2(x) = 1$. La derivada de φ en $1/\sqrt{2}$, será

$$\varphi'(1/\sqrt{2}) = -\frac{f_x(1/\sqrt{2}, 1/\sqrt{2})}{f_y(1/\sqrt{2}, 1/\sqrt{2})} = -\frac{2(1/\sqrt{2})}{2(1/\sqrt{2})} = -1$$

ya que $f_x(x, y) = 2x$. Sin embargo es más fácil "derivar" directamente la ecuación $x^2 + \varphi^2(x) = 1$. Efectivamente derivando obtenemos $2x +$

$2\varphi(x)\varphi'(x) = 0$. Evaluando en $x = 1/\sqrt{2}$ tenemos

$$\frac{2}{\sqrt{2}} + 2\varphi(\sqrt{2})\varphi'(\sqrt{2}) = 0 \Rightarrow \frac{2}{\sqrt{2}} + \frac{2}{\sqrt{2}}\varphi'(\sqrt{2}) = 0 \Rightarrow \varphi'(\sqrt{2}) = -1.$$

Es importante observar que en la última ecuación es posible “despejar” $\varphi'(\sqrt{2})$ gracias a que $2\varphi(\sqrt{2}) \neq 0$, que es exactamente la condición $f_y(p) \neq 0$ del teorema.

En realidad en la práctica uno procede “derivando” la ecuación $x^2 + y^2 = 1$ de la siguiente forma $2x + 2yy' = 0$ de donde $y' = -x/y$. Si $x = y = 1/\sqrt{2}$, tenemos $y' = -1$. De forma similar podemos obtener y'' : derivando $x + yy' = 0$ tenemos $1 + y'y' + yy'' = 0$ de donde si $x = y = 1/\sqrt{2}$ y $y' = -1$ tenemos $1 + 1 + (1/\sqrt{2})y'' = 0 \Rightarrow y'' = -2\sqrt{2}$, que es negativa como puede observarse gráficamente.

2. Dada la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ determinar la pendiente de la recta tangente en un punto (x_0, y_0) , con $y_0 \neq 0$. La pendiente es $\varphi'(x)$, donde $\varphi(x)$ es la función que se obtiene despejando y en función de x ($f_x = 2x/a^2$, $f_y = 2y/b^2$, $f_y(x_0, y_0)$). Por el Teorema 6.22

$$\varphi'(x) = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)} = -\frac{2x_0/a^2}{2y_0/b^2} = \left(\frac{b}{a}\right)^2 \frac{x_0}{y_0}.$$

3. Podemos reobtener el teorema de la función inversa de funciones de una variable: sea $F : \mathbb{R} \rightarrow \mathbb{R}$, $x_0 = F(y_0)$ y $F'(y_0) \neq 0$. Si $f(x, y) = x - F(y)$, entonces $f(x_0, y_0) = 0$ y $f_y(x, y) = F'(y)$ no nula en y_0 , por lo tanto, el Teorema 6.22 no asegura que existirán I, J y $\varphi : I \rightarrow J$ tales que $x_0 \in I$, $y_0 \in J$ y $\forall (x, y) \in I \times J$, $f(x, y) = 0$ sii $y = \varphi(x)$. Es decir, $\forall (x, y) \in I \times J$, $x - F(y) = 0$ sii $y = \varphi(x)$ es decir $\forall (x, y) \in I \times J$, $x = F(y)$ sii $y = \varphi(x)$, por lo tanto F es inversa de φ y viceversa. Además

$$\varphi'(x) = -\frac{f_x(x, y)}{f_y(x, y)} = -\frac{1}{-F'(y)} = \frac{1}{F'(\varphi(x))}.$$

Como dice el teorema correspondiente en una variable.

1. Extremos Condicionados (o Ligados)

En muchos problemas se busca calcular el máximo o mínimo de una función de varias variables, sujetos a alguna condición adicional (ligadura).

Por ejemplo, se quiere hallar el punto de la curva $g(x, y) = 0$ que este a menor distancia del origen. Esto equivale a minimizar la función $d = \sqrt{x^2 + y^2}$ entre los puntos que verifican $g(x, y) = 0$.

Definición 26. Sean $f, g : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$, $\mathcal{C} = \{(x, y) \in D : g(x, y) = 0\}$, $p \in \mathcal{C}$.

Entonces decimos que f presenta en p un $\left\{ \begin{array}{l} \text{máximo} \\ \text{mínimo} \end{array} \right\}$

- *absoluto condicionado* a $g(x, y) = 0$ sii $\forall v \in \mathcal{C}$, $f(v) \left\{ \begin{array}{l} \leq \\ \geq \end{array} \right\} f(p)$.
- *relativo condicionado* a $g(x, y) = 0$ sii $\exists B_p \subset D : \forall v \in B_p \cap \mathcal{C}$, $f(v) \left\{ \begin{array}{l} \leq \\ \geq \end{array} \right\} f(p)$.

p249 de [5]

7.1.0. Ejemplo: Sea $f(x, y) = x \log x + y \log y$, $x > 0, y > 0$. Hallar el mínimo de condicionado a $x + y - 1 = 0$.

En los puntos de $x + y - 1 = 0, y = 1 - x$. Escribimos entonces

$$f(x, 1 - x) = x \log x + (1 - x) \log(1 - x) = z(x).$$

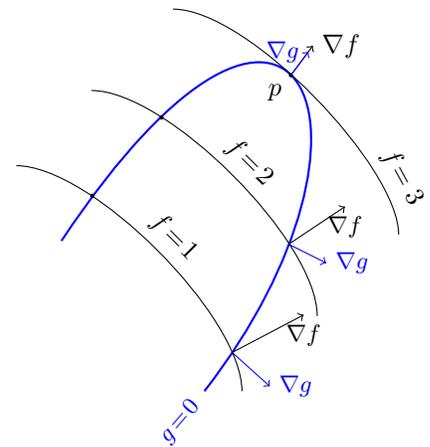
Estudiamos el crecimiento de $z(x)$ en $(0, 1)$. Hallamos $z'(x) = \log(x/(1 - x))$ de donde el signo de z' es -1 si $x < 1 - x$ sii $x < 1/2$ y positivo si $x > 1/2$. de donde el mínimo (absoluto) de $z(x)$ es $f(1/2, 1/2) = -\log 2$.

7.1.1. Observación: en el ejemplo, el éxito de la solución depende de haber podido “despejar” $y = \varphi(x)$ de la ecuación $g(x, y) = 0$, y así reducir el problema a un estudio de extremos sin condicionar de una función de una variable. No siempre será fácil, o posible, despejar en forma explícita $y = \varphi(x)$, pero el teorema de la función implícita nos permitirá eludir la dificultad. Esto origina el siguiente método del multiplicador de Lagrange.

Para tener una idea geométrica del resultado que sigue,

p250 de [5]

dibujamos las curvas de nivel $f = 0, f = 1$, etc. y la curva $g = 0$. En la figura, el punto p sería un máximo relativo de f condicionado a $g = 0$. Se observa que las curvas de nivel de f y g son tangentes en p . Por la Proposición 6.16, deducimos que los vectores $\nabla f, \nabla g$ deben ser colineales en p . Formalizamos estas ideas a continuación.



Teorema 6.23 (Método del multiplicador de Lagrange). Sean f, g de clase C^1 tales que $g(p) = 0, \nabla g(p) \neq 0$ y f presenta en p un extremo relativo condicionado a $g = 0$. Entonces $\exists \lambda \in \mathbb{R} : \nabla f(p) + \lambda \nabla g(p) = 0$. (λ es llamado *multiplicador de Lagrange*)

Demostración. Como $\nabla g(p) \neq 0$ entonces existe B_p entorno de $p = (x_0, y_0)$ y $\alpha : (-\epsilon; \epsilon) \subset \mathbb{R} \rightarrow \mathbb{R}^2$ tal que: $\alpha(0) = p, \alpha'(t) \neq 0$ y $\forall (x, y) \in B_p, g(x, y) = 0 \iff \exists t \in I : (x, y) = \alpha(t)$. Por lo tanto tenemos $g(\alpha(t)) = 0 \forall t \in I$, de donde, derivando tenemos $\nabla g(\alpha(t)) \cdot \alpha'(t) = 0 \forall t \in I$. En particular para $t = 0$ tenemos

$$(*) \quad \nabla g(p) \cdot \alpha'(0) = 0.$$

Por otro lado, si $h(t) = f(\alpha(t))$, por definición h tendrá un extremo relativo en 0, por lo tanto $h'(0) = 0$, pero $h'(t) = \nabla f(\alpha(t)) \cdot \alpha'(t)$. Por lo tanto $\nabla f(\alpha(0)) \cdot \alpha'(0) = 0$ es decir

$$(**) \quad \nabla f(p) \cdot \alpha'(0) = 0.$$

De (*) y (**), $\nabla g(p)$ y $\nabla f(p)$ son ortogonales a $\alpha'(0) \neq 0$, por lo tanto $\nabla g(p)$ y $\nabla f(p)$ son colineales. Como además $\nabla g(p) \neq 0$, existirá λ como en las hipótesis. \square

$g(x, y) = 0$, y $\nabla g(p) \neq 0$, es necesario que exista λ tal que:

$$\begin{cases} f_x(x_0, y_0) + \lambda g_x(x_0, y_0) = 0, \\ f_y(x_0, y_0) + \lambda g_y(x_0, y_0) = 0, \\ g(x_0, y_0) = 0. \end{cases}$$

Si podemos resolver el sistema anterior de tres ecuaciones en las incógnitas x_0, y_0, λ tendremos los puntos (x_0, y_0) donde puede ocurrir un extremo condicionado, (además del multiplicador λ que en general no nos interesa).

La condición anterior no es suficiente. Harán falta otros argumentos para decidir si los “candidatos” (x_0, y_0) obtenidos de las ecuaciones corresponden a extremos.

Una forma habitual de sintetizar las ecuaciones anteriores es escribir

$$\Phi(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

e imponer $\nabla \Phi = 0$.

Recalcamos que para que lo anterior valga, hace falta $\nabla g \neq 0$. Puntos con $\nabla g = 0$ deben considerarse aparte.

7.1.3. Ejemplos: 1) Máximo y mínimo absolutos de $f(x, y) = x \cdot y$ en $x^2 + y^2 = 1$.

Como f es continua y el conjunto $\mathcal{C} : \{(x, y) / x^2 + y^2 = 1\}$ es compacto, se sabe que deben existir el máximo y el mínimo absolutos (que son además relativos).

Buscamos entonces por el procedimiento del Teorema 23 los candidatos posibles, y comparando valores obtenemos máximo y mínimo. $\nabla g = (2x, 2y)$ no se anula en

$$\Phi(x, y, \lambda) = xy + \lambda(x^2 + y^2 - 1).$$

De donde $\nabla \Phi = (y + 2\lambda x, x + 2\lambda y, x^2 + y^2 - 1)$. Haciendo $\nabla \Phi = (0, 0, 0)$,

obtenemos tres ecuaciones

$$\begin{cases} y + 2\lambda x = 0, & (1) \\ x + 2\lambda y = 0, & (2) \\ x^2 + y^2 - 1 = 0. & (3) \end{cases}$$

De despejando x de (1) y sustituyendo en (2) se obtiene $y(1 - 4\lambda^2) = 0$ o $y = 0$ o $\lambda = \pm 1/2$. Si ocurriera lo primero, entonces de (2) obtenemos $x = 0$, pero $y = x = 0$ no verifica (3). Por lo tanto se cumple lo segundo ($\lambda = \pm 1/2$).

Si $\lambda = 1/2$, tenemos

$$\begin{cases} x + y = 0 \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{matrix} p_1 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \\ p_2 = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right). \end{matrix}$$

Si $\lambda = -1/2$, tenemos

$$\begin{cases} x - y = 0 \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{matrix} p_3 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \\ p_4 = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right). \end{matrix}$$

Los puntos p_1, p_2, p_3 y p_4 , son los "candidatos" a extremo. Calculamos los va-

lores de f : $f(p_1) = f(p_2) = -1/2 \Rightarrow$ Mínimo Absoluto,

$d(p_3) = f(p_4) = 1/2 \Rightarrow$ Máximo Absoluto.

2) Mínima distancia el origen de puntos de $\mathcal{C} = \{(x, y)/(x - 1)^3 - y^2 = 0\}$.

Tomamos $f(x, y) = x^2 + y^2$ (cuadrado de la distancia)

$$\Phi(x, y, \lambda) = x^2 + y^2 + \lambda[(x - 1)^3 - y^2].$$

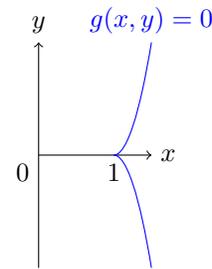
$$\text{Condiciones} \begin{cases} 2x + 3\lambda(x - 1)^2 = 0, & (1) \\ 2y - 2\lambda y = 0, & (2) \\ (x - 1)^3 - y^2 = 0. & (3) \end{cases}$$

De (2), se obtiene $y = 0$ o $\lambda = 1$. Si $y = 0 \stackrel{(3)}{\Rightarrow} x = 1$ no cumple (1). Si $\lambda = 1$, sustituida en (1) nos da:

$$2x + 3(x - 1)^2 = 0 \iff 3x^2 - 4x + 3 = 0.$$

La ecuación $3x^2 - 4x + 3 = 0$ no tiene raíces en \mathbb{R} , por lo tanto el sistema es incompatible. No tenemos entonces, candidatos a extremo. Sin embargo, intuitivamente, el mínimo debía existir. Lo que ocurre es que no hemos verificado la condición $\nabla g = 0$. $\nabla g = (3(x - 1)^2, -2y)$ se anula en $(1, 0)$, que cumple $g(1, 0) = 0$.

El punto $(1, 0)$ es entonces un candidato más a extremo, que será el mínimo buscado. La curva $g(x, y) = 0$ es la de la figura.



2. Extensiones del Método

1) Para más de dos variables.

Se buscan extremos de $f(x_1, \dots, x_n)$ condicionado a $g(x_1, \dots, x_n) = 0$.

Si $\nabla g \neq 0$, la condición necesaria es $\nabla \Phi = 0$, con $\Phi(x_1, \dots, x_n, \lambda) = f + \lambda g$.
(demostración análoga)

7.2.0. Ejemplo: Se quiere construir un paralelogramo de cartón con $2k$ metros cuadrados, que tenga el máximo volumen posible.

Tenemos $V = x \cdot y \cdot z$, que queremos maximizar condicionado a

$$2(xy + yz + xz) = 2k, \quad x > 0, y > 0, z > 0$$

$g(x, y, z) = xy + yz + xz - k$, $\nabla g = (y + z, x + z, x + y)$ solo se anula en $(0, 0, 0)$.

$$\phi(x, y, z, \lambda) = xyz + \lambda(xy + yz + xz - k).$$

Condiciones

$$yz + \lambda(y + z) = 0, \quad (1)$$

$$xz + \lambda(x + z) = 0, \quad (2)$$

$$xy + \lambda(x + y) = 0, \quad (3)$$

$$xy + yz + xz = k. \quad (4)$$

De (1), (2) y (3) tenemos

$$-\lambda = \frac{yz}{y+z} = \frac{xz}{x+z} = \frac{yx}{y+x},$$

de donde $x = y = z \stackrel{(4)}{=} \sqrt{k/3}$. El candidato a extremo es el cubo. Para demostrar que es el máximo, hace falta algún argumento adicional. Una forma de hacerlo es demostrar que el máximo debe existir, considerando la función continua $V = x.y.z$ y el conjunto compacto $\{(x, y, z) : x \geq 0, y \geq 0, z \geq 0, xy + yz + xz = k\}$.

2) Más de dos variables, más de una condición.

Se buscan extremos de $f(x_1, \dots, x_n)$ condicionado a las ligaduras

$$\begin{cases} g_1(x_1, \dots, x_n) = 0, \\ \vdots \\ g_h(x_1, \dots, x_n) = 0. \end{cases} \quad h < n,$$

Para hallar una condición necesaria en este caso hace falta una generalización el Teorema 6.22 de la función implícita que permite “despejar”, en el sistema anterior, $n - h$ variables en función de las restantes h . Esta versión (Teorema 6.29) se verá más adelante, pero enunciamos aquí la condición necesaria de extremos condicionados, que se deduce en forma similar al Teorema 6.23.

p255 de [5]

Si en $p = (x_{10}, \dots, x_{n0})$ hay extremo condicionado, y $\nabla g_1(p), \dots, \nabla g_h(p)$ son li-

nealmente independientes entonces existen escalares $\lambda_1, \dots, \lambda_h$ tales que

$$\nabla f(p) + \lambda_1 \nabla g_1(p) + \dots + \lambda_h \nabla g_h(p) = 0.$$

De otro modo, la función

$$\Phi(x_1, \dots, x_n, \lambda_1, \dots, \lambda_h) = f(p) + \lambda_1 g_1(p) + \dots + \lambda_h g_h(p).$$

tiene un punto crítico para algunos λ_i y $p = (x_1, \dots, x_n)$.

Los λ_i se llaman **multiplicadores de Lagrange**.

Veamos la necesidad de dicha condición para $n = 3$ y $h = 2$, asumiendo que se pueden despejar dos variables en función de la tercera. Sea entonces la función $f(x, y, z)$ y las condiciones $g_1(x, y, z) = 0$, $g_2(x, y, z) = 0$. Supongamos que podemos despejar y y z en función de x en un entorno de p : $y = y(x)$, $z = z(x)$. Sustituyendo en f tenemos $h(x) = f(x, y(x), z(x))$. Como f tiene un extremo en $p = (x_0, y_0, z_0)$, entonces h tendrá un extremo en x_0 , por lo que $h'(x_0) = 0$. Calculando h' tenemos

$$0 = h'(x_0) = f_x(p) + f_y(p)y_x(x_0) + f_z(p)z_x(x_0). \quad (*_1)$$

Por otro lado, sustituyendo $y(x)$, $z(x)$ en las condiciones y derivando tenemos

$$g_{1x}(p) + g_{1y}(p)y_x(x_0) + g_{1z}(p)z_x(x_0) = 0, \quad (*_2)$$

$$g_{2x}(p) + g_{2y}(p)y_x(x_0) + g_{2z}(p)z_x(x_0) = 0. \quad (*_3)$$

Las ecuaciones $(*_1)$, $(*_2)$, y $(*_3)$, nos dice que los gradientes $\nabla f(p)$, $\nabla g_1(p)$ y $\nabla g_2(p)$ son perpendiculares al vector $(1, y_x(x_0), z_x(x_0))$ (no nulo), por lo tanto pertenecerán al plano perpendicular él. En el caso de que $\nabla g_1(p)$ y $\nabla g_2(p)$ sean L.I., todos los vectores del plano se podrán expresar como combinación lineal de ellos, en particular $\nabla f(p)$. De dicha combinación surgen λ_1 y λ_2 .

7.2.1. Ejemplo: Hallar puntos de

$$\mathcal{C} = \begin{cases} x^2 + y^2 + z^2 = 1, \\ x + y + 2z = 0, \end{cases}$$

de altura máxima y mínima.

\mathcal{C} representa en el espacio, la intersección de una esfera y un plano, es decir una circunferencia.

Como $\{\nabla g_1, \nabla g_2\} = \{(2x, 2y, 2z), (1, 1, 2)\}$, es linealmente dependiente sii $x = y = z/2$. Se verifica fácilmente que el sistema

$$\begin{cases} x^2 + y^2 + z^2 = 1, \\ x + y + 2z = 0, \\ x = y = \frac{z}{2}, \end{cases}$$

es incompatible, de modo que sobre \mathcal{C} , ∇g_1 y ∇g_2 son linealmente independiente.

Tomamos $f(x, y, z) = z$ (altura) y $\Phi = f + \lambda_1 g_1 + \lambda_2 g_2$. Imponiendo $\nabla \Phi = 0$, se obtiene

$$\begin{cases} 2\lambda_1 x + \lambda_2 = 0, & (1) \\ 2\lambda_1 y + \lambda_2 = 0, & (2) \\ 1 + 2\lambda_1 z + 2\lambda_2 = 0, & (3) \\ x^2 + y^2 + z^2 = 1, & (4) \\ x + y + 2z = 0 & (5). \end{cases}$$

De (1) y (2), tenemos $x = y$ o $\lambda_1 = 0$. Ésto último es incompatible con (1) y (3). Luego de (5) y $x = y$, tenemos $z = -x$. Sustituyendo en (4) se deduce $x = y = -z = \pm 1/\sqrt{3}$. Hay dos candidatos a extremo $A = (1/\sqrt{3})(1, 1, -1)$ y $B = (1/\sqrt{3})(-1, -1, 1)$, mínimo y máximo respectivamente.

6.8. Funciones de \mathbb{R}^n en \mathbb{R}^m

Extendemos la teoría desarrollada para $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$ a funciones con recorrido en \mathbb{R}^m .

Sea $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^m$,

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))^t,$$

donde $f_i : D \rightarrow \mathbb{R}$ son funciones como las estudiadas hasta ahora y se llaman **funciones coordenadas** de la función f .

8.0.2. Ejemplo: Se considera un cuerpo de masa m sometido a la atracción gravitatoria de la tierra, de masa M , situada en el origen.

Si la posición de la partícula es (x, y, z) , buscamos la función $\vec{F}(x, y, z)$ que nos de el vector fuerza de atracción sobre el cuerpo.

Siendo $r = \sqrt{x^2 + y^2 + z^2}$ la distancia al origen, la fuerza apunta hacia el origen y su módulo es $|\vec{F}| = GMm/r^2$, de acuerdo con la Ley de la Gravitación Universal.

Si $\vec{u} = (1/r)(x, y, z)$, la fuerza es $\vec{F} = (GMm/r^2)\vec{u}$, o sea:

$$F(x, y, z) = \left(\frac{-GMmx}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-GMmy}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-GMmz}{(x^2 + y^2 + z^2)^{3/2}} \right)^t.$$

Tenemos entonces una función $\vec{F} : D \rightarrow \mathbb{R}^3$, con $D = \mathbb{R}^3 \setminus \{(0, 0, 0)\}$.

1. Continuidad

Definición 27. $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^m$ es **continua** en p sii $p \in D$ y dado $B_{f(p)}$ existe $B_p / f(B_p \cap D) \subset B_{f(p)}$.

La definición es idéntica a la Definición 15; en este caso, la $B_{f(p)}$ es una bola de \mathbb{R}^m . Con demostración análoga tenemos una versión de la continuidad en términos de sucesiones.

Proposición 6.24. $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^m$ continua en p sii $p \in D$ y para toda sucesión $(x_k) \subset D$, con $x_k \rightarrow p$, se cumple $f(x_k) \rightarrow f(p)$.

Demostración. Demostración: Análoga a Proposición 6.7. □

Proposición 6.25. $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^m$, $f = (f_1, f_2, \dots, f_m)^t$, entonces f continua en p sii f_i continua en p para $i = 1, \dots, m$.

p258 de [5]

Demostración. (\Rightarrow) Supongamos $x_k \rightarrow p$, $(x_k) \subset D$, f continua en p entonces $f(x_k) \rightarrow f(p)$.

Como la convergencia de una sucesión en \mathbb{R}^m equivale a la de sus coordenadas (Proposición 6.2), $f_i(x_k) \rightarrow f_i(p)$, $i = 1, \dots, m$.

Entonces f_i continua en p para $i = 1, \dots, m$.

El recíproco es análogo. □

8.1.0. Ejemplo: La función del ejemplo anterior es continua en $\mathbb{R}^2 \setminus \{(0, 0)\}$ porque sus funciones coordenadas lo son. (Por ejemplo $-GMm x / (x^2 + y^2 + z^2)^{3/2}$ es continua en $\mathbb{R}^2 \setminus \{(0, 0)\}$)

2. Diferenciabilidad

Como se vio, una función con recorrido en \mathbb{R} es diferenciable en $p \in D$ si su incremento $f(p+v) - f(p)$ puede ser aproximado por una función $df|_p(v)$ lineal en v .

Es decir

$$f(p+v) - f(p) = df|_p(v) + r(v), \quad r(v) \xrightarrow{v \rightarrow 0} 0 \quad (\text{vector nulo de } \mathbb{R}^m)$$

$$df|_p(v) = \nabla f(p) \cdot v = \underset{\substack{\uparrow \\ \text{para } n=2}}{f_x(p)\Delta x + f_y(p)\Delta y}$$

Hacemos énfasis en el carácter lineal de $df|_p$.

p259 de [5]

$df|_p(v)$ es una función (o transformación) lineal de \mathbb{R}^n en \mathbb{R} . Es decir,

- $df|_p(\lambda v) = \lambda df|_p(v)$ (homogeneidad),
- $df|_p(v_1 + v_2) = df|_p(v_1) + df|_p(v_2)$ (aditividad)

De otro modo: la función $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}$, que no es en general lineal, se aproxima localmente (cerca de p) por una función lineal. Extendemos este concepto para f con recorrido en \mathbb{R}^m .

Definición 28. Sea $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^m$, p interior a D . Entonces f es *diferenciable en p* sii existe una transformación lineal $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que

$$\lim_{v \rightarrow 0} \frac{f(p+v) - f(p) - Tv}{\|v\|} = 0.$$

La transformación lineal T se llama *diferencial* (o a veces, *derivada*) de f en p , y se denota $T = df|_p$ o simplemente $f'(p)$. A la resta $f(p+v) - f(p) - Tv$ se le llama *resto* (de primer orden) de f en p . Si $m = 1$, entonces $Tv = \nabla f(p) \cdot v$.

En general, tenemos el siguiente

Teorema 6.26. $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_m)^t$, p interior a D .

- 1) f diferenciable en p sii f_i diferenciable en p , $i = 1, \dots, m$
- 2) En las condiciones de 1), la transformación $T = df|_p$ tiene la matriz

$$J_{fp} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix}$$

en las bases canónicas de \mathbb{R}^n y \mathbb{R}^m .

Observación: La matriz $m \times n$ J_{fp} que tiene por filas a los gradientes $\nabla f_1(p), \dots, \nabla f_m(p)$, se llama *matriz Jacobiana* de f en p .

Demostración. 1) Si f es diferenciable en p ,

$$f(p+v) = f(p) + T(v) + r(v), \quad \text{con } T \text{ lineal y } \frac{r(v)}{\|v\|} \xrightarrow{v \rightarrow 0} 0.$$

La igualdad anterior y el límite son sobre vectores de \mathbb{R}^m . Considerando la coordenada i -ésima obtenemos:

$$f_i(p+v) = f_i(p) + T_i(v) + r_i(v), \quad \text{con } T_i \text{ lineal y } \frac{r_i(v)}{\|v\|} \xrightarrow{v \rightarrow 0} 0. \quad (*)$$

que es la definición de diferenciabilidad para f_i , ya que la coordenada i -ésima de una función lineal es lineal también.

El recíproco es similar.

2) La ecuación (*) nos dice de $df_i|_p(v) = T_i(v)$, coordenada i -ésima de $T(v)$.

Como sabemos $df_i|_p(v) = \nabla f_i(p) \cdot v$, pero

$$\nabla f_i(p) = \left(\frac{\partial f_i}{\partial x_1}(p), \dots, \frac{\partial f_i}{\partial x_n}(p) \right).$$

Por lo tanto $e_i^t \cdot T \cdot e_j = T_i \cdot e_j = \frac{\partial f_i}{\partial x_j}(p)$, donde e_i^t es el **vector coordenado i -ésimo** (que tiene un 1 en el lugar i -ésimo y el resto ceros) \square

Para el caso particular $m = n$, la matriz $J_f p$ es cuadrada. Podemos calcular su determinante $\det(J_f p)$, que se llama **Jacobiano** de f en p y se denota $\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}(p)$.

Definición 29. $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es de clase C^k sii cada una de sus funciones coordenadas f_i lo es, o sea tiene derivadas parciales continuas hasta orden k .

Del teorema anterior y el Teorema 6.13, deducimos que si f es de clase C^k , es diferenciable.

Teorema 6.27 (Regla de la cadena). Sean $D \subset \mathbb{R}^n, D' \subset \mathbb{R}^m$,

$f : D \rightarrow \mathbb{R}^m$, diferenciable en p , $f(p) \in D'$,

$g : D' \rightarrow \mathbb{R}^p$, diferenciable en $f(p)$, entonces

$g \circ f : D \rightarrow \mathbb{R}^p$, es diferenciable en p y además

$$d(g \circ f)|_p(v) = dg|_{f(p)} \circ df|_p,$$

o equivalentemente, $J_{g \circ f} p = J_g f(p) \cdot J_f p$ (producto de matrices).

Demostración. Daremos solo un esbozo de la demostración, ya que las ideas son las mismas que en el Teorema 6.14. Como f diferenciable en p ,

$$f(p+v) = f(p) + J_f p v + r_f(v) \quad \frac{r_f(v)}{\|v\|} \xrightarrow{v \rightarrow 0} 0$$

Aplicamos g a ambos lados, y usamos la diferenciabilidad de g en $q = f(p)$:

$$\begin{aligned} g \circ f(p+v) &= g\left(q + \overbrace{J_f p \cdot v + r_f(v)}^w\right) \\ &= g(q) + J_g q \cdot w + r_g(w), \quad \frac{r_g(w)}{\|w\|} \xrightarrow{w \rightarrow 0} 0 \end{aligned}$$

de donde

$$g \circ f(p+v) = g(q) + (J_g q) \cdot (J_f p) \cdot v + \underbrace{J_g q \cdot r_f(v) + r_g(w)}_{=r(v)}.$$

Como la aplicación $v \mapsto (J_g q) \cdot (J_f p) \cdot v$ es lineal en v (es la compuesta de $dg|_q$ y $df|_p$), para terminar basta demostrar que $r(v)/\|v\| \xrightarrow{v \rightarrow 0} 0$:

p262 de [5]

$$\frac{r(v)}{\|v\|} = J_g q \underbrace{\frac{r_f(v)}{\|v\|}}_{\rightarrow 0} + \frac{r_g(w)}{\|v\|}$$

El primer término tiende a 0 por ser el resto de f . Respecto al segundo, procedemos como en el Teorema 6.14, si $w = 0$, entonces $r_g(w) = 0$. Si $w \neq 0$, entonces, recordando que $w = J_f p \cdot v + r_f(v)$,

$$\frac{r_g(w)}{\|v\|} = \frac{r_g(w)}{\|w\|} \frac{\|J_f p \cdot v + r_f(v)\|}{\|v\|} = \underbrace{\frac{r_g(w)}{\|w\|}}_{\rightarrow 0} \underbrace{\left\| \underbrace{J_f p \cdot \frac{v}{\|v\|}}_{\text{acotado}} + \underbrace{\frac{r_f(v)}{\|v\|}}_{\rightarrow 0} \right\|}_{\text{acotado}}.$$

□

8.2.0. Nota: Un caso particular de la regla anterior es el del Teorema 6.15:

En ese caso, se tenía $f(x(u, v), y(u, v)) = g(u, v)$, que puede interpretarse como la composición de dos funciones $f(x, y)$ de \mathbb{R}^2 en \mathbb{R} y $h(u, v) = (x(u, v), y(u, v))$ de \mathbb{R}^2 en \mathbb{R}^2 . Las ecuaciones (1) y (2) dadas allí pueden reescribirse en forma matricial:

$$\begin{pmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$$

o sea $\nabla(f \circ h) = \nabla f \cdot Jh$;

3. Funciones Inversas e Implícitas

Si tenemos una función $f(x)$ de una variable, de clase C^1 con $f'(x_0) \neq 0$, sabemos (ver Capítulo 3 de [5]) que existe una función inversa $x = f^{-1}(y)$ definida en un entorno $y_0 = f(x_0)$. Además, f^{-1} es derivable en y_0 y $(f^{-1})'(y_0) = 1/f'(x_0)$.

Generalizamos lo anterior para f de \mathbb{R}^n en \mathbb{R}^m , diferenciable en p . La condición $f'(x_0) \neq 0$ anterior equivale a pedir que $f'(x_0)$ tenga inverso o que $df|_{x_0}$ sea una transformación lineal inversible.

Parece natural exigir ahora que J_{fp} sea una matriz $n \times n$ inversible, lo que equivale a decir que $df|_p$ es inversible. Tenemos el siguiente:

Teorema 6.28 (de la función inversa). $f : D(\subset \mathbb{R}^n) \rightarrow \mathbb{R}^n$ de clase C^k , p interior a D , J_{fp} inversible. Entonces existen abiertos U, V tales que $p \in U, V = f(U)$, f es inyectiva en U y su inversa $f^{-1} : V \rightarrow U$, $(f \circ f^{-1} = id_V, f^{-1} \circ f = id_U)$ es de clase C^k . Además,

$$J_{f(p)}f^{-1} = (J_{fp})^{-1}.$$

Demostración. No probaremos este teorema, ya que la demostración supera los

objetivos del curso (ver, por ejemplo, [6]). \square

8.3.0. Ejemplo: $f(x, y) = (e^x \cos y, e^x \sin y), f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$J_f(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}.$$

Inversible $\forall x, y$, ya que su determinante es $e^{2x} \neq 0$. Por lo tanto existe una inversa local en un entorno de cualquier punto. En $(0, 0)$, por ejemplo,

$$f(0, 0) = (1, 0), \quad J_{(0,0)}f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Existe entonces f^{-1} en un entorno de $(1, 0)$ de clase C^∞ (f lo es) y $J_{(1,0)}f^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

8.3.1. Nota: En el caso de una variable, si $f'(x) \neq 0$ en un intervalo, por ejemplo, puede definirse una inversa global en el recorrido (ver Teo. 15 en Cap. 3 de [5]).

Uno puede intentar extender esto al caso de \mathbb{R}^n . Sin embargo, aquí el resultado no es cierto: la función del Ejemplo 3 anterior, considerada en todo \mathbb{R}^2 , no admite inversa global porque no es inyectiva: $f(x, y) = f(x, y + 2kn) \quad \forall k \in \mathbb{Z}$.

Lo que ocurre es que en el caso de \mathbb{R} , se explota la monotonía de la función, concepto que, como ya dijimos, no tiene extensión a \mathbb{R}^n .

Come Corolario del Teorema 6.28, obtenemos un **teorema de la función implícita** para sistemas de ecuaciones. Sean

$$\begin{cases} f_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0, \\ \vdots \\ f_m(x_1, \dots, x_n, y_1, \dots, y_m) = 0, \end{cases}$$

m ecuaciones en las $n + m$ variables $x_1, \dots, x_n, y_1, \dots, y_m$.

Por comodidad, sea $X = (x_1, \dots, x_n)^t, Y = (y_1, \dots, y_m)^t, f = (f_1, \dots, f_m)^t,$

tenemos $F(X^t, Y^t) = 0$.

p265 de [5]

Observación 6.8.1. *A partir de ahora dejaremos a cargo del lector el uso de la transposición, ya que la notación recarga mucho la escritura. Así, por ejemplo en el caso anterior, escribiremos “ $F(X, Y) = 0$ ”, en lugar de “ $F(X^t, Y^t) = 0$ ”.*

Supongamos $f(X_0, Y_0) = 0$. Queremos, en un entorno de (X_0, Y_0) , obtener $f(X, Y) = 0 \iff Y = \varphi(X)$, con φ de \mathbb{R}^n en \mathbb{R}^m .

Para extender el Teorema 6.22 de la función implícita, pediremos f de clase C^1 . Necesitamos una hipótesis que juegue el papel de la condición $f_y(x_0, y_0) \neq 0$ que teníamos allí. Para motivarla, supongamos $f(X, Y) = AX + BY$, donde A matriz $m \times n$, B matriz $m \times m$. Es decir, f es lineal. Si queremos “despejar” Y del sistema $AX + BY = 0$, la condición natural es pedir B inversible, para tener $T = B^{-1}AX$.

En el caso general, f no será lineal pero su incremento $f(p + v) - f(p)$ se aproxima por el término lineal $J_f p v$, donde $p = (X_0, Y_0)$ y $v = (\Delta X, \Delta Y) = (\Delta x_1, \dots, \Delta x_n, \Delta y_1, \dots, \Delta y_m)$.

Si $f(p) = 0$, la ecuación $f(X, Y) = 0$, con $X = X_0 + \Delta X$, $Y = Y_0 + \Delta Y$, es

$$0 = f(X_0 + \Delta X, Y_0 + \Delta Y) \approx J_f p \cdot \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} = J_1 \Delta X + J_2 \Delta Y,$$

$$\text{donde } J_f p = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} & \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{pmatrix}$$

La hipótesis natural a exigir es entonces, J_2 inversible, o sea $\det J_2 \neq 0$.

p266 de [5]

Teorema 6.29 (Teorema de la función implícita). $f : D(\subset \mathbb{R}^{n+m}) \rightarrow \mathbb{R}^m$ de clase

C^k , $p = (X_0, Y_0)$ interior a D , $f(p) = 0$ y

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{pmatrix} \neq 0$$

Entonces, existen $W_1 \subset \mathbb{R}^n$, $W_2 \subset \mathbb{R}^m$, abiertos que contienen a X_0 y Y_0 respectivamente, y $\varphi : W_1 \rightarrow W_2$ de clase C^k tal que $\forall (X, Y) \in W_1 \times W_2$

$$f(X, Y) = 0 \iff Y = \varphi(X).$$

Demostración. La idea de la demostración es extender f a una función de \mathbb{R}^{n+m} en \mathbb{R}^{n+m} de Jacobiano no nulo, y aplicarle el teorema de la función inversa. Sea $F : D \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$, dada por $F(X, Y) = (X, f(X, Y))$

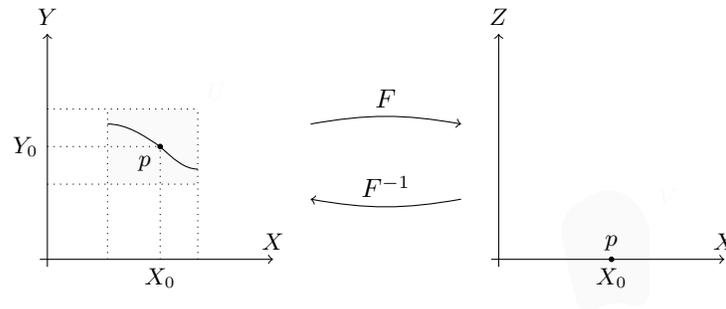
F es de clase C^k , $F(p) = (X_0, f(X_0, Y_0)) = (X_0, 0) = q$, y la matriz Jacobiana de F en p es

$$J_{Fp} = \begin{pmatrix} I & 0 \\ J_1 & J_2 \end{pmatrix}$$

donde I es la matriz identidad $n \times n$ y $[J_1 J_2] = J_{fp}$. Por lo tanto $\det J_{Fp} = \det J_2 \neq 0$, por hipótesis.

Entonces, por el Teorema 6.28, existe una función inversa $F^{-1} : V \rightarrow U$, de clase C^k , donde U, V abiertos que contienen p y q respectivamente (tomamos $U = W_1 \times W_2$).

Representamos la situación en el siguiente diagrama donde hemos identificado (X, Y) , $F(X, Y)$ con puntos de \mathbb{R}^2 , para poder tener una idea gráfica.



p268 de [5]

Aquí $F(X, Y) = (X, Z)$ con $Z = F(X, Y)$. Para $(X, Z) \in V$, tenemos $F^{-1}(X, Z) = (X, Y)$, y sea $Y = g(x, z)$ las últimas coordenadas de F^{-1} . Como g es de clase C^k , y para $(X, Y) \in W_1 \times W_2$, $(X, Z) \in U$, se cumple:

$$(X, Z) = F(X, Y) \iff (X, Y) = F^{-1}(X, Z),$$

o también

$$Z = f(X, Y) \iff Y = g(X, Z).$$

Tomando $Z = 0$, y definiendo $\varphi(X) = g(X, 0)$, se tiene para $(X, Y) \in W_1 \times W_2$, $f(X, Y) = 0$ si $Y = \varphi(X)$. \square

8.3.3. Notas:

- De imponer $f(X, \varphi(X)) \equiv 0$, puede deducirse por diferenciación, fórmulas para las derivadas de φ .

- Como aplicación del teorema anterior puede deducirse, en forma similar el Teorema 6.23, el método de multiplicadores de Lagrange para extremos con varias ligaduras.

p269 de [5]

Capítulo 7

INTEGRALES MÚLTIPLES

7.1. Integrales Dependientes de un Parámetro

Consideremos en esta sección, integrales de la forma

$$\int_c^d f(x, y) dy,$$

en las que la función $f(x, y)$ se integra respecto de una variable y la otra variable permanece constante, comportándose como un parámetro. Los extremos de integración pueden depender también del parámetro, en este caso x .

1. Continuidad

Sea en primera instancia $f(x, y)$ definida en el rectángulo $[a, b] \times [c, d]$.

Teorema 7.1. Si f continua en $[a, b] \times [c, d]$, entonces $F(x) = \int_c^d f(x, y) dy$ es continua en $[a, b]$.

Demostración. Sean $x, x + h \in [a, b]$, entonces (suponiendo $c < d$)

$$|F(x + h) - F(x)| = \left| \int_c^d f(x + h, y) - f(x, y) dy \right| \leq \int_c^d |f(x + h, y) - f(x, y)| dy.$$

Para acotar ésta última integral, es necesario acotar la diferencia $f(x+h, y) - f(x, y)$ para todo $y \in [c, d]$. Como $[a, b] \times [c, d]$ es compacto, sabemos por el Teorema 6.11 (de Cantor), que f es uniformemente continua,

$$\Rightarrow \exists \delta > 0 : \|(x_1, y_1) - (x_2, y_2)\| < \delta \Rightarrow |f(x_1, y_1) - f(x_2, y_2)| < \frac{\varepsilon}{d-c}.$$

Si $|h| < \delta \Rightarrow \|(x+h, y) - (x, y)\| < \delta \Rightarrow |f(x+h, y) - f(x, y)| < \frac{\varepsilon}{d-c} \Rightarrow |F(x+h, y) - F(x, y)| < \int_c^d \frac{\varepsilon}{d-c} dy = \varepsilon. \quad \square$

Permitiremos ahora que los extremos de integración dependan del parámetro, es decir $c = c(x), d = d(x)$.

Por simplicidad, supondremos f definida en un rectángulo $[a, b] \times [c, d]$, con $c(x), d(x) \in [a, b] \forall x$.

Lema 7.2. f continua en $[a, b] \times [c, d]$. Para $x \in [a, b], u, v \in [c, d]$, definimos

$$\phi(u, v, x) = \int_u^v f(x, y) dy$$

entonces ϕ es continua en $[c, d] \times [c, d] \times [a, b]$.

Demostración.

$$\phi(u, v, x) = \underbrace{\int_c^v f(x, y) dy}_{\psi(v, x)} - \underbrace{\int_c^u f(x, y) dy}_{\psi(u, x)}.$$

Basta probar que $\psi(v, x)$ es continua en $[a, b] \times [c, d]$:

$$\begin{aligned} \psi(v_0 + h, x_0 + k) - \psi(v_0, x_0) &= [\psi(v_0 + h, x_0 + k) - \psi(v_0, x_0 + k)] + [\psi(v_0, x_0 + k) - \psi(v_0, x_0)] = \\ &= \underbrace{\int_{v_0}^{v_0+h} f(x_0 + k, y) dy}_{(I)} + \underbrace{\int_c^{v_0} f(x_0 + k, y) - f(x_0, y) dy}_{(II)}. \end{aligned}$$

Por el teorema del valor medio para integrales $(I) = f(x_0 + k, v_0 + \theta h)h$ para

cierto $\theta \in (0; 1)$. Como f es acotada en $[a, b] \times [c, d]$ por ser continua, entonces (I) tiende a cero si $h \rightarrow 0$. Por otro lado, por el Teorema 7.1, (II) es continua como función de k , por lo tanto $\phi(k) \rightarrow \phi(0) = 0$ cuando $k \rightarrow 0$. \square

Teorema 7.3. Si f es continua en $[a, b] \times [c, d]$ y $c(x), d(x) : [a, b] \rightarrow [c, d]$ continuas.

Entonces,

$$F(x) = \int_{c(x)}^{d(x)} f(x, y) dy$$

es continua en $[a, b]$.

Demostración. Si ϕ es como en el Lema 7.2, entonces $F(x) = \phi(c(x), d(x), x)$, de donde, por la continuidad de la composición de funciones continuas, F lo será. \square

1.1.1. Ejemplos

$$1) \int_0^1 y^x dy = \frac{y^{x+1}}{x+1} \Big|_0^1 = \frac{1}{x+1} \text{ para } x > -1.$$

$$2) \text{ Para } x \geq 0, \int_0^x y^x dy = \frac{x^{x+1}}{x+1}.$$

3) $F(x) = \int_0^1 e^{xy} dy = (e^x - 1)/x$ si $x \neq 0$ y 1 si $x = 0$. Está claro que F es continua en 0.

2. Derivadas

Estudiaremos ahora la existencia de derivadas respecto del parámetro x . Comenzamos con extremos fijos.

Teorema 7.4. Sean $f(x, y), f_x(x, y)$ continuas en $[a, b] \times [c, d]$, entonces $F(x) = \int_c^d f(x, y) dy$ es derivable en $[a, b]$ y

$$F'(x) = \int_c^d f_x(x, y) dy.$$

Demostración.

$$\frac{1}{h} [F(x+h) - F(x)] = \int_c^d \frac{f(x+h, y) - f(x, y)}{h} dy = \int_c^d f_x(x+\theta h, y) dy \quad 0 < \theta < 1,$$

donde θ existe por Lagrange. Por lo tanto

$$\left| \frac{F(x+h) - F(x)}{h} - \int_c^d f_x(x, y) dy \right| \leq \underbrace{\int_c^d |f_x(x+\theta h, y) - f_x(x, y)| dy}_{(I)}.$$

Por el Teorema 7.1, la función $\psi(z) = \int_c^d |f_x(x+z, y) - f_x(x, y)| dy$ es continua respecto de z , por lo tanto, cuando $h \rightarrow 0$, $\psi(\theta h) \rightarrow \psi(0) = 0$, de donde

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \int_c^d f_x(x, y) dy.$$

□

El teorema anterior nos dice que puede cambiarse el orden de la derivación respecto de x y de la integración respecto de y , siempre que $f_x(x, y)$ sea continua.

p273 de [5]

1.2.0. Ejemplo: 3) $F(x) = \int_0^1 y \sin(xy) dy \Rightarrow F'(x) = \int_0^1 y^2 \cos(xy) dy.$

Pasamos nuevamente al caso de extremos dependientes del parámetro,

$$F(x) = \int_{c(x)}^{d(x)} f(x, y) dy = \phi(c(x), d(x), x),$$

con ϕ como en el Lema 7.2. Para derivar $F(x)$, se utilizará la regla de la cadena.

Teorema 7.5. f, f_x continuas en $[a, b] \times [c, d]$, $c(x), d(x)$ derivables en (a, b) , con valores en $[c, d]$. Entonces

$$F(x) = \int_{c(x)}^{d(x)} f(x, y) dy$$

es derivable en $[a, b]$ con

$$F'(x) = \int_{c(x)}^{d(x)} f_x(x, y) + f(x, d(x))d'(x) - f(x, c(x))c'(x).$$

Demostración. Sea $\phi(u, v, x) = \int_u^v f(x, y) dy$.

Del Teorema Fundamental del Cálculo Integral se obtiene

$$\phi_v(u, v, x) = f(x, v), \quad \phi_u(u, v, x) = -f(x, u),$$

funciones continuas. Del Teorema 7.4 (u, v fijos) se obtiene

$$\phi_x(u, v, x) = \int_u^v f_x(x, y) dy$$

continua por Lema 7.2. Entonces ϕ tiene derivadas parciales continuas y podemos aplicar la regla de la cadena para obtener

$$F'(x) = \phi_u(c(x), d(x), x).c'(x) + \phi_v(c(x), d(x), x).d'(x) + \phi_x(c(x), d(x), x).$$

Sustituyendo, obtenemos la tesis. □

1.2.1. Ejemplo: 4) $F(x) = \int_0^x \sin(xy) dy \Rightarrow F'(x) = \int_0^x y \cos(xy) dy + \sin x^2$.

p274 de [5]

3. Integrales Iteradas

Sea f continua en $[a, b] \times [c, d]$.

Como $F(x) = \int_c^d f(x, y) dy$ es continua en $[a, b]$, podemos calcular la integral

$$\int_a^b F(x) dx = \int_a^b dx \int_c^d f(x, y) dy,$$

llamada integral iterada de $f(x, y)$.

Se plantea inmediatamente la posibilidad de realizar las integrales en el

otro orden

$$\int_c^d dy \int_a^b f(x, y) dx.$$

Más adelante (Sección 7.3) veremos que si f es continua, ambas integrales coinciden.

La integral iterada también puede plantearse en el caso de extremos dependientes del parámetro:

$$\int_a^b dx \int_{c(x)}^{d(x)} f(x, y) dy.$$

Muchas veces, el conjunto $D = \{(x, y) / a \leq x \leq b, c(x) \leq y \leq d(x)\}$, puede escribirse en la forma

$$\{(x, y) / c \leq y \leq d, a(y) \leq x \leq b(y)\}.$$

y se escribe la integral iterada

$$\int_c^d dy \int_{a(y)}^{b(y)} f(x, y) dx.$$

p275 de [5]

1.3.0. Ejemplo: 5) Sea $D = \{(x, y) / 0 \leq x \leq 1, 0 \leq y \leq x\} = \{(x, y) / 0 \leq y \leq 1, y \leq x \leq 1\}$.

Sea $f(x, y) = y$. Calculemos las integrales iteradas

$$\int_0^1 dx \int_0^x y dy = \int_0^1 dx \left[\frac{y^2}{2} \right]_0^x = \int_0^1 \frac{x^2}{2} dx = \frac{1}{6}.$$

$$\int_0^1 dy \int_y^1 y dx = \int_0^1 dy y(1 - y) = \int_0^1 y - y^2 dy = \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.$$

En el ejemplo, las dos integrales iteradas obtenidas a partir del mismo dominio D coincidieron (con el volumen de una pirámide de base $1/2$ y altura 1).

Estamos entonces ante un número que depende exclusivamente de la función $f(x, y)$ y del dominio D . Esto nos lleva a pensar en una construcción más simétrica, en que las variables x e y jueguen un papel equivalente entre si.

Esto será abordado en lo que sigue, construyéndose la integral doble.

p276 de [5]

7.2. Medida de Jordan en el Plano

Nos proponemos a continuación, una extensión más completa de la integración a funciones de dos variables.

Para funciones de una variable, la idea consistía en considerar sumas de la forma $\sum f(\tilde{x}_i)\Delta x_i$, donde \tilde{x}_i era un punto del intervalo $[x_i, x_{i+1}]$ y Δx_i su longitud.

En funciones de dos variables, consideraremos en forma análoga sumas $\sum f(\tilde{x}_i, \tilde{y}_i) A(R_i)$, donde $(\tilde{x}_i, \tilde{y}_i)$ es un punto de la región R_i y $A(R_i)$ su área.

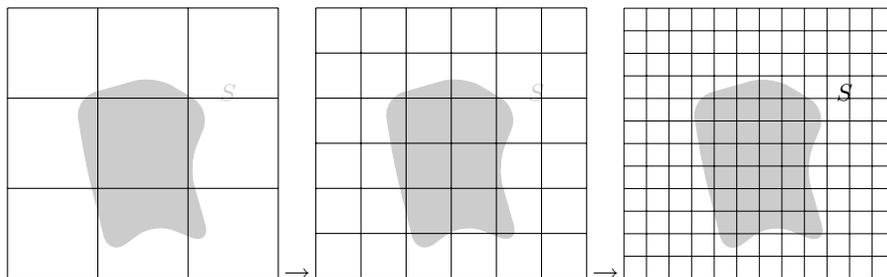
En el Capítulo 4 de [5], se da una definición de área para conjuntos particulares del plano. Nos proponemos ahora (previamente a la construcción de la integral doble) extender esa noción para conjuntos más generales del plano.

Intuitivamente, el área deberá cumplir:

- 1) El área de un rectángulo es el producto de base por altura.
- 2) El área es una propiedad aditiva. Si S y T son conjuntos disjuntos, $A(S \cup T) = A(S) + A(T)$.

Sea $S \subset \mathbb{R}^2$ un conjunto acotado, queremos definir su área $A(S)$. Un camino natural es considerar un "cuadriculado" del plano dado por los puntos de coordenadas enteras, como aproximación por exceso tomar $A_0^+(S)$, número de cuadrados que tienen algún punto en común con S , y como aproximación por defecto, tomar el número $A_0^-(S)$ de cuadrados totalmente contenidos en S .

En la primer figura, $A_0^+(S) = 9$ y $A_0^-(S) = 1$.



p277 de [5]

Luego, podemos dividir en 4 a cada cuadrado (puntos de coordenadas múltiplos de $1/2$) y mejorar las aproximaciones considerando $A_1^+(S) = 19/4$, área de los cuadrados que cortan a S (es decir el número de cuadrados que cortan a S por $1/4$), y $A_1^-(S) = 4/4$, área de los cuadrados contenidos en S . Repetimos el procedimiento, considerando los puntos de coordenadas múltiplos de $1/2^n$, es decir puntos $(i/2^n, j/2^n)$, con i, j enteros. Para la tercer figura $A_2^+(S) = 46/8$ y $A_2^-(S) = 20/8$.

Sea

$$R_{i,j}^n = \left[\frac{i}{2^n}, \frac{i+1}{2^n} \right] \times \left[\frac{j}{2^n}, \frac{j+1}{2^n} \right]$$

de área $1/2^{2n}$. Definimos entonces

$$A_n^+(S) = \sum_{i,j/R_{i,j}^n \cap S \neq \emptyset} \frac{1}{2^{2n}}, \quad A_n^-(S) = \sum_{i,j/R_{i,j}^n \subset S} \frac{1}{2^{2n}}.$$

Está claro que $A_n^+(S) = 2^{-2n} |\{(i,j) : R_{i,j}^n \cap S \neq \emptyset\}|$ y $A_n^-(S) = 2^{-2n} |\{(i,j) : R_{i,j}^n \subset S\}|$.

Proposición 7.6. *Se cumple:*

- 1) $A_n^-(S) \leq A_n^+(S)$.
- 2) Si $S \subset T$, $A_n^-(S) \leq A_n^-(T)$; $A_n^+(S) \leq A_n^+(T)$.
- 3) $A_n^-(S) \nearrow$ esto es $A_n^-(S) \leq A_{n+1}^-(S) \quad \forall n \geq 0$.
- 4) $A_n^+(S) \searrow$ esto es $A_n^+(S) \geq A_{n+1}^+(S) \quad \forall n \geq 0$.

Demostración. 1) y 2) son inmediatas y quedan como ejercicio.

Para 3), observar que si pasamos de n a $n+1$, cada $R_{i,j}^n$ que estaba contenido en S , se divide en cuatro cuadrados que sigue estando contenidos en S ; por lo

tanto, A_n^- no puede decrecer.

Para 4), si un cuadrado $R_{i,j}^n$ no corta a S , tampoco lo hacen los cuadrados resultantes; por lo tanto, $A_n^+(S)$ no puede decrecer. \square

Las propiedades 1), 3) y 4), nos llevan a definir

$$A^-(S) = \lim A_n^-(S) \quad A^+(S) = \lim A_n^+(S).$$

Se cumple que $A^-(S) \leq A^+(S)$.

Definición 30. Si $A^-(S) = A^+(S)$, decimos que S es *medible según Jordan* y definimos su *área* (o *medida de Jordan*) como ese número, que denotamos $A(S)$.

2.0.1. Observación: notar que S es medible sii $A_n^+(S) - A_n^-(S) \xrightarrow{n \rightarrow \infty} 0$.

2.0.2. Ejemplo: 6) Sea S el rectángulo $S = [a, b] \times [c, d]$. Fijado n , se pueden hallar, $\alpha, \beta, \gamma, \delta$ enteros tal que

$$\frac{\alpha}{2^n} \leq a < \frac{\alpha+1}{2^n}, \quad \frac{\beta}{2^n} \leq b < \frac{\beta+1}{2^n}, \quad \frac{\gamma}{2^n} \leq c < \frac{\gamma+1}{2^n}, \quad \frac{\delta}{2^n} \leq d < \frac{\delta+1}{2^n}.$$

Entonces,

$$A_n^-(s) = \frac{(\beta - \alpha - 1)}{2^n} \frac{(\delta - \gamma - 1)}{2^n} \geq \left(b - a - \frac{2}{2^n}\right) \left(d - c - \frac{2}{2^n}\right),$$

$$A_n^+(s) = \frac{(\beta - \alpha + 1)}{2^n} \frac{(\delta - \gamma + 1)}{2^n} \leq \left(b - a + \frac{2}{2^n}\right) \left(d - c + \frac{2}{2^n}\right).$$

Tomando límite con $n \rightarrow +\infty$,

$$\lim A_n^-(S) = \lim A_n^+(S) = (b - a)(d - c) = A(S).$$

La definición dada de área cumple entonces el primer requisito elemental que exigíamos al comenzar. Faltaría ver la propiedad de aditividad, en la que po-

demos flexibilizar un poco el requisito de que S y T sean disjuntos, con la siguiente definición.

Definición 31. Decimos que S y T *no se solapan* sii $\overset{\circ}{S} \cap \overset{\circ}{T} = \emptyset$.

Es decir, S y T no tienen puntos interiores comunes.

Teorema 7.7. S y T medibles Jordan, S y T no se solapan, entonces $S \cup T$ medible Jordan y $A(S \cup T) = A(S) + A(T)$.

Demostración. Sea $n \in \mathbb{N}$. Se cumple:

$$1) A_n^+(S \cup T) \leq A_n^+(S) + A_n^+(T)$$

Para ver esto, observar que todo $R_{i,j}^n$ que corta a $S \cup T$, aparece entre los que cortan a S y los que cortan a T ; la desigualdad se debe a que pueden existir $R_{i,j}^n$ que corten a ambos y aparecería dos veces en el segundo miembro.

$$2) A_n^-(S \cup T) \geq A(S) + A(T)$$

Los $R_{i,j}^n$ contenidos en S o en T , están contenidos en $S \cup T$. Además, un $R_{i,j}^n$ no puede estar contenido a la vez en S y T , porque entonces sería $\overset{\circ}{S} \cap \overset{\circ}{T} \neq \emptyset$, contradiciendo que S y T no se solapan. Pero sí podría estar contenido en $S \cup T$ sin estarlo en S ni en T .

Tomando límite:

$$A^-(S \cup T) \geq A(S) + A(T) \geq A^+(S \cup T).$$

Como $A^-(S \cup T) \leq A^+(S \cup T)$, se cumple la tesis. \square

Nuestra definición de área cumple entonces con los requisitos elementales que exigíamos al principio. Para ser consistente con nuestros estudios anteriores de áreas, veamos que el área de una región $\{(x, y) / x \in [a, b], 0 \leq y \leq f(x)\}$ vale $\int_a^b f(x) dx$, para f seccionalmente continua.

Lema 7.8. Sea $S \subset \mathbb{R}^2$ acotado, tal que $\forall \epsilon > 0$, existen R_ϵ, T_ϵ medibles Jordan con $R_\epsilon \subset S \subset T_\epsilon$, $A(T_\epsilon) - A(R_\epsilon) < \epsilon$. Entonces, S es medible Jordan.

Demostración. Sea $\epsilon > 0$ arbitrario. Elegimos R_ϵ, T_ϵ y se cumple

$$A_n^-(R_\epsilon) \leq A_n^-(S) \leq A^+(S) \leq A_n^+(T_\epsilon) \Rightarrow A_n^+(S) - A_n^-(S) \leq A_n^+(T_\epsilon) - A_n^-(T_\epsilon).$$

Como el 2^{do} miembro tiene límite menor que ϵ , se puede obtener $n_0/\forall n \geq n_0$,

$$0 \leq A_n^+(S) - A_n^-(S) < \epsilon.$$

Por definición, entonces, $A_n^+(S) - A_n^-(S) \xrightarrow{n \rightarrow \infty} 0$. □

Teorema 7.9. Sea $S = \{(x, y) / a \leq x \leq b, 0 \leq y \leq f(x)\}$, f integrable y no nula en $[a, b]$. Entonces, S es medible Jordan y

$$A(S) = \int_a^b f(x) dx$$

p281 de [5]

Demostración. Dado $\epsilon > 0$, considero una partición $P = \{x_0, \dots, x_n\}$ del intervalo $[a, b]$ tal que $S(f, P) - s(f, P) < \epsilon$.

$$s(f, P) = \sum_{i=0}^{m-1} m_i \Delta x_i \quad S(f, P) = \sum_{i=0}^{m-1} M_i \Delta x_i.$$

con $m_i = \inf\{f(x) : x \in [x_i, x_{i-1}]\}$, $M_i = \sup\{f(x) : x \in [x_i, x_{i-1}]\}$.

$$\text{Sea } R_\epsilon = \cup_{i=0}^{m-1} [x_i, x_{i-1}] \times [0, m_i], \quad T_\epsilon = \cup_{i=0}^{m-1} [x_i, x_{i-1}] \times [0, M_i].$$

Aplicando el Teorema 7.7 y el Ejemplo 7.2, tanto R_ϵ como T_ϵ , que son uniones de rectángulos, son medibles Jordan y sus áreas valen

$$A(R_\epsilon) = s(f, P), \quad A(T_\epsilon) = S(f, P).$$

Estamos por tanto en las hipótesis del Lema 7.8 y por lo tanto S es medible

Jordan. Además, $\forall \epsilon$, tenemos para la correspondiente partición,

$$s(f, P) \leq A(S) \leq S(f, P)$$

Como

$$s(f, P) \leq \int_a^b f(x) dx \leq S(f, P) \Rightarrow \left| A(S) - \int_a^b f(x) dx \right| < \epsilon.$$

Como ϵ es arbitrario, se cumple la igualdad de la tesis. \square

p282 de [5]

Nuestra nueva noción de área coincide entonces con la dada anteriormente para conjuntos como en el teorema anterior, pero permite, además, incluir muchos más conjuntos.

Una caracterización útil es que *un conjunto es medible según Jordan si y solo si $A(\partial S) = 0$* , siendo ∂S la frontera de S . Esto resulta de probar que

$$A_n^+(S) - A_n^-(S) \leq A_n^+(\partial S) \leq 9 [A_n^+(S) - A_n^-(S)],$$

cosa que no veremos aquí (ver [1]).

La caracterización anterior permite probar que *una región cuya frontera es la unión de un conjunto finito de curvas $y = f(x)$ o $x = f(y)$, es medible Jordan*.

Además, se prueba que *si S y T son medibles Jordan, también lo son $S \cup T$, $S \cap T$, $S \setminus T$* .

Por lo tanto, la clase de conjuntos medibles Jordan es lo suficientemente amplia para incluir la mayor parte de los conjuntos habituales del plano. Por un estudio más completo, ver [1]. Se puede mejorar la teoría de conjuntos medibles Jordan considerando los conjuntos **medibles Lebesgue**. Si bien dicha teoría excede los límites del presente curso, es interesante notar que los conjuntos medibles Lebesgue son aproximables tanto como se quiera por conjuntos medibles Jordan.

El siguiente, es un conjunto no medible Jordan, pero si medible Lebesgue.

2.0.3. Ejemplo: 7) Sea $S = \{(x, y) \in (0, 1) \times (0, 1), x, y, \text{ racionales}\}$. Es fácil observar que $\forall n, A_n^+(S) = 1, A_n^-(S) = 0$. Observar que $\partial S = [0, 1] \times [0, 1]$ no tiene área 0.

Otro ejemplo no trivial es el siguiente: Considere una sucesión a_n de términos positivos cuya suma $\sum a_n$ sea menor que 1. Luego considere la sucesión de puntos con coordenadas racionales $q_n \in [0, 1]^2$ incluidos en $[0, 1]^2$. El conjunto $\cup(B(q_n; a_n) \cap [0, 1]^2)$ no es medible Jordan, es medible Lebesgue, y su medida es positiva.

7.3. Integrales Dobles

Procedemos ahora a la construcción de la integral doble, en forma muy similar a las integrales simples. Sea $f(x, y)$ continua en un dominio D , que suponemos medible Jordan.

Damos las siguientes definiciones.

Definición 32. Si $S \subset \mathbb{R}^2$ es un conjunto acotado, llamamos *diámetro* de S al número $d(S) = \sup\{\|x - x'\|, x, x' \in S\}$

3.0.4. Ejemplo: 8): $S = (a, b) \times (c, d)$, entonces

$$d(S) = \sqrt{(c - a)^2 + (d - b)^2}$$

Definición 33. Sea $S \subset \mathbb{R}^2$, medible Jordan. Llamamos *partición* de S a una familia de conjuntos $P = \{S_1, \dots, S_n\}$ tal que

- 1) S_i medible Jordan $\forall i$,
- 2) Los S_i no se solapan ($\overset{\circ}{S}_i \cap \overset{\circ}{S}_j = \emptyset$, para $i \neq j$),
- 3) $\cup_{i=1}^n S_i = S$.

Llamamos *norma de la partición* al máximo diámetro:

$$|P| = \max_{i=1, \dots, n} d(S_i).$$

Si P y Q son particiones, y todo conjunto de P es unión de conjuntos de Q , decimos que Q es un *refinamiento* de P .

Un ejemplo típico de partición de un conjunto S es la obtenida intersecándolo con un “cuadrulado” como los utilizados en Sección 7.2. En ese caso al dividir en 4 cada cuadrado se obtiene un refinamiento de la partición.

La norma de una partición obtenida con este método es menor o igual que las diagonal de los cuadrados. Esto, muestra que dado S medible Jordan, existen particiones de S de norma arbitrariamente pequeña.

p284 de [5]

Definición 34. Sea $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ acotada, D medible Jordan, $P = \{D_1, \dots, D_n\}$ partición de D . A las sumas

$$s(f, P) = \sum_{i=0}^{n-1} m_i A(S_i), \quad S(f, P) = \sum_{i=0}^{n-1} M_i A(S_i).$$

con $m_i = \inf_{(x,y) \in S_i} f(x, y)$, $M_i = \sup_{(x,y) \in S_i} f(x)$ les llamamos *sumas inferior y superior*, respectivamente de f respecto a la partición P .

Proposición 7.10. a) Si Q es un refinamiento de P , entonces $s(f, P) \leq s(f, Q)$ y $S(f, P) \geq S(f, Q)$.

b) $s(f, P) \leq S(f, Q)$ para cualesquiera particiones P, Q .

Demostración. La prueba es muy similar a la utilizada en la Proposición 1 del Capítulo 4 de [5], las ligeras modificaciones se deben a que aquí le llamaremos elemento de la partición D_i a un conjunto, mientras que en una variable utilizábamos los extremos de los intervalos. Esta distinción no es para nada esencial y no afectará los resultados. Veamos los detalles.

a) Supongamos por simplicidad que Q refina a P , dividiendo el conjunto D_k de P en dos conjuntos D'_k y D''_k :

$$P = \{D_1, \dots, D_{k-1}, D_k, \dots, D_n\}, \quad Q = \{D_1, \dots, D_{k-1}, D'_k, D''_k, \dots, D_n\},$$

$$m'_k = \inf\{f(x, y) : (x, y) \in D'_k\} \geq m_k,$$

$$m''_k = \inf\{f(x, y) : (x, y) \in D''_k\} \geq m_k,$$

de donde $m'_k A(D'_k) + m''_k A(D''_k) \geq m_k A(D_k)$.

Agregando los restantes términos, $s(f, Q) \geq s(f, P)$.

b) Dadas dos particiones P y Q , se consigue un refinamiento común a P y Q , considerando las intersecciones (no vacías) de los conjuntos de P y Q . Si R es ese refinamiento,

$$s(f, P) \leq_{(a)} s(f, R) \leq S(f, R) \leq_{(a)} S(f, Q)$$

□

De b) deducimos que

$$\sup_P s(f, P) \leq \inf_P S(f, P).$$

La integral doble podrá definirse en el caso de igualdad. Para funciones continuas en dominios compactos, el siguiente lema (análogo al Lema 3, capítulo 4 de [5]) permite hacerlo.

Lema 7.11. *f continua en D compacto. Dado $\epsilon > 0, \exists \delta > 0, /|P| < \delta$ implica*

$$S(f, P) - s(f, P) < \epsilon.$$

Demostración. Si $A(D) = 0$, resultado es trivial. Si $A(D) > 0$, como f es unifor-

mamente continua, dado $\epsilon > 0$ puedo elegir $\delta > 0$ tal que

$$\|x - x'\| < \delta \Rightarrow |f(x) - f(x')| < \frac{\epsilon}{A(D)}.$$

$$\text{Sea } P = \{D_1, \dots, D_n\}, |P| < \delta, \text{ si } d(D_i) < \delta \Rightarrow M_i - m_i < \frac{\epsilon}{A(D)}$$

$$\Rightarrow S(f, P) - s(f, P) = \sum_{i=1}^n (M_i - m_i)A(D_i) < \frac{\epsilon}{A(D)} \sum_{i=1}^n A(D_i) = \epsilon.$$

□

Corolario 7.12. Si f es continua en D compacto, entonces $\sup_P s(f, P) = \inf_P S(f, P)$.

p286 de [5]

Definición 35. f continua en D compacto. Al número $I = \sup s(f, P) = \inf S(f, P)$, le llamamos *integral doble* de f , y la denotamos

$$I = \iint_D f(x, y) dx dy \quad \text{o} \quad I = \iint_D f.$$

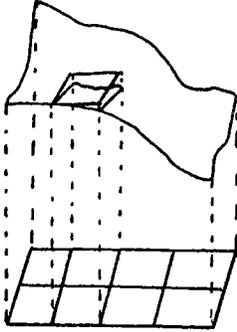
1. Interpretación Geométrica

Si $f(x, y) \geq 0$ representa una superficie en el espacio, las sumas superiores e inferiores son una aproximación al volumen del conjunto

$$\{f(x, y, z) / (x, y) \in D, 0 \leq z \leq f(x, y)\}.$$

La integral es entonces una forma de calcular el volumen comprendido entre el plano $z = 0$ y la superficie $f(x, y)$.

Tenemos también aquí una versión alternativa de la integral como límite de sumas $\sum f(\bar{x}_i, \bar{y}_i)A(D_i)$.



Definición 36. Un conjunto de puntos $X_P = \{(\bar{x}_i, \bar{y}_i), i = 1, \dots, n\}$ es *admisibile para una partición* P si $(\bar{x}_i, \bar{y}_i) \in D_i, \forall i$.

A la suma

$$SR(f, P, X_P) = \sum_{i=1}^n f(\bar{x}_i, \bar{y}_i)A(D_i),$$

le llamamos *suma de Riemann* de f asociada a P y X_P .

p287 de [5]

Teorema 7.13. f continua en D compacto. Dado $\epsilon > 0, \exists \delta > 0$ tal que:

$$|P| < \delta \Rightarrow |SR(f, P, X_P) - \iint_D f(x, y) dx dy| < \epsilon.$$

Demostración. Es idéntica al Teorema 4, capítulo 4 de [5] y queda como ejercicio. □

Corolario 7.14. f continua en D compacto, P_m sucesión de particiones tales que $|P_m| \xrightarrow{m} 0, X_m$ admisibile para P_m , entonces

$$SR(f, P_m, X_m) \xrightarrow{m} \iint_D f(x, y) dx dy.$$

2. Propiedades de la Integral Doble

1) Linealidad: (f, g continuas en D)

$$\iint_D \alpha f + \beta g = \alpha \iint_D f + \beta \iint_D g.$$

2) Aditividad respecto al dominio de integración: D, D' compactos, f continua en $D \cup D'$, D y D' no se solapan

$$\iint_{C \cup C'} f = \iint_C f + \iint_{C'} f.$$

p288 de [5]

3) $f \geq 0$ en $D \Rightarrow \iint_D f \geq 0$.

4) $f \geq g$ en $D \Rightarrow \iint_D f \geq \iint_D g$.

5) $|\iint_D f| \leq \iint_D |f|$.

Las demostraciones se deducen del Corolario 7.14 en forma similar al capítulo 4 de [5] y quedan como ejercicio. Para 2), observar que si D y D' no se solapan, se obtiene una partición de $D \cup D'$ uniendo particiones de D y D' .

La integral doble introducida hasta aquí es una extensión bastante natural de la realizada en la Sección 4.1 de [5].

Una vez más, es posible definir la integral doble para funciones más generales que las continuas, llamando **integrables** a aquellas funciones tales que

$$\sup_P s(f, P) = \inf_P S(f, P).$$

Se puede ver, por ejemplo, que *son integrables las funciones acotadas cuyos puntos de discontinuidad formen un conjunto de área nula*. Se puede demostrar que *las funciones integrables son exactamente aquellas funciones acotadas tales que para cualquier $\epsilon > 0$ sus puntos de discontinuidad se pueden cubrir con una cantidad numerable de conjuntos cuyas áreas forman una serie con suma menor que ϵ* .

Un caso particular es

$$\chi_S(x, y) = \begin{cases} 1, & (x, y) \in S, \\ 0, & (x, y) \notin S, \end{cases}$$

para S medible Jordan. En este caso es fácil ver que $\iint_D \chi_S = A(S)$, para todo $D \supset S$.

Los puntos de discontinuidad de χ_S son los puntos de ∂S y forman un conjunto de área nula.

Nos limitaremos en lo que sigue a funciones continuas. Por una teoría más general de las funciones integrables ver [3]. Para continuar, precisamos métodos prácticos que nos permitan calcular integrales dobles sin tener que recurrir a límites de sumas.

En integrales simples, ese papel lo jugaba el cálculo de primitivas.

Aquí, veremos que es posible reducir el problema de cálculo a integrales iteradas, introducidas en la Sección 7.1, que son de hecho integrales simples.

En efecto, la integral doble en un determinado dominio $D = \{(x, y)/a \leq x \leq b, c(x) \leq y \leq d(x)\}$, es igual a la integral iterada $\int_a^b dx \int_{c(x)}^{d(x)} f(x, y) dy$, si f es continua. Esto, se verá a continuación.

Teorema 7.15 (Fubini). *f continua en $D = [a, b] \times [c, d] = D$. Entonces,*

$$\iint_D f(x, y) dx dy \stackrel{(1)}{=} \int_a^b dx \int_c^d f(x, y) dy = \int_c^d dy \int_a^b f(x, y) dx.$$

Demostración. Veamos la igualdad (1), la otra se demuestra en forma simétrica.

Sea $\varphi(x) = \int_c^d f(x, y) dy$.

Consideramos particiones $P_1 = \{x_0, \dots, x_n\}$ y $P_2 = \{y_0, \dots, y_n\}$ de los intervalos $[a, b]$ y $[c, d]$ respectivamente y formamos con ellas la partición $P = \{R_{i,j}, i = 0, \dots, n-1, j = 0, \dots, n-l\}$, del conjunto D , siendo $R_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$. Probaremos:

$$(*) \quad s(f, P) \leq s(\varphi, P_1) \leq S(\varphi, P_1) \leq S(f, P).$$

$$s(f, P) \leq \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} m_{ij} \Delta x_i \Delta y_j = \sum_{i=0}^{n-1} \Delta x_i \sum_{j=0}^{n-1} m_{ij} \Delta y_j$$

donde $m_{ij} = \min\{f(x, y) : x \in [x_i, x_{i+1}], y \in [y_j, y_{j+1}]\}$, $\Delta x_i = x_{i+1} - x_i$ y $\Delta y_j = y_{j+1} - y_j$.

Observemos que si $x_0 \in [x_i, x_{i+1}]$ y $g(y) = f(x_0, y)$ entonces

$$m_{ij} \leq \min\{g(y) : y \in [y_j, y_{j+1}]\}$$

de donde, si $\varphi(x) = \int_c^d f(x, y) dy$, tenemos

$$\sum_{j=0}^{n-1} m_{ij} \Delta y_j \leq s(g, P_2) \leq \varphi(x_0).$$

Tomando mínimo en $x \in [x_i, x_{i+1}]$,

$$\sum_{j=0}^{n-1} m_{ij} \Delta y_j \leq \min_{x \in [x_i, x_{i+1}]} \varphi(x).$$

p291 de [5]

$$s(f, P) \leq \sum_{i=0}^{n-1} \Delta x_i \min_{x \in [x_i, x_{i+1}]} \varphi(x) = s(\varphi, P_1).$$

Análogamente se prueba $S(f, P) \geq S(\varphi, P_1)$, de donde se deduce (*).

Si ahora tomamos sucesiones de particiones $P_1^{(n)}$ de $[a, b]$ y $P_2^{(n)}$ de $[c, d]$ tales que $|P_1^{(n)}| \xrightarrow{n} 0$ y $|P_2^{(n)}| \xrightarrow{n} 0$, entonces la partición de rectángulos $P^{(n)}$ correspondiente cumple con $|P^{(n)}| \xrightarrow{n} 0$.

Tomando límite en $s(f, P^{(n)}) \leq s(\varphi, P_1^{(n)}) \leq S(\varphi, P_1^{(n)}) \leq S(f, P)$, se deduce que

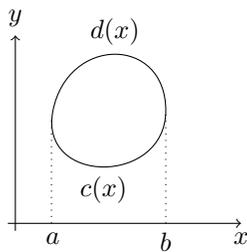
$$\int_D f(x, y) dx dy = \int_c^d \varphi(x) dx.$$

□

Es posible extender el teorema para una integral doble en una región $D = \{(x, y) / a \leq x \leq b, c(x) \leq y \leq d(x)\}$ donde $c(x), d(x)$ son continuas en $[a, b]$.

Obtenemos

$$\iint_D f(x, y) dx dy = \int_a^b dx \int_{c(x)}^{d(x)} f(x, y) dy.$$



No haremos aquí el detalle de la demostración, que puede hacerse, por ejemplo, aproximando la región D por una unión finita de rectángulos, aplicando el teorema anterior a cada uno y acotando la diferencia con la integral en D .

Si D puede además ponerse en la forma

$$D = \{(x, y) / c \leq y \leq d, a(y) \leq x \leq b(y)\}.$$

se tiene en forma simétrica

$$\iint_D f(x, y) dx dy = \int_c^d dy \int_{a(y)}^{b(y)} f(x, y) dx.$$

3.2.0. Ejemplo: 9) En el ejemplo 5), se probó que $\iint_D f(x, y) dx dy = 1/6$, siendo $D = \{(x, y) / 0 \leq x \leq 1, 0 \leq y \leq x\}$.

3.2.1. Ejemplo: 10) $D = [0, \pi] \times [0, \pi]$

$$\iint_D \cos(x+y) dx dy = \int_0^\pi dx \int_0^\pi \cos(x+y) dy dx = \int_0^\pi dx \sin(x+y)|_0^\pi =$$

$$= \int_0^{\pi} (\operatorname{sen}(\pi + x) - \operatorname{sen} x) dx = \int_0^{\pi} -2 \operatorname{sen} x dx = 2 \cos x \Big|_0^{\pi} = -4.$$

p293 de [5]

7.4. Cambios de Variable en Integrales Dobles

A la herramienta básica de cálculo de integrales doble (reducción a integrales iteradas) agregamos ahora el cambio de variables, que permite pasar de una integral doble a otra más conveniente.

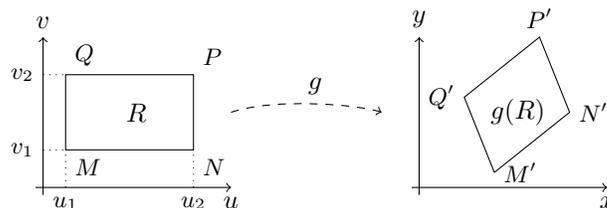
Si queremos calcular $\iint_D f(x, y) dx dy$, un cambio de variable es una función $(x, y) = g(u, v)$, $g : U \rightarrow V$, tal que V contiene al dominio de integración. Exigiremos además que g sea biyectiva y cumpla condiciones técnicas.

Definición 37. U, V , abiertos en el plano, $g : U \rightarrow V$ es un *cambio de variable (difeomorfismo)* de U a V sii g es biyectiva, diferenciable y $\det J_g a \neq 0, \forall a \in U$. Donde $J_g a$ es la jacobiana de g en a .

Del Teorema de la Función Inversa (Cap. 6, Teorema 6.28) deducimos que existe una función inversa $g^{-1} : V \rightarrow U$, que es también diferenciable y de Jacobiano no nulo.

El ejemplo más sencillo es el cambio lineal, $g(u, v) = (\alpha u + \beta v, \gamma u + \delta v) + (x_0, y_0)$, donde la matriz $T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ tiene determinante no nulo.

Veamos como se transforman las áreas por un cambio lineal. Sea $R = [u_1, u_2] \times [v_1, v_2]$ rectángulo en el plano (u, v) . Por ser lineal, g transforma a R en un paralelogramo $g(R)$.



Aquí $M' = g(M)$, etc.

p294 de [5]

Calculamos el área del paralelogramo $g(R)$ utilizando la fórmula (ver curso de Algebra Lineal)

$$A(g(R)) = |(N' - M') \wedge (Q' - M')|$$

en que \wedge denota el producto vectorial.

Se tiene

$$M' = (\alpha u_1 + \beta v_1, \gamma u_1 + \delta v_1) + (x_0, y_0),$$

$$N' = (\alpha u_2 + \beta v_1, \gamma u_2 + \delta v_1) + (x_0, y_0),$$

$$Q' = (\alpha u_1 + \beta v_2, \gamma u_1 + \delta v_2) + (x_0, y_0),$$

de donde

$$A(g(R)) = |(\alpha \Delta u, \gamma \Delta u) \wedge (\beta \Delta u, \delta \Delta v)| = |(\alpha \delta - \beta \gamma)| \Delta u \Delta v = |\det T| \Delta u \Delta v.$$

siendo $\Delta u = u_2 - u_1$, $\Delta v = v_2 - v_1$.

En resumen, $A(g(R)) = |\det T| A(R)$, es decir que las áreas en un cambio lineal se transforman multiplicando por el valor absoluto del determinante de T , siendo T la matriz de la transformación lineal (lo anterior lo muestra para rectángulos, pero por aditividad y paso al límite se extiende en forma sencilla para todo R medible Jordan).

Para calcular $\iint_D f(x, y) dx dy$, consideramos una partición

$P = \{D_1, \dots, D_n\}$ de D y formamos una suma de Riemann

$$\sum f(\bar{x}_i, \bar{y}_i) A(D_i) = \sum f(g(\bar{u}_i, \bar{v}_i)) |\det T| A(g^{-1}(D_i))$$

donde $(\bar{u}_i, \bar{v}_i) = g^{-1}(\bar{x}_i, \bar{y}_i)$.

$g^{-1}(P) = \{g^{-1}(D_i), D_i \in P\}$ es una partición de $g^{-1}(D)$, como surge de la continuidad de g , g^{-1} . Por lo tanto, el segundo miembro es una suma de

Riemann de $f(g(u, v))|\det T|$.

No es difícil ver que si $|P| \rightarrow 0$, la norma de la partición $g^{-1}(P)$ tiende a cero y tenemos

$$\iint_D f(x, y) dx dy = \iint_{g^{-1}(D)} f(g(u, v))|\det T| du dv$$

fórmula de cambio de variable para el caso de un cambio lineal.

Para una $g(u, v)$ que no sea lineal, la diferenciabilidad nos permite escribir

$$g(u, v) = g(\bar{u}_i, \bar{v}_i) + dg|_{(\bar{u}_i, \bar{v}_i)}(u - \bar{u}_i, v - \bar{v}_i) + r(u, v)$$

Es decir que g se aproxima localmente por una transformación lineal de matriz $T = J_g(\bar{u}_i, \bar{v}_i)g$.

Es natural pensar que al afinarse la partición, y hacerse $r(u, v)$ despreciable frente al término lineal, *las áreas tiendan a transformarse como en el caso lineal*. Es decir

$$A(D_i) \approx |\det(J_g(\bar{u}_i, \bar{v}_i))|A(g^{-1}(D_i))$$

si la partición es fina. Por lo tanto,

$$\sum f(\bar{x}_i, \bar{y}_i)A(D_i) \approx \sum f(g(\bar{x}_i, \bar{y}_i))|\det(J_g(\bar{u}_i, \bar{v}_i))|A(g^{-1}(D_i)).$$

Al tender la norma de la partición a cero, las consideraciones anteriores nos llevan a formular los siguiente.

Teorema 7.16. $f : D(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ continua, U, V , abiertos de \mathbb{R}^2 y $g : U \rightarrow V$ es un cambio de variable, $D \subset V$. Entonces

$$\iint_D f(x, y) dx dy = \iint_{g^{-1}(D)} f(g(u, v))|\det(J_g(u, v))| du dv$$

Demostración. La omitimos. Ver [1, 3]. □

Si escribimos $g(u, v) = (x(u, v), y(u, v))$ es más mnemotécnico usar la notación $\partial(x, y)/\partial(u, v)$ para el Jacobiano pues nos queda:

$$\iint_D f(x, y) dx dy = \iint_{g^{-1}(D)} f(g(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv,$$

fórmula que sugiere que los $du dv$ “se van” con los $\partial(u, v)$ del jacobiano para dejarnos los $dx dy$.

4.0.2. Ejemplos: 11) $D = \{(x, y) : x \geq 0, y \geq 0, x + y \leq 1\}$

Calculamos $\iint_D \exp(x - y)/(x + y) dx dy$.

Es natural escribir el cambio lineal $u = x - y, v = x + y$, o también $x = (u + v)/2, y = (-u + v)/2$.

Se obtiene $g^{-1}(D) = \{(u, v) : u + v \geq 0, v - u \geq 0, v \leq 1\}$.

La matriz jacobiana del cambio es $T = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}$ cuyo determinante es $1/2$. Por lo tanto,

$$\iint_D \exp\left(\frac{x - y}{x + y}\right) dx dy = \iint_{g^{-1}(D)} \frac{1}{2} \exp\left(\frac{u}{v}\right) du dv = \frac{1}{2} \int_0^1 dv \int_{-v}^v -e^{u/v} du =$$

$$\frac{1}{2} \int_0^1 dv [ve^{u/v}]_{-v}^v = \frac{1}{2} \int_0^1 (e - e^{-1})v dv = \frac{1}{4}(e - e^{-1}) = \frac{1}{2} \sinh 1.$$

Nota: el Teorema 7.16 puede extenderse permitiendo que $J_g(u, v) = 0$ en un conjunto de área nula, e incluso que g no sea inyectiva en un conjunto de área nula.

4.0.3. Ejemplo: 12) Coordenadas polares

Consideramos el cambio $(x, y) = g(\rho, \varphi)$, en qué $x = \rho \cos \varphi$ y $y = \rho \sin \varphi$ con $\rho \geq 0, \varphi \in [-\pi, \pi]$. Se verifica que $J_g(\rho, \varphi) = \rho$. Se lo aplicamos a

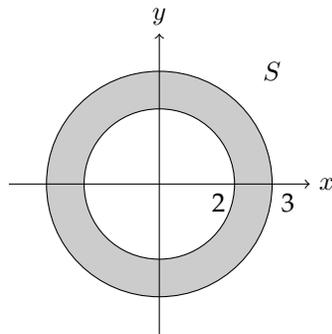
$$\iint_S (x^2 + y^2) dx dy$$

donde $S = \{(x, y) / 4 \leq x^2 + y^2 \leq 9\}$

$$\iint_S (x^2 + y^2) dx dy = \iint_{g^{-1}} (\rho^2) \rho d\rho d\varphi$$

con $g^{-1}(S) = [2, 3] \times [-\pi, \pi]$

$$\iint_S (x^2 + y^2) dx dy = \int_{-\pi}^{\pi} d\varphi \int_2^3 \rho^2 \rho d\rho = \int_{-\pi}^{\pi} d\varphi \int_2^3 \frac{\rho^4}{4} = 2\pi \left(\frac{3^4}{4} - \frac{2^4}{4} \right).$$



Aquí, el cambio g no es inyectivo en $\varphi = \pi, \varphi = -\pi$, y J_g se anula en $\rho = 0$. Siendo estos conjuntos de área nula, es válido el cambio realizado.

p298 de [5]

7.5. Integrales Múltiples de Dimensión Mayor que

2

La extensión de lo anterior para funciones de más de dos variables es totalmente directa y no ofrece nuevas dificultades. Nos limitaremos a enunciar en forma resumida los resultados y a resolver algunos ejemplos.

El primer paso es definir la noción de medida de Jordan (o volumen) en el espacio n -dimensional. Para ello se consideran los **cubos de dimensión n** , determinados por los puntos de coordenadas múltiplo de $1/2^k$. Un tal **n -cubo**

es de la forma

$$R_{i_1, \dots, i_n}^k = \left[\frac{i_1}{2^k}, \frac{i_1 + 1}{2^k} \right] \times \left[\frac{i_2}{2^k}, \frac{i_2 + 1}{2^k} \right] \times \dots \times \left[\frac{i_n}{2^k}, \frac{i_n + 1}{2^k} \right]$$

con i_1, i_2, \dots, i_n enteros.

Dado $S \subset \mathbb{R}^n$, se considera

$$V_k^+(S) = \sum_{(i_1, \dots, i_n) / R_{i_1, \dots, i_n}^k \cap S \neq \emptyset} \left(\frac{1}{2^n} \right)^k$$

volúmen de los n -cubos que cortan a S y $V_k^-(S)$, volumen de los n -cubos que están contenidos en S .

Una vez más, S será medible Jordan sii

$$\lim_k V_k^-(S) = \lim_k V_k^+(S) = V(S)$$

y esta definición es aditiva.

Sea ahora $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, función acotada, D medible Jordan. Se define en forma análoga como en la Sección 7.3:

$P = \{D_1, \dots, D_n\}$, partición de D ($D = \cup_{i=1}^n D_i$, los D_i no se solapan), $|P| = \max d(D_i)$, donde el diámetro $d(S)$ se define como en Def. 32.

p299 de [5]

$$s(f, P) = \sum_{i=1}^n \inf_{x \in D_i} f(x) \cdot V(D_i),$$

$$S(f, P) = \sum_{i=1}^n \sup_{x \in D_i} f(x) \cdot V(D_i)$$

y

$$SR(f, P, X_P) = \sum_{i=1}^n f(\bar{x}_i) V(D_i), \text{ con } X_P = \{\bar{x}_1, \dots, \bar{x}_n\} \text{ tal que } \bar{x}_i \in D_i.$$

Si f es continua y D es compacto, se prueba aquí que $\sup_P s(F, P) = \inf_P S(f, P) = I$ y además que dado $\epsilon > 0$, $\exists \delta / |P| < \delta \Rightarrow |I - SR(f, P, X_P)| < \epsilon$.

Al número I así introducido se le llama integral múltiple de f en D y se denota

$$\iint_D \cdots \int f \quad \circ \quad \iint_D \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

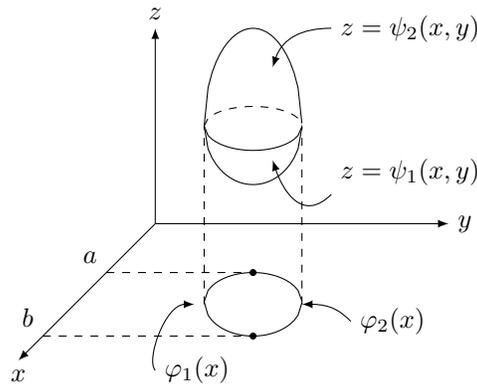
La herramienta básica de cálculo aquí también es la reducción a integrales iteradas. Si $D = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$, se obtiene para f continua

$$\iint_D \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_n}^{b_n} f(x_1, x_2, \dots, x_n) dx_n.$$

y el orden de la integral iterada es indiferente.

Lo anterior también es cierto para dominios más complicados. Veámoslo para el caso de una integral triple. Sea $D = \{(x, y, z) \in \mathbb{R}^3 / a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x), \psi_1(x, y) \leq z \leq \psi_2(x, y)\}$, donde $\varphi_1, \varphi_2, \psi_1, \psi_2$ son continuas.

p300 de [5]



Se tiene:

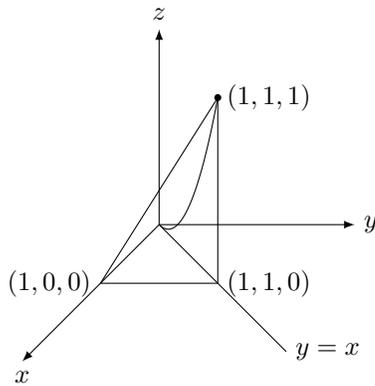
$$\iiint_D f(x, y, z) dx dy dz = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} dy \int_{\psi_1(x, y)}^{\psi_2(x, y)} f(x, y, z) dz$$

donde se resuelve en primer lugar la integral de la derecha, obteniendo una

función de x , y , etc.

5.0.4. Ejemplo: 13) Calcular $\iiint_D x y^2 z^3 dx dy dz$, siendo

$$D = \{(x, y, z) \in \mathbb{R}^3 / 0 \leq x \leq 1, 0 \leq y \leq x, 0 \leq z \leq xy\}$$



Se tiene:

p301 de [5]

$$\begin{aligned} \iiint_D x y^2 z^3 dx dy dz &= \int_0^1 \left[\int_0^x \left(\int_0^{xy} x y^2 z^3 dz \right) dy \right] dx = \\ &= \int_0^1 \left(\int_0^x x y^2 \frac{x^4 y^4}{4} dy \right) dx = \int_0^1 \frac{x^5}{4} \frac{x^7}{7} dx = \frac{1}{364}. \end{aligned}$$

Por último, tenemos también aquí, la posibilidad de un cambio de variables.

Un **cambio de variables** es una función $g : U \rightarrow V$, donde U, V son abiertos de \mathbb{R}^n , tal que g es biyectiva, diferenciable y $\det(J_g) \neq 0$ en U (o se anula en un conjunto de volumen nulo). Si $(x_1, \dots, x_n) = g(u_1, \dots, u_n)$ es un cambio de variables y $D \subset V$, se obtiene la siguiente fórmula de cambio de variables en la integral de una función f :

$$\iiint_D \dots \int f(x_1, \dots, x_n) dx_1 \dots dx_n = \iiint_{g^{-1}(D)} \dots \int f(g(x_1, \dots, x_n)) \left| \frac{\partial(x_1, \dots, x_n)}{\partial(u_1, \dots, u_n)} \right| du_1 \dots du_n.$$

Algunos cambios de variables importantes son los siguientes (para la elección

del cambio, tener en cuenta la forma de D y la expresión de f).

1) Cambio lineal Sea T matriz $n \times n$, con $\det T \neq 0$. La función $(x_1, \dots, x_n) = g(u_1, \dots, u_n)$ dada por $g(\vec{u}) = T\vec{u}$ es un cambio de variable en \mathbb{R}^n .

2) Coordenadas cilíndricas en \mathbb{R}^3 Tomemos $U = (0, +\infty) \times (0, 2\pi) \times \mathbb{R}$ y $g(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$. Se tiene $\det J_g(r, \theta, z) = r$.

3) Coordenadas esféricas en \mathbb{R}^3 $U = (0, +\infty) \times (0, 2\pi) \times (0, \pi)$ $g(\rho, \theta, \varphi) = (\rho \cos \theta \sin \varphi, \rho \sin \theta \sin \varphi, \rho \cos \varphi)$. En este caso se obtiene: $\det(J_g(\rho, \theta, \varphi)) = -\rho^2 \sin \varphi$.

5.0.5. Ejemplo: 14) $D = \{(x, y, z) / z \geq 0, x^2 + y^2 + z^2 \leq r^2\}$

Calculamos $\iiint_D z \, dx \, dy \, dz = I$

En coordenadas esféricas, $g^{-1}(D) = [0, r] \times [0, 2\pi] \times [0, \pi/2]$

$$\Rightarrow I = \int_0^{2\pi} d\theta \int_0^r d\rho \int_0^{\pi/2} \underbrace{\rho \cos \varphi}_z \underbrace{\rho^2 \sin \varphi}_{|\det J_g|} d\varphi = \int_0^{2\pi} d\theta \int_0^r \rho^3 d\rho \int_0^{\pi/2} \sin \varphi \cos \varphi d\varphi = (2\pi) \left(\frac{r^4}{4}\right) \frac{1}{2}.$$

Se verán más ejemplos en la sección siguiente.

7.6. Aplicaciones de las Integrales Múltiples

1) Áreas En el Capítulo 4 de [5], se calcularon áreas para regiones particulares por medio de integrales simples.

Para S medible Jordan cualquiera,

$$A(S) = \iint_S 1 \, dx \, dy$$

El cálculo de la integral doble mediante integrales iteradas nos conduce de todas formas a integrales simples, pero se debe observar que antes se pueden realizar cambios de variable.

6.0.6. Ejemplo: 15) Área del círculo $S = \{(x, y)/x^2 + y^2 \leq r^2\}$

$$\iint_S 1 \, dx \, dy = \int_0^{2\pi} d\varphi \int_0^r \rho \, d\rho = (2\pi) \frac{r^2}{2} = \pi r^2.$$

16) Área de la elipse $E = \{(x, y)/(x/a)^2 + (y/b)^2 \leq 1\}$. $A(E) = \iint_E 1 \, dx \, dy$.

Con el cambio $x = au$, $y = bv$ lineal, de Jacobiano ab , la preimagen de E es la Cfa. $u^2 + v^2 \leq 1$, lo que implica

$$A(E) = \iint_{u^2+v^2 \leq 1} ab \, du \, dv = a \cdot b \cdot \pi.$$

p304 de [5]

2) Volúmenes El volumen de una región S medible Jordan del espacio es

$$V(S) = \iiint_S 1 \, dx \, dy \, dz.$$

Si S es de la forma $S = \{(x, y, z)/(x, y) \in D, 0 \leq z \leq f(x, y)\}$, se lo puede calcular como una integral doble

$$V(S) = \iint_D f(x, y) \, dx \, dy$$

6.0.7. Ejemplos: 17) Volumen de la esfera $x^2 + y^2 + z^2 \leq R^2$ (S). En coordenadas esféricas, la esfera, es el conjunto $0 \leq \varphi \leq \pi, 0 \leq \theta \leq 2\pi, 0 \leq \rho \leq R$. El Jacobiano es $-\rho^2 \sin \varphi$, por lo tanto

$$\begin{aligned} V &= \iiint_V 1 \, dx \, dy \, dz = \int_0^{2\pi} d\theta \int_0^\pi d\varphi \int_0^R \rho^2 \sin \varphi \, d\rho = \int_0^{2\pi} d\theta \int_0^\pi \sin \varphi \, d\varphi \int_0^R \rho^2 \, d\rho \\ &= 2\pi \cdot 2 \cdot \frac{R^3}{3} = \frac{4}{3} \pi R^3. \end{aligned}$$

18) Volúmenes de revolución

En el Capítulo 4 de [5] se obtuvieron fórmulas para los volúmenes de revo-

p305 de [5]

lución mediante argumentos directos. Los reencontramos ahora utilizando un cambio a coordenadas cilíndricas.

Si S es el conjunto engendrado por la rotación de $\{(y, z)/z \in [a, b], 0 \leq y \leq f(z)\}$ en torno a Oz

$$V(S) = \iiint_S 1 \, dx \, dy \, dz = \int_0^{2\pi} d\theta \int_a^b dz \int_0^{f(z)} r \, dr = 2\pi \int_a^b \frac{f^2(z)}{2} \, dz.$$

6.0.8. Ejercicio: Si S' se obtiene girando el conjunto $\{(y, z)/y \in [a, b], 0 \leq z \leq f(y)\}$ en torno a Oz , verificar que

$$V(S') = 2\pi \int_a^b r f(r) \, dr.$$

3) Masa de un cuerpo

En Física, se asocia una magnitud llamada **masa** (m) a una porción de materia que ocupa una región S del espacio. Se denomina **densidad media** (ρ) de la sustancia al cociente entre la masa y el volumen $V(S)$. ($m = \rho V$)

Esta densidad media es suficiente en la medida en que la sustancia sea homogénea, es decir que los volúmenes iguales tengan la misma masa.

Una situación más general es considerar una partición $\{S_1, \dots, S_n\}$ de S , donde cada S_i tiene asociada una densidad ρ_i y la masa total se calcula como

p306 de [5]

$$m = \sum_{i=1}^n \rho_i V(S_i).$$

En el límite, podemos considerar que la densidad es una función $\rho(x, y, z)$ de cada punto y la masa se calcula como

$$m(S) = \iiint \rho(x, y, z) \, dx \, dy \, dz.$$

Por lo tanto, cualquier integral triple puede considerarse como una masa, en la que la función integrando se toma como la densidad $\rho(x, y, z)$ (Sin embargo, es

habitual considerar solo masas y densidades no negativas).

Por último, si $\rho(x, y, z)$ es continua, se puede obtener

$$(*) \quad \rho(x_0, y_0, z_0) = \lim_{d(\Delta S) \rightarrow 0} \frac{m(\Delta S)}{V(\Delta S)},$$

donde ΔS son regiones del espacio que contienen (x_0, y_0, z_0) y cuyo diámetro $d(\Delta S)$ tiende a 0.

En efecto,

$$\min_{(x,y,z) \in \Delta S} \rho(x, y, z) \leq \frac{m(\Delta S)}{V(\Delta S)} = \frac{1}{V(\Delta S)} \iiint_{\Delta S} \rho(x, y, z) dx dy dz = \max_{(x,y,z) \in \Delta S} \rho(x, y, z)$$

lo que nos da (*) por continuidad. El anterior es un proceso de “derivación en el espacio” que es análogo al que relaciona integrales y derivadas en el caso de integrales simples.

4) Centro de masa o baricentro

Supongamos dos partículas de masas m_1 y m_2 que ocupan posiciones $P = (x_1, y_1, z_1)$ y $P_2 = (x_2, y_2, z_2)$. Es natural definir un punto en que podamos concentrar la masa total $m_1 + m_2$, como

$$G = \frac{m_1}{m_1 + m_2} P_1 + \frac{m_2}{m_1 + m_2} P_2.$$

De esta forma, G es un punto situado en el segmento $P_1 P_2$ y sus distancias a los puntos P_1, P_2 son inversamente proporcionales a las masas respectivas.

G se llama baricentro o centro de masas y tiene aplicaciones en Mecánica.

La noción de centro de masa se extiende inmediatamente a un conjunto m_1, \dots, m_n de masas con puntos P_1, \dots, P_n , como

$$G = \frac{\sum m_i P_i}{m_i}.$$

Sus coordenadas son

$$x_G = \frac{\sum m_i x_i}{\sum m_i}, \quad y_G = \frac{\sum m_i y_i}{\sum m_i}, \quad z_G = \frac{\sum m_i z_i}{\sum m_i}.$$

Para el caso de una masa distribuida en una región S del espacio, con densidad $\rho(x, y, z)$, es natural definir G como el punto $G = (x_G, y_G, z_G)$, con

$$x_G = \frac{\iiint_S x \rho(x, y, z) dx dy dz}{\iiint_S \rho(x, y, z) dx dy dz}, \quad y_G = \frac{\iiint_S y \rho(x, y, z) dx dy dz}{\iiint_S \rho(x, y, z) dx dy dz}, \quad z_G = \frac{\iiint_S z \rho(x, y, z) dx dy dz}{\iiint_S \rho(x, y, z) dx dy dz}.$$

Para el caso homogéneo ($\rho(x, y, z) = \rho_0$, constante), G resulta independiente de ρ_0 y depende exclusivamente de la geometría de S .

6.0.9. Ejemplo: 19) Baricentro de una hemiesfera

p308 de [5]

homogénea:

$$S = \{(x, y, z) / z \geq 0, x^2 + y^2 + z^2 \leq r^2\}$$

Por argumentos de simetría es intuitivo que $x_G = y_G = 0$. (Verificarlo como ejercicio. Sugerencia: pasar a coordenadas cilíndricas.) Calculamos

$$z_G = \frac{\iiint_S z \rho(x, y, z) dx dy dz}{\iiint_S \rho(x, y, z) dx dy dz}.$$

Como $V(S) = (2/3)r^3\pi$ y el numerador fue calculado en el Ejemplo 14, entonces

$$z_G = \frac{\frac{\pi}{4}r^4}{\frac{2}{3}r^3\pi} = \frac{3}{8}r.$$

5) Momentos de inercia. El momento de inercia de una masa m situada en un punto $P = (x, y, z)$, respecto de una recta r se define como $I_r = md^2$, siendo d la distancia de P a r . En el caso en que r sea uno de los ejes coordenados, se tiene,

$$I_{Ox} = m(y^2 + z^2) \quad I_{Oy} = m(x^2 + z^2) \quad I_{Oz} = m(y^2 + x^2)$$

Si tenemos un conjunto de masas m_i en puntos $P_i = (x_i, y_i, z_i)$ definimos el momento de inercia total como la suma de los momentos de inercia de cada partícula.

$$I_{Ox} = \sum m_i(y_i^2 + z_i^2), \quad I_{Oy} = \sum m_i(x_i^2 + z_i^2), \quad I_{Oz} = \sum m_i(y_i^2 + x_i^2).$$

Para el caso de una masa distribuida en una región S del espacio, con densidad $\rho(x,y,z)$, definimos:

p309 de [5]

$$\begin{aligned} I_{Ox} &= \iiint_S (y_i^2 + z_i^2) \rho(x, y, z) \, dx \, dy \, dz, \\ I_{Oy} &= \iiint_S (x_i^2 + z_i^2) \rho(x, y, z) \, dx \, dy \, dz, \\ I_{Oz} &= \iiint_S (y_i^2 + x_i^2) \rho(x, y, z) \, dx \, dy \, dz. \end{aligned}$$

6.0.10. Ejemplo: 20) Momento de inercia de un cilindro homogéneo de masa M , radio R y altura h , respecto a su eje.

Sea $S = \{(x, y, z) / x^2 + y^2 \leq R^2, 0 \leq z \leq h\}$, con densidad ρ_0 .

$$I_{Oz} = \rho_0 \iiint_S (x^2 + y^2) \, dx \, dy \, dz = \rho_0 \int_0^{2\pi} d\theta \int_0^h dz \int_0^R (r^2) r \, dr = \rho_0 2\pi h \frac{R^4}{4} = \frac{\rho_0 \pi h R^4}{2}$$

Como el volumen $V(S) = R^2 \pi h$, se tiene $M = \rho_0 R^2 h$, de donde

$$I = \frac{MR^2}{2}.$$

p310 de [5]

7.7. Integrales Múltiples Impropias

La teoría anterior nos permite integrar funciones continuas en un dominio D compacto. Nos interesa ahora considerar extensiones de la definición a casos

en que D no sea compacto. Esto incluye dos casos particulares importantes: (a) el de dominios que no son acotados, como por ejemplo $D = \{(x, y)/x \geq 0, y \geq 0\}$ y (b) el de puntos de discontinuidad infinita, en que la función es continua en todo un compacto salvo en un punto, como por ejemplo $D = \{(x, y)/x^2 + y^2 \leq 1\} \setminus \{(0, 0)\}$.

Ambos corresponden a las integrales de 1^{era.} y 2^{da.} especie, ya encontradas en el caso de una variable.

La idea es aquí, como en una variable, integrar sobre dominios compactos D_n que "crezcan a D_n " cuando $n \rightarrow +\infty$. Sin embargo, no aparece aquí, como en el caso de una variable, una forma natural de elegir los D_n . Por ejemplo, en el caso en que D sea primer cuadrante, podemos tomar $D = [0, n] \times [0, n]$ o $D'_n = \{(x, y)/x, y \geq 0, x^2 + y^2 \leq n^2\}$.

No parece haber motivos para preferir una u otra elección o incluso otras imaginables. Sin embargo, como veremos a continuación, la elección puede no ser indiferente.

7.0.11. Ejemplo: 21) Sea D el primer cuadrante, queremos estudiar

$$\iint_D \operatorname{sen}(x^2 + y^2) dx dy.$$

Sea $D_n = [0, n] \times [0, n]$, $I_n = \iint_{D_n} \operatorname{sen}(x^2 + y^2) dx dy$. Se tiene:

$$I_n = \int_0^n dx \int_0^n (\operatorname{sen} x^2 \cos y^2 + \operatorname{sen} y^2 \cos x^2) dy = 2 \left(\int_0^n \operatorname{sen} u^2 du \right) \left(\int_0^n \cos u^2 du \right).$$

Las integrales impropias $\int_0^n \operatorname{sen} u^2 du$ y $\int_0^n \cos u^2 du$ son convergentes (en el capítulo 4 de [5] se vió la primera, la segunda es similar). Por lo tanto

$$I_n \xrightarrow{n \rightarrow +\infty} 2 \int_0^{+\infty} \operatorname{sen} u^2 du \int_0^{+\infty} \cos u^2 du.$$

Tomamos ahora D'_n como el cuadrante de círculo de radio n y $I'_n = \iint_{D'_n} \operatorname{sen}(x^2 +$

$y^2) dx dy$. Pasando a coordenadas polares

$$I'_n = \int_0^{\pi/2} d\varphi \int_0^n \operatorname{sen} \rho^2 \rho d\rho = \frac{\pi}{2} \cdot [-\cos \rho^2]_0^n = \frac{\pi}{4}(1 - \cos n^2).$$

Pero I'_n no posee límite. Por lo tanto, para definir una integral impropia en el caso de varias variables seremos algo más exigentes que en una variable.

p312 de [5]

Definición 38. Sea $f : D(\subset \mathbb{R}^k) \rightarrow \mathbb{R}$, continua; D tal que existe una sucesión (D_n) de conjuntos compactos tales que $D_n \subset D_{n+1}$, $D = \cup_n D_n$ y todo conjunto compacto de D está contenido en algún D_n .

Decimos que

$$\iint_D \dots \int f(x_1, \dots, x_k) dx_1 \dots dx_k$$

es *convergente* sii el siguiente límite

$$\lim_n \iint_{D_n} \dots \int f(x_1, \dots, x_k) dx_1 \dots dx_k$$

existe y vale lo mismo para cualquier (D_n) en las condiciones de arriba.

Con esta definición, la integral impropia del Ejemplo 21 no es convergente. Aparentemente, con la Definición 38 sería difícil saber si una integral impropia converge, ya que haría falta probar con cualquier elección de los (D_n) . Para funciones no negativas, sin embargo, una elección es suficiente.

Teorema 7.17. Si $f : D \rightarrow \mathbb{R}$ es continua, D como en la Definición 38. Si $f \geq 0$ en D , entonces $\lim_n \iint \dots \int_{D_n} f$ no depende de la elección de la sucesión (D_n) y solo tiene dos posibilidades: o es finito o es $+\infty$.

Demostración. Claramente, $I_n = \iint \dots \int_{D_n} f$ es una sucesión creciente, por ser $D_n \subset D_{n+1}$ y $f \geq 0$. Por lo tanto, solo puede ser $I_n \rightarrow I$ finito o $I_n \rightarrow +\infty$.

(a) Supongamos que para una sucesión (D_n) particular, se cumple $\iint \dots \int_{D_n} f \rightarrow I$, finito. Sea ahora (D'_n) otra sucesión de conjuntos como en la Def. 38, y

p313 de [5]

$$I'_n = \iint \dots \int_{D'_n} f.$$

Sea m fijo. Como D'_n es compacto, existe (Def. 38) un primer $n = n(m)$ tal que $D'_n \subset D$, entonces $I'_m \leq I_n \leq I$. Por lo tanto $I'_m \leq I \forall m$ y como I'_n es creciente, se deduce que $I'_m \rightarrow I' \leq I$. Cambiando los papeles de D_n y D'_n , deducimos que $I \leq I'$, de donde $I = I'$.

(b) Por otra parte, si para una (D_n) particular, el límite es $+\infty$, deberá serlo en cualquier otra sucesión (D'_n) ya que de otro modo, la parte (a) nos conduce a un absurdo. \square

7.0.12. Ejemplo: 27) Sea $D = \mathbb{R}^2$, $f(x, y) = e^{-(x^2+y^2)}$.

Como la función es no negativa elegimos una sucesión, particular (D_n) , tomando D_n como el círculo de centro $(0, 0)$ y radio n .

$$\begin{aligned} \iint_{D_n} e^{-(x^2+y^2)} dx dy &= \int_0^{2\pi} d\varphi \int_0^n e^{-\rho^2} \rho d\rho = 2\pi \left[-\frac{e^{-\rho^2}}{2} \right]_0^n = \pi(1 - e^{-n^2}) \xrightarrow{n \rightarrow +\infty} \pi \\ &\Rightarrow \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy \rightarrow \pi. \end{aligned}$$

Si ahora tomo $D'_n = [-n, n] \times [-n, n]$, el Teorema 7.17 nos dice que

$$\lim_n \iint_{D'_n} e^{-(x^2+y^2)} dx dy = \pi.$$

Pero

$$\iint_{D'_n} e^{-(x^2+y^2)} dx dy = \int_{-n}^n dx \int_{-n}^n e^{-x^2} e^{y^2} dy = \int_{-n}^n e^{-x^2} dx \int_{-n}^n e^{y^2} dy = \left(\int_{-n}^n e^{-x^2} dx \right)^2.$$

Como $\int_{-n}^n e^{-x^2} dx$ es convergente, tomando límite tenemos

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \pi \Rightarrow \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Este resultado es bastante llamativo, ya que obtuvimos una expresión exacta para $\int_{-\infty}^{\infty} e^{-x^2} dx$ pese a que tu integrando no posee primitiva elemental.

7.0.13. Ejemplo: 23) $D = \{(x, y) / 0 < x^2 + y^2 \leq 1\}$. Estudiamos

$$\iint_D \frac{1}{(\sqrt{x^2 + y^2})^\alpha} dx dy.$$

Tomamos $D_n = \{(x, y) : \frac{1}{n} \leq x^2 + y^2 \leq 1\}$;

$$\iint_{D_n} \frac{1}{(\sqrt{x^2 + y^2})^\alpha} dx dy = \int_0^{2\pi} d\varphi \int_{1/n}^1 \frac{1}{\rho^\alpha} \rho d\rho \xrightarrow{n} 2\pi \int_0^1 \frac{1}{\rho^{\alpha-1}} d\rho.$$

p315 de [5]

Recordando la integral impropia de dimensión 1, tenemos que

$$\iint_D \frac{1}{(\sqrt{x^2 + y^2})^\alpha} dx dy \text{ converge sii } \alpha < 2.$$

7.0.14. Ejercicio: Sea $D = \{(x, y) : x^2 + y^2 \geq 1\}$ Mostrar que

$$\iint_D \frac{1}{(\sqrt{x^2 + y^2})^\alpha} dx dy \text{ converge sii } \alpha > 2.$$

Por último veamos que ocurre si el Integrand f no es de signo constante.

El estudio de la integral impropia de $|f|$ nos conduce, como en el caso de dimensión 1, a la noción de convergencia absoluta. Hay, sin embargo aquí, una diferencia importante.

Teorema 7.18. Si f es continua en D , entonces $\iint \dots \int_D f$ converge sii $\iint \dots \int_D |f|$ converge.

Omitimos la demostración, que puede verse en [3].

En el caso de una variable, la integral impropia podrá converger pero no hacerlo absolutamente. Aquí, sin embargo, la convergencia equivale a la convergencia absoluta.

p316 de [5]

El origen de esta diferencia está en la definición adoptada para la convergencia. Al exigir que el límite sea el mismo para cualquier elección de los D'_n , permitimos que las integrales de las partes positivas y negativas de f sean “sumadas en cualquier orden” y, aparece aquí un efecto similar al que aparecía en la reordenación de series, que no alteraba la suma, solo en el caso de convergencia absoluta.

Los dos teoremas nos dan entonces el procedimiento para estudiar las integrales impropias con signo cualquiera:

- En 1^{er} lugar, estudiar la convergencia de $\iint \dots \int_D |f|$ a través de una sucesión particular D_n .
- En el caso de convergencia, si se quiere hallar la integral, hacer $\lim_n \iint \dots \int_{D_n} f$ para esa sucesión D_n o cualquier otra.

Bibliografía

- [1] Courant, R. y John, F. *Introducción al Cálculo y al Análisis Matemático. Vol. 1* 7a edición, Limusa. 1988
- [2] Do Carmo, M. *Differential Geometry of Curves and Surfaces*. Prentice-Hall.
- [3] Kudriavtsev, L.D. *Curso de Análisis Matemático*. Mir. 1994
- [4] Lages Lima, E. *Curso de Analise*. Projeto Euclides. 1991
- [5] F. Pagannini, *Análisis Matemático 1*. 1991-1992. Oficina de Publicaciones del Centro de Estudiantes de Ingeniería.
- [6] Rudin, W. *Principios de Análisis Matemático*. McGraw-Hill. 1967.

Índice de Símbolos

- $\|x\|$, 8
 \mathring{A} , 10
 $[p, p + v]$, 44
 $\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}(p)$, 73
 $\frac{\partial f}{\partial \bar{V}}$, 25
 $\frac{\partial f}{\partial \bar{V}}(x_1, \dots, x_n)$, 26
 $\frac{\partial f}{\partial x_i}(x_1, \dots, x_n)$, 26
 $\frac{\partial f}{\partial x}(x, y)$, 24
 \bar{A} , 10
 \iint_D , 96
 $\nabla f(p)$, 29
 ∂A , 10

 A' , 12
 $A(S)$, 89
 $A^+(S)$, 89
 $A_n^+(S)$, 88
 $A^-(S)$, 89
 $A_n^-(S)$, 88
 A^c , 9
 A^{ext} , 10

 $B(p; r)$, 9
 $B^*(p; r)$, 9
 B_p , 9
 $B_p(r)$, 9

 C_i^k , 41

 $d(x, y)$, 8
 $d^2 f|_p(\Delta x, \Delta y)$, 41
 $d^3 f|_p(\Delta x, \Delta y)$, 41
 $d^k f|_p(\Delta x, \Delta y)$, 41
 df , 33
 $df|_p$, 72

 $E(p; r)$, 9
 e_i , 73

 $f''(p)$, 48
 $f_x(x, y)$, 24

 $J_f p$, 72

 $R_{i,j}^n$, 88

 $S(f, P)$, 95
 $s(f, P)$, 95

 x_k , 13
 $x_{k i}$, 13

Índice de Nombres

- n -cubo, 107
- área, 87, 89
- espacios métricos, 11
- bola
 - abierta, 9
 - reducida, 9
- Bolzano-Weierstrass, 17
- cambio de variable, 102
- centro, 9
- clausura, 10
- coeficiente binomial, 41
- complemento, 9
- completo, 18
- conjunto
 - abierto, 10
 - acotado, 12
 - cerrado, 12
 - compacto, 13
 - medible Lebesgue, 93
- continua, 19
 - uniformemente, 22
- cubos de dimensión n , 107
- curva, 34
- curvas de nivel, 36
- de la función inversa, 75
- derivada, 71
 - direccional, 25
 - parcial, 24
- desigualdad
 - triangular, 8
- diámetro, 94
- difeomorfismo, 102
- diferenciabilidad
 - Condición suficiente de, 30
- diferencial, 33, 71
 - de orden k , 41
- distancia, 8
- dominio, 18
- entorno
 - abierto, 9
 - reducido, 9
- extremo, 54
 - absoluto, 45
 - relativo, 45
- Extremo Relativo
 - Condición Necesaria de, 46
- forma cuadrática, 48
 - definida
 - negativa, 48
 - positiva, 48
 - indefinida, 48

- semidefinida
 - negativa, 48
 - positiva, 48
- función
 - continua
 - en un punto, 19, 70
 - de clase C^k , 41, 73
 - de clase C^∞ , 41
 - diferenciable, 27
 - en un punto, 71
 - holomorfa, 24
 - integrable, 99
- funciones coordenadas, 69
- integral doble, 96
- integral impropia convergente, 118
- isotermas, 37
- Jacobiana
 - matriz, 72
- Jacobiano, 73
- k -veces diferenciable, 42
- límite, 18
 - de f en p , 18
 - de una sucesión, 14
- Lagrange
 - Método del mutiplicador de, 62
 - multiplicador de, 62
 - multiplicadores de, 67
- máximo
 - absoluto, 45, 54
 - condicionado, 61
 - relativo, 45
 - condicionado, 61
- mínimo
 - absoluto, 54
 - condicionado, 61
 - relativo, 45
 - condicionado, 61
- mapas de relieve, 36
- matriz
 - Hessiana, 48
- matriz Jacobiana, 72
- medible según Jordan, 89
- medida de Jordan, 89
- norma, 8
- paraboloide hiperbólico, 46
- partición, 94
 - norma de una, 94
- punto
 - crítico, 46
 - de acumulación, 12
 - exterior, 10
 - frontera, 10
 - interior, 10
- punto de silla, 50
- punto silla, 52

- radio, 9
- refinamiento, 94
- Regla de la cadena, 34
- resto, 27, 72
- resto de Lagrange, 42

- silla de montar, 46
- solapar, 90
- subsucesión, 15
- sucesión, 13
 - convergente, 14
 - de Cauchy, 18
 - parcial, 15
- suma
 - de Riemann, 97
 - inferior, 95
 - superior, 95

- Teorema
 - de Cantor, 23
 - de la función implícita, 56, 76,
77
- tiende a, 18

- vector
 - gradiente, 29
 - coordenado i -ésimo, 73
- versor, 51

- Weierstrass, 23